

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

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

MEMOIRE DE MASTER
en
MATHEMATIQUES
Spécialité
Analyse mathématique et application

Thème
Estimation et représentation ARCH(∞) des modèle GARCH

Présenté par
AZRAK Imane

Devant le jury d'examen composé de :

Mohamed MORSLI	Professeur	UMMTO	Président
Abdelghani HAMAZ	M. Conf B	UMMTO	Rapporteur
Fazia KHELLAS	Professeur	UMMTO	Examinatrice
Farida ACHEMINE	M. Conf A	UMMTO	Examinatrice

Soutenu le 23 /09 / 2018

Remerciements

En préambule à ce mémoire, je souhaite adresser mes remerciements les plus sincères à mon encadreur Mr Abdelghani HAMAZ, pour l'orientation, la confiance, la patience qui ont constitué un apport considérable sans lequel ce travail n'aurait pas pu être mené au bon port. Qu'il trouve dans ce travail un hommage vivant à sa haute personnalité.

Mes vifs remerciements vont également aux membres du jury : Mr MORSLI, Mme KHELLAS et Mme ACHEMINE pour l'intérêt qu'ils ont porté à ma recherche en acceptant d'examiner mon travail et de l'enrichir par leurs propositions.

Je n'oublie pas mes très chers parents qui ont toujours été là pour moi, et qui m'ont donné un magnifique modèle de labeur et de persévérance, ainsi que leur contribution, leur soutien et leur patience.

Je remercie ma grande soeur Ryma, son mari Zakaria et leur petit ange Yahia, ainsi que mon petit frère Amine et ma tante Akila pour tout leur soutien inconditionnel et leur encouragement.

Enfin, j'adresse mes plus sincères remerciements à mes meilleures amies Lisa et Lynda et à tous mes proches, qui m'ont toujours encouragée au cours de la réalisation de ce mémoire.

Table des matières

Introduction	4
1 Généralités sur les modèles ARCH/GARCH	6
1.1 Introduction	6
1.2 Modèles Autorégressifs Conditionnellement Hétéroscédastiques ARCH	7
1.3 Modèle ARCH(1)	8
1.3.1 Propriétés des processus ARCH(1)	10
1.4 Modèles ARCH(q)	13
1.5 Modèle avec erreurs ARCH(q)	14
1.6 Les modèles ARCH Généralisées :GARCH	16
1.7 Modèle GARCH(p,q) faible	17
1.8 Etude de la stationnarité	18
1.9 Modèle GARCH(1,1)	18
1.10 Modèle GARCH(p,q)	22
1.11 Modèles FIGARCH et ARFIMA	23
1.11.1 Modèle FIGARCH	24
2 Estimation des paramètres ARCH/GARCH	31
2.1 Estimation des modèles ARCH par la méthode des moindres carrés	32
2.2 Méthode d'estimation du maximum de vraisemblance	35
2.2.1 Quasi-vraisemblance conditionnelle	36
2.2.2 Equations de vraisemblance	37
2.2.3 Propriétés asymptotiques de l'estimateur du QMV	38
3 Représentation ARCH(∞)	41
3.1 Introduction	41

Table des matières **3**

3.2	Représentation ARCH(∞)	42
3.2.1	Représentation d'un GARCH(p,q)	43
3.2.2	Réurrences liées à la représentation infinie de GARCH(p,q)	50
3.3	Méthode d'estimation et la normalité asymptotique	54
	Conclusion	60
	Bibliographie	61

Introduction

La très grande fréquence des données financières a favorisé l'étude des impacts de court et de long terme d'une brusque fluctuation sur la dynamique des séries. Elle met en évidence le concept de mémoire longue dans les séries financières et microéconomiques. Pendant de nombreuses années, le modèle de marche aléatoire a été considéré le meilleur prédicteur de la dynamique de la volatilité sur les marchés de change. Ce modèle présume qu'on ne peut pas battre systématiquement le marché : les marchés sont efficients.

L'étude de séries financières révèle une dépendance temporelle du risque qui disparaît souvent progressivement. Les recherches se sont orientées vers l'étude des dynamiques conditionnelles. Toutefois, ces travaux ne différencient pas la dépendance conditionnelle révélant de la moyenne de celle de la variance. Deux volets sont alors issus. Le premier étudie l'équation de la moyenne conditionnelle avec les modèles ARMA, ARIMA, AR-FIMA... Le second s'est focalisé sur la variance conditionnelle. Deux classes de modèles non linéaires se sont développées pour caractériser cette variance.

- La première classe se base sur les modèles d'hétéroscédasticité conditionnelle autoregressive (ARCH) de Engle (1982) et ses extensions, le modèle d'hétéroscédasticité conditionnelle autoregressif généralisé (GARCH) de Bollerslev (1986) et le modèle GARCH intégré de Engle et de Bollerslev (1986).
- La deuxième classe fait référence aux modèles de volatilité stochastique (SV) introduit par Taylor (1986).

Contrairement aux modèles économétriques standards qui supposent que la variance est toujours constante, les modèles ARCH supposent que la variance conditionnelle est une fonction des carrés des innovations passées. L'application de ces modèles sur les taux d'inflation montre que la volatilité change au cours du temps. Dans tous ces travaux, on suppose que plus l'innovation est lointaine dans le passé plus faible sera son impact sur la

variance conditionnelle : la valeur de la volatilité est une fonction décroissante de l'ordre de retard. La spécification linéaire de la variance conditionnelle semble saisir l'effet des erreurs passées ou de mémoire longue trouvé dans les travaux empiriques. Cette spécification ne pose pas de problèmes d'estimation particuliers. Toutefois, pour que la variance conditionnelle soit positive, des contraintes supplémentaires sur les paramètres du modèle sont nécessaires. Pour contourner cette dernière difficulté, Engle (1982, 1983) et Engle et Kraft (1983) proposent de fixer le nombre de retards à prendre en compte au lieu de travailler avec des ARCH infinis. En se basant sur les travaux faits sur l'équation de la moyenne conditionnelle, une extension généralisée des modèles ARCH a été développée : le modèle GARCH. Ce dernier tient compte non seulement de la volatilité courante exprimée par les carrés des résidus passés mais aussi de la volatilité passée, il présente ainsi une spécification plus flexible de la variance conditionnelle.

Toutefois, le modèle GARCH ne permet de modéliser que la dépendance temporelle de court terme : l'effet des chocs décroît exponentiellement dans le temps. Pour étendre la mémoire du modèle GARCH, Engle et Bollerslev (1986) introduisent une extension du modèle GARCH dite intégrée (IGARCH) qui présente une mémoire "explosive" : les effets des chocs sont persistants à l'infini. Compte tenu des propriétés opposées des modèles traditionnels : décroissance exponentielle de l'impact des chocs sur la volatilité des modèles GARCH par opposition leur persistance à l'infinie dans les modèles IGARCH, un troisième modèle plus flexible a été introduit. Sur le marché de change, Baillie, Bollerslev et Mikkelsen (1996) ont proposé une version ajustée des deux modèles : c'est le processus GARCH fractionnement intégré (FIGARCH). Ces modèles sont spécifiés par analogie avec les modèles ARFIMA dans l'équation de la moyenne.

Le but de notre travail est de fournir une introduction aux modèles ARCH, GARCH, IGARCH et FIGARCH, le plus souvent utilisés dans la modélisation des marchés financiers. A base de l'article de Berkes *etal* (2003), nous allons traiter en particulier l'estimation des paramètres et l'existence d'une représentation ARCH(∞).

Le travail s'organise comme suit : dans le chapitre 1, nous étudions les modèles ARCH, GARCH, IGARCH et FIGARCH. Dans le chapitre 2, nous présentons les méthodes d'estimations des paramètres GARCH en montrant la consistance et la normalité asymptotique. Dans le chapitre 3, nous introduisons la représentation ARCH(∞) d'un modèle GARCH(p,q), en se basant sur les conditions d'existence et d'unicité et la consistance et la normalité asymptotique de l'estimateur de quasi-maximum de vraisemblance des paramètres sous des conditions moins restrictives.

Chapitre 1

Généralités sur les modèles ARCH/GARCH

1.1 Introduction

L'approche de modélisation Autoregressifs Conditionnellement Hétéroscédastique (ARCH) et leur extension (GARCH) (ARCH généralisé), a été proposée pour prendre en compte des variances conditionnelles dépendant du temps. L'idée générale consiste donc à remettre en cause la propriété d'homoscédasticité que l'on retient généralement dans le cadre du modèle linéaire. Le principe général proposé par Engle (1982) consiste à supposer que la variance dépend de l'ensemble informationnel dont on dispose. Il propose une spécification ARCH(q) où le carré des perturbations suit un processus autorégressif d'ordre q . Les modèles ARCH sont donc des modèles autorégressifs conditionnellement hétéroscédastiques. Engle (1982) a donc proposer ces processus pour pallier aux insuffisances de la classe des représentations ARMA, notamment en ce qui concerne les séries financières qui présentent une volatilité (ou variabilité instantanée mesurée par la variance conditionnelle) fonction du temps et par des ajustements asymétriques. Ainsi, les modèles ARCH sont basés sur une paramétrisation endogène de la variance conditionnelle.

Il faut souligner que dans la famille des modèles ARCH, différents types de processus ont été développés par d'autres chercheurs. Dès lors on peut distinguer les modèles ARCH linéaires et les modèles ARCH non linéaires. Les modèles ARCH linéaires reposent sur une spécification quadratique de la variance conditionnelle ; on y trouve les modèles : ARCH(q), GARCH(p, q) (ARCH Généralisé) et IGARCH(p, q)(GARCH Intégré) introduit par Engle et Bollerslev (1986) ; tant dis que les modèles ARCH non linéaires sont présentés par des spécifications asymétriques des perturbation. Ce sont les modèles EGARCH(p, q) (GARCH

Exponentiels) introduit par Nelson(1991), TARCH(q) et TGARCH(p, q) "GARCH à seuil" (Threshold GARCH(p,q)),(Bresson et Pirote, Séries temporelles). Par contre, le modèle FIGARCH (GARCH Fractionnellement Intégré) récemment introduit par Baillie, Bollerslev et Mikkelsen (1996) (Baillie et al. 1996 par la suite) fournit une mesure directe de cette persistance à travers le paramètre d'intégration fractionnelle. Les premières applications aux taux de change nominaux (Baillie, Bollerslev et Mikkelsen, (1996), Tse (1998), Beine, Lecourt et Laurent 1999) ont montré l'intérêt de cette modélisation.

Dans ce chapitre nous présentons un rappel des principaux définitions des modèles ARCH/GARCH, ainsi que leurs propriétés en précisant les conditions de stationnarité. Nous étudions aussi, le modèle FIGARCH et ses propriétés.

1.2 Modèles Autorégressifs Conditionnellement Hétéroscédastiques ARCH

Engle (1982) propose une nouvelle classe de modèles autorégressifs conditionnellement hétéroscédastiques (ARCH) apte à capter le comportement de la volatilité dans le temps. Le modèle est formé de deux équations. La première met en relation le rendement et certaines variables qui l'expliquent et la seconde modélise la variance conditionnelle des résidus. Le principe proposé par Engle consiste à introduire une dynamique dans la détermination de la volatilité en supposant que la variance est conditionnelle aux informations dont nous disposons. Il avance une spécification ARCH (q) où le carré des innovations, c'est-à-dire la variance du terme d'erreur au temps t, dépend de l'importance des termes d'erreur au carré des q périodes passées. Le modèle ARCH (q) permet de générer des épisodes de volatilité importante suivis d'épisodes de volatilité plus faibles. Pour bien comprendre le processus ARCH, nous allons présenter ce processus tel qu'il a été introduit par Engel (1982).

$$Y_t = \beta X_t + \varepsilon_t$$

avec

$$\varepsilon_t | I_{t-1} \sim \mathcal{N}(0, h_t)$$

où le terme Y_t correspond aux variables expliquant les rendements, il peut être un modèle ARMA(p, q). L'expression $I_{t-1} = \sigma(X_{t-s})_{s \leq t}$ désigne la tribu engendrée par les X_{t-s} $s \leq t$. Dans la modélisation ARCH, le processus $\{\varepsilon_t\}$ peut s'écrire sous la forme :

$$\varepsilon_t = z_t h_t$$

$$\text{avec } h_t = \sqrt{\alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2}$$

où $\{z_t\}$ est une suite de variables aléatoires indépendantes et identiquement distribuées $(0,1)$, avec $\alpha_0 > 0$ et $\alpha_i \geq 0$.

1.3 Modèle ARCH(1)

Le cas particulier $q = 1$, correspond à un ARCH(1) caractérisé par la définition suivante :

Définition 1.1. Un processus $\{\varepsilon_t\}$ satisfait une représentation ARCH(1) si :

$$\varepsilon_t = z_t h_t \tag{1.1}$$

avec

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

et $\{z_t\}$ désigne une suite de variables aléatoires indépendantes et identiquement distribuées (*i.i.d.*) d'espérance nulle et de variance σ_z^2 et $\alpha_0 > 0$, $\alpha_1 \geq 0$.

La composante h_t désigne une variable qui, conditionnellement à l'ensemble d'information des valeurs passées de ε_t , i.e. à $I_{t-1} = \sigma\{\varepsilon_{t-1}, \varepsilon_{t-2}, \dots, \varepsilon_{t-j}, \dots\}$, est déterministe et positive. La variance conditionnel :

$$\begin{aligned} V(\varepsilon_t | I_{t-1}) &= V(z_t h_t | I_t) \\ &= h_t^2 V(z_t | I_t) \\ &= h_t^2 \sigma_z^2 \end{aligned}$$

$\Rightarrow h_t^2$ est la variance conditionnelle de ε_t .

Dans ce modèle, le processus $\{\varepsilon_t\}$ est caractérisé par des autocorrélations nulles $\mathbb{E}(\varepsilon_t \varepsilon_s) = 0$ pour $t \neq s$ ce qui signifie que les ε_t sont non corrélées dans le temps. En effet, $\{\varepsilon_t\}$ reste un bruit blanc mais faible.

Des résultats intéressantes sont établis, en considérant le processus autorégressif sur les ε_t^2 . Pour simplifier, on se limite au cas du ARCH(1). Dans ces conditions :

$$h_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 \Leftrightarrow \varepsilon_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + (\varepsilon_t^2 - h_t^2) \tag{1.2}$$

soit encore :

$$\varepsilon_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \mu_t$$

$$\varepsilon_t^2 = h_t^2 + \mu_t$$

où $\mu_t = (\varepsilon_t^2 - h_t^2)$

Et en ayant les informations disponibles jusqu'au temps $t - 1$:

$$\mathbb{E}(\mu_t | I_{t-1}) = 0$$

avec I_{t-1} c'est l'ensemble de l'information jusqu'à $t - 1$, en effet

$$\begin{aligned} \mathbb{E}(\mu_t / I_{t-1}) &= \mathbb{E}(\varepsilon_t^2 / I_{t-1}) - \mathbb{E}(h_t^2 / I_{t-1}) \\ &= V(\varepsilon_t / I_{t-1}) - \mathbb{E}(h_t^2 / I_{t-1}) \\ &= h_t^2 - h_t^2 \\ &= 0 \end{aligned}$$

μ_t est un processus d'innovation pour ε_t^2 . Ainsi, cette écriture précédente correspond à celle d'un processus AR(1) sur le carré ε_t^2 .

$$\varepsilon_t^2 = h_t^2 + \mu_t$$

$$\varepsilon_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \mu_t \tag{1.3}$$

On sait que ce processus ε_t^2 est stationnaire au second ordre si et seulement si $|\alpha_1| < 1$. Cela s'énonce clairement par le théorème suivant.

Théorème 1.1. *Si un processus $\{\varepsilon_t\}$ satisfait une représentation ARCH(1), alors $\{\varepsilon_t^2\}$ satisfait une représentation AR(1) telle que :*

$$\varepsilon_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \mu_t$$

où $\mu_t = \varepsilon_t^2 - h_t^2$ vérifiant $\mathbb{E}(\mu_t / I_{t-1}) = 0$ est un processus d'innovation pour ε_t^2 .

1.3.1 Propriétés des processus ARCH(1)

Propriété 1.1. *On peut noter que pour tout $s \geq 1$:*

$$\mathbb{E}(\varepsilon_t/I_{t-s}) = 0 \quad (1.4)$$

Preuve. Pour démontrer cela, utilisons la propriété des espérances itérées. On a :

$$\begin{aligned} \mathbb{E}(\varepsilon_t/I_{t-s}) &= \mathbb{E}(\mathbb{E}(\varepsilon_t/I_{t-1})/I_{t-s}) \\ &= \mathbb{E}(0/I_{t-s}) \\ &= 0 \quad \forall s > 1. \end{aligned}$$

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

Propriété 1.2. *La variance conditionnelle du processus $\{\varepsilon_t\}$ satisfaisant une représentation ARCH(1), où $\alpha_0 > 0$ et $0 < \alpha_1 < 1$ définie par l'équation $\varepsilon_t = z_t h_t$ est non constante dans le temps et vérifie :*

$$V(\varepsilon_t/I_{t-s}) = \alpha_0 \left[\frac{1 - \alpha_1^s}{1 - \alpha_1} \right] + \alpha_1^s \varepsilon_{t-s}^2 \quad \forall t. \quad (1.5)$$

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

Preuve. On sait que $\mathbb{E}(\varepsilon_t/I_{t-s}) = 0$ dès lors, $V(\varepsilon_t/I_{t-s}) = \mathbb{E}(\varepsilon_t^2/I_{t-s})$, en utilisant la relation (1.3) où μ_t est un bruit blanc faible.

Par itération successive, on a :

$$\begin{aligned} \varepsilon_t^2 &= \alpha_0 + \alpha_1(\alpha_0 + \alpha_1 \varepsilon_{t-2}^2 + \mu_{t-1}) + \mu_t \\ &= \alpha_0 + \alpha_1 \alpha_0 + \alpha_1^2 \varepsilon_{t-2}^2 + \alpha_1 \mu_{t-1} + \mu_t \\ &= \alpha_0 + \alpha_1 \alpha_0 + \alpha_1^2 (\alpha_0 + \alpha_1 \varepsilon_{t-3}^2 + \mu_{t-2}) + \alpha_1 \mu_{t-1} + \mu_t \\ &= \alpha_0 + \alpha_1 \alpha_0 + \alpha_1^2 \alpha_0 + \alpha_1^3 \varepsilon_{t-3}^2 + \alpha_1^2 \mu_{t-2} + \alpha_1 \mu_{t-1} + \mu_t \\ &= \alpha_0 + \alpha_1 \alpha_0 + \alpha_1^2 \alpha_0 + \alpha_1^3 \alpha_0 + \dots + \alpha_1^{s-1} \alpha_0 + \mu_t + \alpha_1 \mu_{t-1} + \alpha_1^2 \mu_{t-2} + \dots \\ &\quad + \alpha_1^{s-1} \mu_{t-s+1} + \alpha_1^s \varepsilon_{t-s}^2. \end{aligned}$$

On aura donc :

$$\varepsilon_t^2 = \alpha_0 [1 + \alpha_1 + \alpha_1^2 + \dots + \alpha_1^{s-1}] + \mu_t + \alpha_1 \mu_{t-1} + \alpha_1^2 \mu_{t-2} + \dots + \alpha_1^{s-1} \mu_{t-s+1} + \alpha_1^s \varepsilon_{t-s}^2$$

En considérant l'esperance conditionnelle de chacun de ces nombres, il s'ensuit,

$$\mathbb{E}(\varepsilon_t^2/I_{t-s}) = \alpha_0 \left[\frac{1 - \alpha_1^s}{1 - \alpha_1} \right] + \sum_{j=0}^{s-1} \alpha_1^j \mathbb{E}(\mu_{t-j}/I_{t-s}) + \alpha_1^s \mathbb{E}(\varepsilon_{t-s}^2/I_{t-s}).$$

Par définition du bruit blanc, on a :

$$\mathbb{E}(\mu_{t-j}/I_{t-s}) = 0 \forall j = 0, 1, \dots, s-1, \text{ et par définition } \mathbb{E}(\varepsilon_{t-s}^2/I_{t-s}) = \varepsilon_{t-s}^2,$$

on obtient ainsi la formule de la variance :

$$V(\varepsilon_t/I_{t-s}) = \alpha_0 \left[\frac{1 - \alpha_1^s}{1 - \alpha_1} \right] + \alpha_1^s \varepsilon_{t-s}^2, \quad \forall t$$

Lorsque s tend vers l'infini, ces variances conditionnelles convergent vers la variance non conditionnelle, et l'on retrouve alors la formule :

$$\begin{aligned} V(\varepsilon_t) &= \lim_{s \rightarrow \infty} V(\varepsilon_t/I_{t-s}) \\ &= \lim_{s \rightarrow \infty} \left(\alpha_0 \left[\frac{1 - \alpha_1^s}{1 - \alpha_1} \right] + \alpha_1^s \varepsilon_{t-s}^2 \right) \\ &= \frac{\alpha_0}{1 - \alpha_1}. \end{aligned}$$

car $\alpha_0 > 0$ et $0 < \alpha_1 < 1$

Le résultat selon lequel la variance non conditionnelle est définie par la relation

$$V(\varepsilon_t) = \frac{\alpha_0}{1 - \alpha_1}.$$

qui explique les contraintes sur les paramètres de la représentation ARCH $\alpha_0 > 0$ et $0 < \alpha_1 < 1$.

Propriété 1.3. *Les autocovariances conditionnelles du processus $\{\varepsilon_t\}$, ARCH (1), définies par l'équation (1.1) sont nulles*

$$\text{cov}(\varepsilon_t, \varepsilon_{t+k}/I_{t-s}) = 0, \quad \forall k, s \geq 1. \quad (1.6)$$

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

Preuve. Cette propriété s'obtient de la façon suivante :

$$\begin{aligned}
 \text{cov}(\varepsilon_t, \varepsilon_{t+k}/I_{t-s}) &= \mathbb{E}(\varepsilon_t \varepsilon_{t+k}/I_{t-s}) - \mathbb{E}(\varepsilon_t/I_{t-s})\mathbb{E}(\varepsilon_{t+k}/I_{t-s}) \\
 &= \mathbb{E}(\varepsilon_t \varepsilon_{t+k}/I_{t-s}) \\
 &= \mathbb{E}[\mathbb{E}(\varepsilon_t \varepsilon_{t+k}/I_{t+k-1})/I_{t-s}] \\
 &= \mathbb{E}[\varepsilon_t \mathbb{E}(\varepsilon_{t+k}/I_{t+k-1})/I_{t-s}] \\
 &= \mathbb{E}[\varepsilon_t \cdot 0/I_{t-s}] \\
 &= 0.
 \end{aligned}$$

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.

Propriété 1.4. i)- *Le moment conditionnel centré d'ordre quatre du processus $\{\varepsilon_t\}$ vérifie*

$$\mathbb{E}(\varepsilon_t^4/I_{t-1}) = 3(\alpha_0 + \alpha_1 \varepsilon_{t-1}^2)^2.$$

ii)- *Sous l'hypothèse $3\alpha_1^2 < 1$, le moment non conditionnel centré d'ordre quatre du processus $\{\varepsilon_t\}$ est égal à :*

$$\begin{aligned}
 \mathbb{E}(\varepsilon_t^4) &= 3 \left[\alpha_0^2 + \frac{2\alpha_1 \alpha_0^2}{1 - \alpha_1} + \alpha_1^2 \mathbb{E}(\varepsilon_{t-1}^4) \right] \\
 &= 3 \frac{\alpha_0^2 (1 + \alpha_1)}{(1 - 3\alpha_1^2)(1 - \alpha_1)}.
 \end{aligned}$$

iii)- *La kurtosis non conditionnelle associée au processus ARCH(1) est*

$$K_u = \frac{\mathbb{E}(\varepsilon_t^4)}{\mathbb{E}^2(\varepsilon_t^2)} = 3 \left[\frac{1 - \alpha_1^2}{1 - 3\alpha_1^2} \right] > 3.$$

Preuve. i)- On rappelle que si une variable aléatoire centrée X suit une loi normale centrée, alors

$$\mathbb{E}(X^4) = 3(V(X))^2 = 3(\mathbb{E}(X^2))^2$$

donc

$$\begin{aligned}
 \mathbb{E}(\varepsilon_t^4/I_{t-1}) &= 3(\mathbb{E}(\varepsilon_t^2/I_{t-1}))^2 \\
 &= 3(\alpha_0 + \alpha_1 \varepsilon_{t-1}^2)^2.
 \end{aligned}$$

ii)- Sous l'hypothèse $3\alpha_1^2 < 1$ on a :

$$\begin{aligned}
 \mathbb{E}(\varepsilon_t^4) &= \mathbb{E}[\mathbb{E}(\varepsilon_t^4/I_{t-1})] \\
 &= \mathbb{E}[3(\alpha_0 + \alpha_1\varepsilon_{t-1}^2)^2] \\
 &= 3\mathbb{E}[\alpha_0^2 + 2\alpha_0\alpha_1\varepsilon_{t-1}^2 + \alpha_1^2\varepsilon_{t-1}^4] \\
 &= 3[\alpha_0^2 + 2\alpha_0\alpha_1\mathbb{E}(\varepsilon_{t-1}^2) + \alpha_1^2\mathbb{E}(\varepsilon_{t-1}^4)] \\
 &= 3[\alpha_0^2 + 2\alpha_0\alpha_1\frac{\alpha_0}{1-\alpha_1} + \alpha_1^2\mathbb{E}(\varepsilon_{t-1}^4)] \\
 &= 3\frac{\alpha_0^2(1+\alpha_1)}{(1-3\alpha_1^2)(1-\alpha_1)}
 \end{aligned}$$

iii)- La kurtosis non conditionnelle associée au processus ARCH(1) est définie par :

$$K_u = \frac{\mathbb{E}(\varepsilon_t^4)}{\mathbb{E}^2(\varepsilon_t^2)}$$

et d'après les résultats obtenus précédemment on obtient :

$$\begin{aligned}
 K_u &= \frac{3\alpha_0^2(1+\alpha_1)}{(1-3\alpha_1)(1-\alpha_1)} \cdot \frac{(1-\alpha_1)^2}{\alpha_0^2} \\
 &= 3\frac{1-\alpha_1^2}{(1-3\alpha_1^2)} > 3.
 \end{aligned}$$

Toutes ces propriétés peuvent être généralisées dans le cas d'un processus ARCH(q). En effet, dans la section qui suit, nous montrons la représentation d'un modèle ARCH(q) et ces propriétés, ainsi que le modèle avec erreurs ARCH(q).

1.4 Modèles ARCH(q)

Définition 1.2. Un processus $\{\varepsilon_t\}$ satisfait une représentation ARCH(q) si : $\varepsilon_t = z_t h_t$, $\alpha_0 > 0$ et $\alpha_i \geq 0$

avec

$$h_t = \sqrt{\alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2}$$

où $\{z_t\}$ désigne un bruit blanc faible, tel que $\mathbb{E}(z_t) = 0$ et $V(z_t) = \sigma_z^2$

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) $\mathbb{E}(\varepsilon_t/I_{t-s}) = 0$ et la propriété de variance conditionnelle variable dans le temps puisque :

$$V(\varepsilon_t/I_{t-1}) = h_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2,$$

de plus

$$\varepsilon_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \mu_t.$$

$$\mathbb{E}(\mu_t/I_{t-1}) = 0.$$

1.5 Modèle avec erreurs ARCH(q)

On considère dorénavant le résidu d'un modèle linéaire. Prenons l'exemple d'un modèle linéaire autorégressif d'ordre p avec résidus de type ARCH (q) : $X_t = \alpha_1 X_{t-1} + \dots + \alpha_p X_{t-p} + \varepsilon_t$. où :

$$\varepsilon_t = z_t h_t$$

avec

$$h_t = \sqrt{\alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2}$$

où $\{z_t\}$ est une suite de variables aléatoires *i.i.d* (0,1).

On a donc un modèle qui décrit à la fois l'évolution de l'espérance conditionnelle et de 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) :

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

$$\varepsilon_t = z_t \sqrt{\alpha_0 + \alpha_1 \varepsilon_{t-1}^2}.$$

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

i) Propriété de différence de martingale : $\mathbb{E}(\varepsilon_t/I_{t-1}) = 0$ et de façon générale

$$\mathbb{E}(\varepsilon_t/I_{t-s}) = 0 \quad \forall s \geq 1 \tag{1.7}$$

ii) Variance conditionnelle dépendante du temps

$$V(\varepsilon_t/I_{t-s}) = \alpha_0 \left[\frac{1 - \alpha_1^s}{1 - \alpha_1} \right] + \alpha_1^s \varepsilon_{t-s}^2.$$

$$V(\varepsilon_t) = \left[\frac{\alpha_0}{1 - \alpha_1} \right]. \quad (1.8)$$

iii) La propriété d'orthogonalité implique que les corrélations conditionnelles sont nulles :

$$\text{cov}(\varepsilon_t, \varepsilon_{t+k}/I_{t-s}) = 0 \quad \forall k, s \geq 1.$$

iii) Sous l'hypothèse $3\alpha_1^2 < 1$, la distribution des résidus est leptokurtique puisque :

$$K_u = 3 \frac{1 - \alpha_1^2}{(1 - 3\alpha_1^2)} > 3.$$

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 ε_t n'est pas constante, la variance non conditionnelle est constante.

On peut, en outre, en déduire un certain nombre de conclusions quant aux processus $\{X_t\}$ lui même. On peut montrer tout d'abord, que l'espérance conditionnelle de X_t vérifie :

$$\mathbb{E}(X_t/I_{t-s}) = \delta + \alpha \mathbb{E}(X_{t-1}/I_{t-s}) \quad \forall s \geq 1.$$

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

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

En effet :

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

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 quelle dépend du processus ε_{t-s}^2 de la façon suivante.

Propriété 1.5. La variance conditionnelle du processus $AR(1)$ avec erreur ARCH (1), X_t , s'écrit :

$$V(X_t/I_{t-s}) = \frac{\delta}{1-\alpha_1} \left[\left(\frac{1-\alpha^{2s}}{1-\alpha^2} \right) - \alpha_1 \left(\frac{\alpha_1^s - \alpha^{2s}}{\alpha_1 - \alpha^2} \right) \right] + \alpha_1 \left[\frac{\alpha_1^s - \alpha^{2s}}{\alpha_1 - \alpha^2} \right] \varepsilon_{t-1}^2 \quad (1.9)$$

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

$$V(X_t/I_{t-1}) = \delta + \alpha_1 \varepsilon_{t-1}^2 \quad (1.10)$$

Preuve.

$$\begin{aligned} V(X_t/I_{t-s}) &= V(\delta \frac{1-\alpha^s}{1-\alpha} + \alpha^s X_{t-s} + \varepsilon_t + \alpha \varepsilon_{t-1} + \dots + \alpha^{s-1} \varepsilon_{t-s+1} / I_{t-s}) \\ &= V(\varepsilon_t / I_{t-s} + \alpha^2 V(\varepsilon_{t-1} / I_{t-s}) + \dots + \alpha^{2(s-1)} V(\varepsilon_{t-s+1} / I_{t-s})) \\ &= \sum_{j=0}^{s-1} \alpha^{2j} \left[\delta \frac{1-\alpha^{s-j}}{1-\alpha} + \alpha_1^{s-j} \varepsilon_{t-s}^2 \right] \\ &= \frac{\delta}{1-\alpha_1} \left[\left(\frac{1-\alpha^{2s}}{1-\alpha^2} \right) - \alpha_1 \left(\frac{\alpha_1^s - \alpha^{2s}}{\alpha_1 - \alpha^2} \right) \right] + \alpha_1 \left[\frac{\alpha_1^s - \alpha^{2s}}{\alpha_1 - \alpha^2} \right] \varepsilon_{t-1}^2. \end{aligned}$$

En conclusion, si l'on désire prévoir le processus $\{X_t\}$ dans le cas d'erreur ARCH(1), l'erreur de prévision à un horizon d'une période admet une variance $V(X_t/I_{t-1})$ qui varie dans le temps en fonction de la valeur de ε_{t-s}^2 .

Remarque 1.1. La variance d'une erreur de prévision sur un processus avec erreur ARCH dépend du temps.

$$V(X_t/I_{t-s}) = f(\varepsilon_{t-s}), \quad s \geq 1 \quad (1.11)$$

L'amplitude des intervalles de confiance associée à cette prévision n'est donc pas constante dans le temps.

1.6 Les modèles ARCH Généralisés :GARCH

Le modèle Autorégressif Conditionnellement Hétéroscédastique Généralisé (GARCH), est suggéré par Bollerslev (1986). Sa 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.

1.7 Modèle GARCH(p,q) faible

Définition 1.3. On dit que $\{\varepsilon_t\}$ est un processus GARCH(p,q) faible si ses deux premiers moments conditionnels existent et vérifient :

(i) $\mathbb{E}(\varepsilon_t/I_{t-s}) = 0, \quad \forall s \geq 1.$

(ii) il existe des constantes $\alpha_0 > 0$, $\alpha_i, i = 1, \dots, q$, et $\beta_j, j = 1, \dots, p$ telles que

$$V(\varepsilon_t/I_{t-s}) = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j h_{t-j}^2 \quad (1.12)$$

L'équation (1.12) peut être écrite de manière symbolique sous la forme plus compacte

$$h_t^2 = \alpha_0 + \alpha(L)\varepsilon_t^2 + \beta(L)h_t^2, \quad t \in \mathbb{Z} \quad (1.13)$$

où L est l'opérateur retard ($L^i \varepsilon_t^2 = \varepsilon_{t-i}^2$ et $L^i h_t^2 = h_{t-i}^2$ pour tout entier i), α et β sont des polynômes de degrés q et p :

$$\alpha(L) = \alpha_1 L + \alpha_2 L^2 + \dots + \alpha_q L^q = \sum_{i=1}^q \alpha_i L^i,$$

$$\beta(L) = \beta_1 L + \beta_2 L^2 + \dots + \beta_p L^p = \sum_{j=1}^p \beta_j L^j.$$

Le processus d'innovation ε_t^2 est par définition la variable $\mu_t = \varepsilon_t^2 - h_t^2$. En remplaçant dans l'équation (1.12), les variables h_{t-j}^2 par $\varepsilon_{t-j}^2 - \mu_{t-j}$ on obtient la représentation :

$$\begin{aligned} \varepsilon_t^2 - \mu_t &= \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j (\varepsilon_{t-j}^2 - \mu_{t-j}) \\ \varepsilon_t^2 &= \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j \varepsilon_{t-j}^2 - \sum_{j=1}^p \beta_j \mu_{t-j} + \mu_t \\ \varepsilon_t^2 &= \alpha_0 + \sum_{i=1}^r (\alpha_i + \beta_i) \varepsilon_{t-i}^2 - \sum_{j=1}^p \beta_j \mu_{t-j} + \mu_t, \quad t \in \mathbb{Z}, \end{aligned}$$

où $r = \max(p, q)$ avec la convention $\alpha_i = 0$ (resp. $\beta_j = 0$) si $i > q$ (resp. $j > p$)

La définition 1 ne fournit pas directement de processus la vérifiant. La définition plus restrictive suivante permettra d'obtenir explicitement des processus solutions.

Définition 1.4 (GARCH(p,q) fort). Soit $\{z_t\}$ une suite de variables aléatoires *i.i.d.* On dit que $\{\varepsilon_t\}$ est un processus GARCH(p,q) au sens fort (relativement à la suite z_t) s'il vérifie

$$\begin{aligned} \varepsilon_t &= z_t h_t \\ h_t^2 &= \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j h_{t-j}^2 \end{aligned} \quad (1.14)$$

où les α_i et β_j sont des constantes positives et α_0 est une constante strictement positive.

Il est clair qu'un processus GARCH fort tel que h_t^2 est mesurable par rapport à la tribu I_{t-1} est un processus GARCH faible. La réciproque n'est cependant pas vraie.

1.8 Etude de la stationnarité

Nous allons chercher sous quelles conditions il existe des processus stationnaires (au sens strict et au second-ordre) vérifiant les définitions 1.3 et 1.4. On s'intéresse plus particulièrement aux solutions non anticipatives du modèle (1.14) c'est-à-dire aux processus (ε_t) tel que ε_t soit une fonction mesurable des variables z_{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, pour $x > 0$, $\log^+ x = \max(\log x, 0)$.

1.9 Modèle GARCH(1,1)

Dans le cas où $p = q = 1$, le modèle (1.14) s'écrit :

$$\begin{aligned} \varepsilon_t &= z_t h_t \\ h_t^2 &= \alpha_0 + \alpha \varepsilon_{t-1}^2 + \beta h_{t-1}^2 \end{aligned} \quad (1.15)$$

avec $\alpha_0 > 0$; $\alpha \geq 0$ et $\beta \geq 0$. On pose $a(y) = \alpha y^2 + \beta$.

Théorème 1.2 (stationnarité stricte du modèle GARCH(1,1) fort). Si

$$-\infty \leq \gamma := E \log\{\alpha z_t^2 + \beta\} < 0, \quad (1.16)$$

la série

$$\eta_t = \left\{ 1 + \sum_{i=1}^{\infty} a(z_{t-1}) \dots a(z_{t-i}) \right\} \alpha_0 \quad (1.17)$$

converge presque sûrement (p.s) et le processus (ε_t) défini par $\varepsilon_t = \sqrt{\eta_t} z_t$ est l'unique solution strictement stationnaire du modèle (1.15). Cette solution est non anticipative et ergodique.

Si $\gamma \geq 0$ et $\alpha_0 > 0$, il n'existe pas de solution strictement stationnaire.

Preuve. on a : $\varepsilon_t = z_t h_t$ et $h_t^2 = \alpha_0 + \alpha \varepsilon_{t-1}^2 + \beta h_{t-1}^2$ Utilisant de manière itérative cette équation, on obtient, pour $N \geq 1$

$$\begin{aligned}
 h_t^2 &= \alpha_0 + \alpha \cdot z_{t-1}^2 \cdot h_{t-1}^2 + \beta h_{t-1}^2 \\
 &= \alpha_0 + (\alpha \cdot z_{t-1}^2 + \beta) h_{t-1}^2 \\
 &= \alpha_0 + a(z_{t-1}) h_{t-1}^2 \\
 &= \alpha_0 + a(z_{t-1}) [\alpha_0 + a(z_{t-2}) \cdot h_{t-2}^2] \\
 &= \alpha_0 + \alpha_0 a(z_{t-1}) + a(z_{t-1}) a(z_{t-2}) h_{t-2}^2 \\
 &= \alpha_0 + \alpha_0 a(z_{t-1}) + a(z_{t-1}) a(z_{t-2}) [\alpha_0 + a(z_{t-3}) h_{t-3}^2] \\
 &= \alpha_0 + \alpha_0 a(z_{t-1}) + \alpha_0 a(z_{t-1}) a(z_{t-2}) + a(z_{t-1}) a(z_{t-2}) a(z_{t-3}) h_{t-3}^2 \\
 &= \cdot \\
 &= \cdot \\
 &= \cdot \\
 &= \alpha_0 [1 + a(z_{t-1}) + a(z_{t-1}) a(z_{t-2}) + \dots + a(z_{t-1}) a(z_{t-2}) a(z_{t-3}) \dots a(z_{t-N})] + a(z_{t-1}) a(z_{t-2}) + \dots \\
 &\quad + a(z_{t-N-1}) \cdot h_{t-N-1}^2 \\
 &= \alpha_0 \left[1 + \sum_{i=1}^N a(z_{t-1}) \dots a(z_{t-i}) \right] + a(z_{t-1}) \dots a(z_{t-N-1}) h_{t-N-1}^2 \\
 &= \eta_t(N) + a(z_{t-1}) \dots a(z_{t-N-1}) h_{t-N-1}^2
 \end{aligned} \tag{1.18}$$

Le processus limite $\eta_t = \lim_{N \rightarrow +\infty} \eta_t(N)$ 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 $\eta_t(N) = \alpha_0 + a(z_{t-1}) \eta_{t-1}(N-1)$, on obtient :

$$\eta_t = \alpha_0 + a(z_{t-1}) \eta_{t-1}$$

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

$$\begin{aligned}
 [a(z_{t-1}) \cdots a(z_{t-n})]^{\frac{1}{n}} &= \exp\left[\frac{1}{n} \log\{a(z_{t-1}) \cdots a(z_{t-n})\}\right] \\
 &= \exp\left[\frac{1}{n} \sum_{i=1}^n \log a(z_{t-i})\right] \rightarrow e^\gamma \text{ p.s.}
 \end{aligned} \tag{1.19}$$

quand $n \rightarrow \infty$, par application de la loi forte des grands nombres à la suite *i.i.d* ($\log\{a(z_t)\}$).

La série définie en (1.17) converge alors presque sûrement dans \mathbb{R} . Par suite, le processus $\{\varepsilon_t\}$ défini par :

$$\varepsilon_t = \sqrt{\eta_t} z_t = \left\{ \alpha_0 + \sum_{i=1}^{\infty} a(z_{t-1}) \dots a(z_{t-i}) \alpha_0 \right\}^{1/2} z_t \tag{1.20}$$

est strictement stationnaire et ergodique et non anticipative (car comme (1.17) fonction mesurable des variables z_{t-i} , $i \geq 0$). De plus $\varepsilon_t = h_t z_t$.

Nous montrons maintenant l'unicité. Soit $\tilde{\varepsilon}_t = \sigma_t z_t$ une autre solution strictement stationnaire. D'après (1.18) on a

$$h_t^2 = \eta_t(N) + a(z_{t-1}) \dots a(z_{t-N-1}) h_{t-N-1}^2.$$

Par suite

$$h_t^2 - \eta_t = \{\eta_t(N) - \eta_t\} + a(z_{t-1}) \dots a(z_{t-N-1}) h_{t-N-1}^2.$$

Le terme entre accolades à droite de l'égalité tend vers 0 p.s. quand $N \rightarrow \infty$. Par ailleurs, puisque la série définissant η_t converge p.s., on a $a(z_{t-1}) \dots a(z_{t-n}) \rightarrow 0$ avec probabilité 1 quand $n \rightarrow \infty$. De plus la loi de h_{t-N-1}^2 est indépendante de N par stationnarité. Par suite $a(z_{t-1}) \dots a(z_{t-N-1}) h_{t-N-1}^2 \rightarrow 0$ en probabilité lorsque $N \rightarrow \infty$. De plus, h_{t-N-1}^2 est indépendant de N et donc on a $h_t^2 = \eta_t$ pour tout t , p.s.

Si $\gamma > 0$, par application de la règle de Cauchy, $\sum_{n=1}^N a(z_{t-1}) \dots a(z_{t-n}) \rightarrow +\infty$ p.s. lorsque $N \rightarrow \infty$. Donc si $\alpha_0 > 0$, $\eta_t = +\infty$, p.s. D'après (1.18), il est clair que $h_t^2 = +\infty$, p.s. Par suite, il n'existe pas de solution finie p.s. pour (1.15).

Dans le cas $\gamma = 0$, nous procéderons par l'absurde. Supposons qu'il existe une solution strictement stationnaire (ε_t, h_t^2) de (1.15). Nous avons pour $n > 0$,

$$h_0^2 \geq \alpha_0 \left\{ 1 + \sum_{i=1}^n a(z_{-1}) \dots a(z_{-i}) \right\}$$

d'où on déduit que le terme général $a(z_{-1}) \dots a(z_{-i}) \alpha - 0$ converge vers zéro, p.s., quand $n \rightarrow \infty$.

Remarque 1.2. 1. Le coefficient $\gamma = \mathbb{E} \log\{a(z_t)\}$ existe toujours dans $[-\infty, +\infty[$ car $\mathbb{E} \log^+ \{a(z_t)\} \leq \mathbb{E} a(z_t) = \alpha + \beta$.

2. Dans le cas où $\alpha_0 = 0$ et $\gamma < 0$, il est clair d'après (1.17) que la seule solution stationnaire du modèle est $\varepsilon_t = 0$. Il est donc naturel d'imposer $\alpha_0 > 0$ dans la pratique.
3. On voit que la condition (1.16) dépend de la loi du processus z_t et qu'elle n'est pas symétrique en α et β .
4. La condition (1.16) implique $\beta < 1$. Inversement, si

$$\alpha + \beta < 1$$

(1.16) est vérifiée, car par application de l'inégalité de Jensen

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

5. Si (1.16) est satisfaite, elle l'est également pour tout couple (α_1, β_1) tel que $\alpha_1 = \alpha$ et $\beta_1 = \beta$. En particulier la stationnarité stricte du modèle GARCH implique celle du modèle ARCH obtenu en supprimant β .
6. Dans le cas ARCH(1) ($\beta = 0$), la contrainte de stationnarité stricte s'écrit

$$0 \leq \alpha < \exp\{-\mathbb{E}(\log z_t^2)\}. \quad (1.21)$$

Théorème 1.3 (Stationnarité au second ordre du GARCH(1,1)). *Supposons $\alpha_0 > 0$.*

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

Si $\alpha + \beta < 1$, le processus (ε_t) défini par (1.21), est stationnaire au second ordre. Plus précisément, (ε_t) est un bruit blanc. De plus, il n'existe pas d'autre solution stationnaire au second ordre et non anticipative.

Preuve. Si $\{\varepsilon_t\}$ est un processus GARCH(1,1), au sens de la définition (1.3), stationnaire au second-ordre et non anticipatif, on a :

$$\begin{aligned} \mathbb{E}(\varepsilon_t^2) &= \mathbb{E}\{\mathbb{E}(\varepsilon_t^2 | \varepsilon_u, u < t)\} = \mathbb{E}(h_t^2) \\ &= \mathbb{E}\{\alpha_0 + \alpha \varepsilon_{t-1}^2 + \beta h_{t-1}^2\} \\ &= \alpha_0 + \alpha \mathbb{E}(\varepsilon_{t-1}^2) + \beta \mathbb{E}(\varepsilon_{t-1}^2) \\ &= \alpha_0 + (\alpha + \beta) \mathbb{E}(\varepsilon_{t-1}^2) \end{aligned}$$

Soit

$$(1 - \alpha - \beta) \mathbb{E}(\varepsilon_t^2) = \alpha_0.$$

Il faut donc $\alpha + \beta < 1$. On obtient de plus : $\mathbb{E}(\varepsilon_t^2) > 0$. Inversement, supposons que $\alpha + \beta < 1$. D'après la remarque 4 précédente, la condition de stationnarité stricte est vérifiée. Il suffit donc de montrer que la solution strictement stationnaire définie en (1.20) admet une variance finie. La variable η_t étant une limite croissante de variables aléatoires positives, d'après le théorème de Beppo Levi, on peut intervertir espérance et somme infinie et écrire

$$\begin{aligned} \mathbb{E}(\varepsilon_t^2) = \mathbb{E}(\eta_t) &= \left[1 + \sum_{n=1}^{+\infty} \mathbb{E}\{a(z_{t-1}) \cdots a(z_{t-n})\} \right] \alpha_0 \\ &= \left[1 + \sum_{n=1}^{+\infty} \{\mathbb{E}a(z_t)\}^n \right] \alpha_0 \\ &= \left[1 + \sum_{n=1}^{+\infty} (\alpha + \beta)^n \right] \alpha_0 \\ &= \frac{\alpha_0}{1 - (\alpha + \beta)}. \end{aligned}$$

Cela suffit pour affirmer que la solution est faiblement stationnaire. De plus cette solution est un bruit blanc car :

$$\mathbb{E}(\varepsilon_t) = \mathbb{E}\{\mathbb{E}(\varepsilon_t|\varepsilon_u; u \leq t)\} = 0$$

et pour tout $j > 0$,

$$\begin{aligned} \text{cov}(\varepsilon_t, \varepsilon_{t-j}) &= \mathbb{E}(\varepsilon_t \varepsilon_{t-j}) \\ &= \mathbb{E}\{\mathbb{E}(\varepsilon_t \varepsilon_{t-j} | \varepsilon_j, j \leq t-1)\} \\ &= \mathbb{E}\{\varepsilon_{t-j} \mathbb{E}(\varepsilon_t | \varepsilon_j, j \leq t-j)\} = 0 \end{aligned}$$

Nous allons montrer l'unicité :

Soit $\tilde{\varepsilon}_t = \sqrt{\tilde{\eta}_t} z_t$ une autre solution stationnaire au second ordre et non anticipative. On a :

$$|\eta_t - \tilde{\eta}_t| = |\alpha_0 + a(z_{t-1})\eta_{t-1} - (\alpha_0 + a(z_{t-1})\tilde{\eta}_{t-1})| = a(z_{t-1}) \cdots a(z_{t-n}) |\eta_{t-n-1} - \tilde{\eta}_{t-n-1}|$$

et par suite

$$\begin{aligned} \mathbb{E}|\eta_t - \tilde{\eta}_t| &= \mathbb{E}\{a(z_{t-1}) \cdots a(z_{t-n})\} \mathbb{E}|\eta_{t-n-1} - \tilde{\eta}_{t-n-1}| \\ &= (\alpha + \beta)^n \mathbb{E}|\eta_{t-n-1} - \tilde{\eta}_{t-n-1}| \end{aligned}$$

Notons que la seconde égalité résulte du caractère non anticipatif des solutions, hypothèse qui n'était pas nécessaire pour établir l'unicité de la solution strictement stationnaire. L'espérance de $|\eta_{t-n-1} - \tilde{\eta}_{t-n-1}|$ étant bornée par $\mathbb{E}|\eta_{t-n-1}| + \mathbb{E}|\tilde{\eta}_{t-n-1}|$, quantité finie et indépendante de n par stationnarité, et $(\alpha + \beta)^n$ tendant vers 0 quand $n \rightarrow \infty$, on obtient $\mathbb{E}|\eta_t - \tilde{\eta}_t| = 0$ et donc $\eta_t = \tilde{\eta}_t$ pour tout t , p.s.

1.10 Modèle GARCH(p,q)

Dans le cas général du GARCH(p,q) fort, l'écriture vectorielle suivante est très utile. On a :

$$Y_t = b_t + A_t Y_{t-1} \tag{1.22}$$

où

$$b_t = b_t(z_t) = \begin{bmatrix} \alpha_0 z_t^2 \\ 0 \\ \vdots \\ \alpha_0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \in \mathbb{R}^{p+q}, \quad Y_t = \begin{bmatrix} \varepsilon_t^2 \\ \vdots \\ \varepsilon_{t-q+1}^2 \\ h_t^2 \\ \vdots \\ h_{t-p+1}^2 \end{bmatrix} \in \mathbb{R}^{p+q},$$

et

$$A_t = \begin{bmatrix} \alpha_1 z_t^2 & \cdots & \alpha_q z_t^2 & \beta_1 z_t^2 & \cdots & \beta_p z_t^2 \\ 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 1 & 0 & 0 & \cdots & 0 \\ \alpha_1 & \cdots & \alpha_q & \beta_1 & \cdots & \beta_p \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 1 & 0 \end{bmatrix} \quad (1.23)$$

est une matrice de dimension $(p+q) \times (p+q)$. L'équation(1.22) constitue un modèle vectoriel autorégressif d'ordre 1, avec coefficients positifs et *i.i.d*. La loi de Y_t conditionnelle à son passé infini coïncide avec sa loi conditionnelle à Y_{t-1} seulement. En itérant (1.22) on obtient

$$\begin{aligned} Y_t &= b_t + A_t Y_{t-1} \\ &= b_t + A_t(b_{t-1} + A_{t-1} Y_{t-2}) \\ &= \cdots \\ Y_t &= b_t + \sum_{k=1}^{\infty} A_t A_{t-1} \cdots A_{t-k+1} b_{t-k} \end{aligned} \quad (1.24)$$

L'objet de ce qui suit est de trouver des conditions justifiant l'existence de cette série. Une condition suffisante pour que, presque sûrement (1.24) soit strictement supérieur à 0, est évidemment :

$$\alpha_0 > 0, \alpha_i \geq 0(i = 1, \dots, q), \beta_j \geq 0(j = 1, \dots, p). \quad (1.25)$$

Cette condition, très simple à utiliser, n'est cependant pas toujours nécessaire.

1.11 Modèles FIGARCH et ARFIMA

La valeur de la volatilité est une fonction décroissante de l'ordre de retard. La spécification linéaire de la variance conditionnelle semble saisir l'effet des erreurs passées ou de mémoire longue trouvé dans les travaux empiriques. Cette spécification ne pose pas de problèmes d'estimation particuliers. Toutefois, pour que la variance conditionnelle soit positive, des contraintes supplémentaires sur les paramètres du modèle sont nécessaires. Pour contourner cette dernière difficulté, Engle (1982, 1983) et Engle et Kraft (1983) proposent de fixer le nombre de retards prendre en compte au lieu de travailler avec des ARCH infinis. En se basant sur les travaux faits sur l'équation de la moyenne conditionnelle, une extension généralisée des modèles ARCH a été développée : Le modèle GARCH tient compte non seulement de la volatilité courante exprimée par les carrés des résidus passés mais aussi de

la volatilité passée : il présente ainsi une spécification plus flexible de la variance conditionnelle. Toutefois, le modèle GARCH ne permet de modéliser que la dépendance temporelle de court terme : l'effet des chocs décroît exponentiellement dans le temps. Pour "étendre" la mémoire du modèle GARCH, Engle et Bollerslev (1986) introduisent une extension du modèle GARCH dite intégrée (IGARCH) qui présente une mémoire "explosive" : les effets de chocs sont persistants à l'infini.

Compte tenu des propriétés opposées des modèles traditionnels : décroissance exponentielle de l'impact des chocs sur la volatilité des modèles GARCH par opposition leur persistance à l'infinie dans les modèles IGARCH, un troisième modèle plus flexible a été introduit. Sur le marché de change, Baillie, Bollerslev et Mikkelsen (1996) ont proposé une version ajustée des deux modèles : c'est le processus GARCH fractionnellement intégré (FIGARCH).

Ces modèles sont spécifiés par analogie avec les modèles ARFIMA dans l'équation de la moyenne. Cette spécification a été discutée par TeyssiLre (1997), Breidt Crato et De Lima (1998). Chung (2001) a émis des réserves sur la spécification FIGARCH adoptée par Baillie, Bollerslev et Mikkelsen (1996) et a proposé une nouvelle version.

Dans ce qui suit nous allons présenter un modèle FIGARCH, et nous étudions les propriétés de la modélisation FIGARCH, et ses différentes spécifications.

1.11.1 Modèle FIGARCH

Soit ε_t un processus GARCH(p,q) tel que :

$$\varepsilon_t = z_t h_t$$

où $\{z_t\}$ désigne une suite de variables aléatoires *i.i.d* , tel que $\mathbb{E}(z_t) = 0$ et $V(z_t) = 1$ et h_t vérifie :

$$h_t^2 = \alpha_0 + \alpha(L)\varepsilon_t^2 + \beta(L)h_t^2, t \in \mathbb{Z}$$

où L est l'opérateur de retard ($L^i \varepsilon_t^2 = \varepsilon_{t-i}^2$ et $L^i h_t^2 = h_{t-i}^2$ pour tout entier i), α et β sont des polynômes de L de degrés respectifs q et p. Pour s'assurer de la stabilité et de la stationnarité du processus $\{\varepsilon_t^2\}$, nous supposons que les racines des deux polynômes $[1 - \alpha(L) - \beta(L)]$ et $[1 - \beta(L)]$ sont à l'extérieur du cercle unitaire. Cette condition est satisfaite si $(\alpha + \beta) < 1$. Le modèle GARCH (p,q) peut s'écrire sous la forme d'un modèle ARCH infini (ARCH(∞)), tel que :

$$h_t^2 = \alpha_0 [1 - \beta(L)]^{-1} + \alpha(L) [1 - \beta(L)]^{-1} \varepsilon_t^2$$

Soit :

$$h_t^2 = \alpha_0[1 - \beta(L)]^{-1} + \lambda(L)\varepsilon_t^2$$

où $\lambda(L) = \alpha(L)[1 - \beta(L)]^{-1} = \sum_{j=1}^{\infty} \lambda_j L^j$ décroît exponentiellement lorsque j tend vers l'infini ($j \rightarrow \infty$).

Le modèle GARCH (p,q) peut être réécrit sous la forme d'un modèle ARMA(r,p) tel que $r = \max p, q$. Soit $\mu_t = \varepsilon_t^2 - h_t^2$ les résidus dans l'équation de la variance conditionnelle (processus d'innovation). En reprenant la définition du modèle GARCH(p,q) et en remplaçant h_t^2 par $\varepsilon_t^2 - \mu_t$, On a :

$$h_t^2 = \alpha_0 + \alpha(L)\varepsilon_t^2 + \beta(L)h_t^2$$

$$\varepsilon_t^2 - \mu_t = \alpha_0 + \alpha(L)\varepsilon_t^2 + \beta(L)[\varepsilon_t^2 - \mu_t]$$

$$\varepsilon_t^2 - \alpha(L)\varepsilon_t^2 - \beta(L)\varepsilon_t^2 = \alpha_0 + \mu_t - \beta(L)\mu_t$$

on obtient :

$$[1 - \alpha(L) - \beta(L)]\varepsilon_t^2 = \alpha_0 + [1 - \beta(L)]\mu_t \tag{1.26}$$

Une large littérature financière suggère que la volatilité conditionnelle des prix d'actions présente une persistance significative des chocs (une longue mémoire). Les modèles standards ARCH, GARCH, GARCH exponentiel (EGARCH) ne permettent pas de modéliser ce type de persistance.

Des études empiriques plus récentes montrent que les deux processus $\{\varepsilon_t^2\}$ et $\{\log \varepsilon_t^2\}$ laissent entrevoir que les chocs de volatilité décroissent lentement et sont persistants. Bien qu'ils soient plus flexibles que d'autres modèles et tiennent compte de plusieurs propriétés spécifiques des marchés financiers, les modèles GARCH présentent une mémoire courte, contredisant certaines propriétés empiriques de la volatilité des séries financières. Mais que se passe-t-il si $[1 - \alpha(L) - \beta(L)]$ admet une racine unitaire ?

Dans ce cas, notre processus est non stationnaire au sens de GARCH. Pour remédier ces limites, Engle et Bollerslev (1986) définissent le modèle GARCH intégré (IGARCH). Il s'écrit :

$$(1 - L)\phi(L)\varepsilon_t^2 = \alpha_0 + [1 - \beta(L)]\mu_t \tag{1.27}$$

où $\phi(L) = \sum_{i=1}^{r-1} \phi_i L^i$ ce modèle présente une persistance explosive des chocs de volatilité sur la dynamique de la série. En se basant sur le modèle IGARCH, Baillie, Bollerslev et

Mikkelsen (1996) s'inspirent de l'analogie entre le modèle ARMA et le modèle ARFIMA, pour définir le modèle FIGARCH à partir du modèle GARCH. L'idée est de permettre au paramètre d'intégration d , qui appartenait à $\{0, 1\}$ dans le modèle GARCH/IGARCH, de varier sur $[0; 1]$. Le modèle résultant est beaucoup plus flexible : les chocs de volatilité décroissent exponentiellement mais sont plus persistants que dans le modèle GARCH et moins persistants que dans le modèle IGARCH. Le modèle FIGARCH est obtenu en remplaçant l'opérateur de retard $(1 - L)$ dans la dernière équation par l'opérateur de retard fractionnel $(1 - L)^d$. Le modèle FIGARCH (p, d, q) s'écrit alors :

$$(1 - L)^d \phi(L) \varepsilon_t^2 = \alpha_0 + [1 - \beta(L)] \mu_t \quad (1.28)$$

où les polynômes $\phi(L)$ et $[1 - \beta(L)]$ admettent des racines à l'extérieur du cercle unitaire. Le processus FIGARCH (p, d, q) , pareillement que pour le GARCH (p, q) , admet une représentation ARCH. En remplaçant μ_t par son expression et en réorganisant les termes de l'équation ci-dessus, on obtient :

$$\begin{aligned} (1 - L)^d \phi(L) \varepsilon_t^2 &= \alpha_0 + [1 - \beta(L)] (\varepsilon_t^2 - h_t^2) \\ h_t^2 - \beta(L) h_t^2 &= \alpha_0 + (1 - L)^d \phi(L) \varepsilon_t^2 + \varepsilon_t^2 - \beta(L) \varepsilon_t^2 \end{aligned}$$

Donc

$$[1 - \beta(L)] h_t^2 = \alpha_0 + [1 - \beta(L) - (1 - L)^d \phi(L)] \varepsilon_t^2$$

ou encore :

$$h_t^2 = \alpha_0 [1 - \beta(L)]^{-1} + [1 - (1 - \beta(L))^d \phi(L)] [1 - \beta(L)]^{-1} \varepsilon_t^2 \quad (1.29)$$

On note par $\lambda(L)$, le polynôme d'ordre infini, tel que :

$$\lambda(L) = 1 - (1 - L)^d \phi(L) [1 - \beta(L)]^{-1}$$

Le modèle FIGARCH s'écrit :

$$h_t^2 = \alpha_0 [1 - \beta(L)]^{-1} + \lambda(L) \varepsilon_t^2 \quad (1.30)$$

On considère $\phi(L)$ dans l'équation (2.3) :

$$\phi(L) = 1 - \sum_{j=1}^q \phi_j L^j$$

Chung analyse l'analogie de base entre le modèle ARFIMA et le modèle FIGARCH et montre qu'il y a des différences structurelles entre les deux spécifications.

Le modèle ARFIMA a été introduit par Granger (1980, 1981), Granger et Joyeux (1980) et Hosking (1981) pour étudier la mémoire longue dans l'équation de la moyenne. Un processus stochastique Y_t suit un processus ARFIMA(a, d_0, r), s'il vérifie :

$$(1 - L)^{d_0} \Psi(L)(Y_t - \mu) = \theta(L)\varepsilon_t$$

où μ est la moyenne non conditionnelle de Y_t , $\Psi(L) = 1 - \sum_{j=1}^a \Psi_j L^j$ et $\theta(L) = 1 - \sum_{j=1}^r \theta_j L^j$ sont deux polynômes de L d'ordre respectifs a et r et ε_t est un bruit blanc de moyenne σ^2 . Pour s'assurer de la stationnarité et de l'inversibilité du processus Y_t , nous supposons que d_0 varie de $-0,5$ à $0,5$. Ce paramètre décrit la mémoire longue du processus Y_t . L'opérateur de différence fractionnelle est défini à partir de son développement sur la base de séries de Maclaurin, tel que :

$$\begin{aligned} (1 - L)^{d_0} &= 1 - d_0 \sum_{j=1}^{\infty} \frac{\Gamma(j - d_0)}{\Gamma(j + 1)\Gamma(1 - d_0)} L^j \\ &= 1 - \sum_{j=1}^{\infty} \pi_j(d_0) L^j \\ &= 1 - \delta_d(L) \end{aligned}$$

Dans certains papiers, on trouve le développement précédent exprimé comme un polynôme d'ordre infini de L donné par :

$$\begin{aligned} (1 - L)^{d_0} &= \sum_{j=0}^{\infty} \frac{\Gamma(j - d_0)}{\Gamma(j + 1)\Gamma(-d_0)} L^j \\ &= \sum_{j=0}^{\infty} \pi_j(d_0) L^j \end{aligned}$$

où $\pi_j(Z) = \frac{\Gamma(j - Z)}{\Gamma(j + 1)\Gamma(-Z)}$ et $\Gamma()$ est la fonction gamma standard.

La constante α_0 du modèle FIGARCH est structurellement différente de dans les modèles ARFIMA. Dans ce dernier, l'opérateur de retard fractionnaire s'applique à la constante μ . De plus, la structure de mémoire longue est différente dans les deux modèles : le paramètre de mémoire d_0 varie entre $0,5$ et $0,5$ alors que d appartient $[0,1]$.

- Si $0 < d_0 < \frac{1}{2}$, le modèle ARFIMA est un modèle stationnaire à mémoire longue. Les autocorrélations sont positives, diminuent hyperboliquement et tendent vers zéro lorsque le retard augmente. La densité spectrale est concentrée autour des faibles fréquences et tend vers l'infini lorsque les fréquences tendent vers zéro.

- Si $d_0 = 0$, le modèle ARFIMA se réduit au modèle ARMA standard.
- Si $-\frac{1}{2} < d_0 < 0$, le modèle est antipersistant, les autocorrélations décroissent hyperboliquement et tendent vers zéro et la densité spectrale est dominée par les composantes des fréquences élevées.

Selon la valeur du paramètre d dans $[0, 1]$, on détermine le profil et la vitesse de propagation des chocs de volatilité. Si on se réfère à l'analyse des fonctions de réponses aux impulsions, dérivées de (2.3), on peut évaluer l'impact de long terme des chocs passés μ_s , $s < t$, sur la série en fonction de la limite des poids cumulés de la fonction de réponses impulsionnelles donnée par :

$$\begin{aligned}\gamma(1) &= \lim_{k \rightarrow +\infty} \sum_{i=0}^k \gamma_i := \lim_{k \rightarrow +\infty} \lambda_k \\ &= F(d-1, 1, 1; 1) \phi^{-1}(1) [1 - \beta(1)]\end{aligned}$$

où $\gamma_k = \frac{\partial E_t(\varepsilon_{t+k}^2)}{\partial \mu_t} - \frac{\partial E_t(\varepsilon_{t+k-1}^2)}{\partial \mu_t}$ et le coefficient d'impulsion, E_t est l'espérance conditionnelle à la date t et $\mu_t = \varepsilon_t^2 - h_t^2$.

- Si $0 \leq d \leq 1$, alors $F(d-1, 1, 1; 1) = 0$ d'où $\gamma(1) = 1$ les chocs de volatilité décroissent hyperboliquement contrairement au modèle GARCH où ces chocs décroissent exponentiellement.
- Si $d = 1$, alors $F(d-1, 1, 1; 1) = 1$ d'où $\gamma(1) = \phi^{-1}(1) [1 - \beta(1)]$ et les chocs sont infiniment persistants.
- Si $d > 1$ alors la variance conditionnelle est explosive et $\gamma(1)$ tend vers $+\infty$.

Le paramètre de mémoire d_0 est appliqué directement sur le carré de l'erreur contrairement au paramètre d s'appliquant sur les h_t^2 dans le FIGARCH. Par ailleurs, un modèle ARFIMA $(a, 0, r)$ se ramène à un modèle ARMA(a,r) tandis qu'un modèle FIGARCH $(p, 0, q)$ ne peut pas être réduit à un GARCH(p,q).

les paramètres du modèle ARFIMA sont moins contraints que ceux du modèle FIGARCH. Ces derniers doivent satisfaire certaines conditions pour assurer de la positivité de la variance conditionnelle.

En se basant sur la première version du modèle, Chung (2001) conclut en la présence de quelques problèmes de spécification. Comme nous l'avons déjà signalé, la constante du modèle FIGARCH peut s'écrire :

$$\alpha_0 = \phi(L)(1 - L)^d h^2 \quad (1.31)$$

où $\phi(L) = 1 - \sum_{j=1}^q \phi_j L^j$ et $(1 - L)^d = \sum_{j=0}^{\infty} \pi_j(d) L^j$. Il suffit de remplacer dans l'équation (2.5) :

$$\alpha_0 = h^2 \phi(1) \sum_{j=0}^{\infty} \pi_j(d) L^j$$

Chung démontre que cette constante est théoriquement égal à 0. Toutefois, des études empiriques montrent que cette constante est faible mais significativement non nulle (voir Baillie, Bollerslev et Mikkelsen (1996), Bollerslev et Mikkelsen (1996) et TeyssiLre, (1997)). Pour contourner ce problème, Chung reprend l'écriture ARMA du modèle GARCH (p,q) (voir Breidt, Crato et De Lima (1998)) en considérant plutôt le processus $\{\varepsilon_t^2\}$:

$$[1 - \alpha(L) - \beta(L)](\varepsilon_t^2 - \sigma^2) = \alpha_0 + [1 - \beta(L)]\mu_t$$

puis, il propose la version suivante du modèle FIGARCH, tel que :

$$\phi(L)(1 - L)^d(\varepsilon_t^2 - \sigma^2) = [1 - \beta(L)]\mu_t \tag{1.32}$$

Par ailleurs, les conditions de positivité de la variance conditionnelle posent un problème. Deux ensembles de conditions suffisantes ont été proposés : le premier est défini dans Baillie, Bollerslev et Mikkelsen(1996) et le deuxième est celui de Chung (2001). Nous notons que ces deux ensembles ne sont pas équivalents bien qu'ils soient tous deux suffisants : certains paramètres vérifient certaines conditions mais pas la totalité de ces conditions. Baillie, Bollerslev et Mikkelsen (1996) considèrent que ces conditions sont à définir au cas par cas. Ils considèrent l'exemple particulier du modèle GARCH(p,q), étudié par Nelson et Cao (1992) alors que Chung, dans une première version de son papier en 1996, considLre l'écriture ARCH du modèle FIGARCH. Il montre que si $\beta \neq 0$, les paramètres de la variance conditionnelle doivent vérifier :

$$\lambda_1 = d - \beta - \phi \tag{1.33}$$

$$\lambda_2 = (d - \beta)(\beta - \phi) + \frac{d(d + 1)}{2} \tag{1.34}$$

$$\lambda_k = \beta^{k-2}(d - \beta)(\beta - \phi) - (\beta - \phi) \sum_{j=2}^{k-1} \pi_j(d) \beta^{k-1-j} - \pi_k(d) \quad k = 3, 4, \dots \tag{1.35}$$

pour que la variance conditionnelle soit positive, il suffit que tous les paramètres λ_k , $k = 1, \dots$ soient positifs. Ceci revient à imposer :

$$0 \leq \phi \leq \beta \leq d < 1$$

Bollerslev et Mikkelsen (1996) quant à eux se basent sur une relation autorégressive entre les paramètres λ_k tel que :

$$\lambda_1 = d - \beta - \phi$$

$$\lambda_k = \beta\lambda_{k-1} + [(k-1-d)k^{-1} - \phi]\pi_k(d) \quad k = 2, 3, \dots$$

où $\pi_k(d) = \pi_{k-1}(d)(k-1-d)k^{-1}$ $k = 2, 3, \dots$ sont les coefficients d'une expansion de séries Maclaurin $(1-L)$. Ils déduisent les inégalités suivantes comme conditions suffisantes pour assurer la positivité de la variance conditionnelle :

$$\beta - d \leq \phi \leq \frac{2-d}{3} \quad (1.36)$$

$$d\left(\phi - \frac{1-d}{2}\right) \leq \beta \leq \beta(d - \beta + \phi) \quad (1.37)$$

Chapitre 2

Estimation des paramètres ARCH/GARCH

Dans ce chapitre, nous allons traiter l'estimation des paramètres d'un modèle (G)ARCH, et plus généralement, d'un modèle de régression avec erreur (G)ARCH. Les modèles introduits reposent sur des formulations des moyennes et variances conditionnelles.

En pratique celle-ci souvent paramétrées de façon que la moyenne conditionnelle $m_t(\theta)$ et la variance conditionnelle $h_t^2(\theta)$ apparaissent comme des fonctions de paramètres inconnus et de valeurs passées du processus. La connaissance de ces moments ne suffit cependant pas sans hypothèse supplémentaire à caractériser la loi conditionnelle du processus.

La vraisemblance est écrite comme si la loi des variables z_t était normale centrée réduite (on parle de pseudo ou quasi-vraisemblance), mais cette hypothèse n'est pas nécessaire pour la convergence forte de l'estimateur. Elle a évidemment un effet sur la variance de la loi normale asymptotique de l'estimateur. Les modèles avec erreurs hétéroscédastiques peuvent être estimés généralement par ces méthodes :

- Estimation par la méthode des moindres carrés ordinaires (MCO)
- Estimateurs de la classe du Maximum de Vraisemblance (MV)

Nous présenterons ici ces deux méthodes d'estimation. Commençons par étudier la méthode des moindres carrés, ainsi que le cas des estimateurs du MV obtenue en cas d'erreur sur la spécification de la loi conditionnelle des résidus.

2.1 Estimation des modèles ARCH par la méthode des moindres carrés

Dans cette partie nous considérons l'estimation par la MCO du modèle ARCH(q) :

$$\begin{aligned}\varepsilon_t &= z_t h_t \\ h_t^2 &= \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 \quad \text{avec } \alpha_0 > 0, \alpha_i \geq 0, i = 1, \dots, q\end{aligned}\quad (2.1)$$

(z_t) est une suite de variables *i.i.d.*, $E(z_t) = 0$, $Var(z_t) = 1$.

La méthode consiste à tirer partie de la représentation AR sur le carré du processus observé et à appliquer la méthode des moindres carrés quasi-généralisés (MCQG). Les estimateurs obtenus sont, au moins pour n grand, moins précis que ceux du quasi-maximum de vraisemblance (QMV), mais plus faciles à obtenir. Ils peuvent également fournir des valeurs initiales pour la procédure d'optimisation utilisée dans l'obtention d'estimateurs du QMV plus précis.

On déduit de (2.1) la représentation AR(q) :

$$\varepsilon_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \mu_t \quad (2.2)$$

où $\mu_t = \varepsilon_t^2 - h_t^2 = (z_t^2 - 1)h_t^2$. La suite $(\mu_t, \mathcal{F}_{t-1})$, constitue une différence de martingale.

On suppose que l'on dispose d'observations $\varepsilon_1, \dots, \varepsilon_n$, réalisation partielle du processus $\{\varepsilon\}$, et de valeurs initiales $\varepsilon_0, \dots, \varepsilon_{1-q}$.

$$\begin{aligned}\varepsilon_n^2 &= \alpha_0 + \alpha_1 \varepsilon_{n-1}^2 + \alpha_2 \varepsilon_{n-2}^2 + \dots + \alpha_q \varepsilon_{n-q}^2 + \mu_n \\ \varepsilon_{n-1}^2 &= \alpha_0 + \alpha_1 \varepsilon_{n-2}^2 + \alpha_2 \varepsilon_{n-3}^2 + \dots + \alpha_q \varepsilon_{n-1-q}^2 + \mu_{n-1} \\ &\vdots \\ \varepsilon_1^2 &= \alpha_0 + \alpha_1 \varepsilon_0^2 + \dots + \alpha_q \varepsilon_{1-q}^2 + \mu_1\end{aligned}$$

Introduisant le vecteur

$$Z'_{t-1} = (1, \varepsilon_{t-1}^2, \dots, \varepsilon_{t-q}^2),$$

On déduit le système

$$\varepsilon_t^2 = Z'_{t-1} \theta_0 + \mu_t, \quad \text{avec } \theta_0 = (\alpha_0, \dots, \alpha_q), \quad t = 1, \dots, n. \quad (2.3)$$

soit

$$Y = X\theta_0 + U$$

en définissant la matrice $n \times q$ et les vecteurs $n \times 1$:

$$X = \begin{bmatrix} Z'_{n-1} \\ \vdots \\ Z'_0 \end{bmatrix}, Y = \begin{bmatrix} \varepsilon_n^2 \\ \vdots \\ \varepsilon_1^2 \end{bmatrix}, U = \begin{bmatrix} \mu_n \\ \vdots \\ \mu_0 \end{bmatrix}.$$

Supposons que la matrice $X'X$ soit inversible. On en déduit l'estimateur des MCO de θ :

$$\hat{\theta}_n = (X'X)^{-1}X'Y. \quad (2.4)$$

Pour établir la consistance et la normalité asymptotique de l'estimateur de moindres carrés, les hypothèses suivantes sont nécessaires

H1 : $\{\varepsilon_t\}$ est solution non anticipative strictement stationnaire.

H2 : $E(\varepsilon_t^4) < +\infty$.

H3 : $P[z_t^2 = 1] \neq 1$.

Théorème 2.1 (Convergence des estimateurs MCO pour un ARCH). *Soit $(\hat{\theta}_n)$ une suite d'estimateurs satisfaisant (2.4). Sous les hypothèses H1 – H3, presque sûrement*

$$\hat{\theta}_n \rightarrow \theta_0, \quad \hat{h}_t^2 \rightarrow h_0^2, \quad \text{quand } n \rightarrow \infty.$$

Preuve. étape1 : Nous avons vu que l'unique solution stationnaire non anticipative $\{\varepsilon_t\}$ est ergodique. Le processus (Z_t) est également ergodique car Z_t s'écrit comme fonction mesurable.

$$\frac{1}{n}X'X = \frac{1}{n} \sum_{t=1}^n Z_{t-1}Z'_{t-1} \rightarrow E(Z_{t-1}Z'_{t-1}), \text{ p.s. quand } n \rightarrow \infty. \quad (2.5)$$

L'existence de l'espérance est assurée par l'hypothèse H3. On a donc

$$\frac{1}{n}X'Y = \frac{1}{n} \sum_{t=1}^n Z_{t-1}\varepsilon_t^2 \rightarrow E(Z_{t-1}\varepsilon_t^2), \text{ p.s. quand } n \rightarrow \infty.$$

étape2 : Montrons par l'absurde l'inversibilité de la matrice $E(Z_{t-1}Z'_{t-1}) = E(Z_tZ'_t)$ Supposons qu'il existe un vecteur non nul $c \in \mathbb{R}^{q+1}$ tel que

$$c'E(Z_tZ'_t) = 0$$

$$E[c'Z_t(c'Z_t)'] = 0, \text{ d'où l'on déduit que } c'Z_t \text{ est p.s. constant.}$$

Par suite, il existe une combinaison linéaire p.s. égale à une constante des variables $\varepsilon_t^2, \dots, \varepsilon_{t-q+1}^2$. Sans perdre de généralité, on suppose que le coefficient de ε_t^2 égale à 1.

Donc (z_t) s'exprime p.s. comme fonction mesurable des variables $\varepsilon_{t-1}, \varepsilon_{t-2}, \dots, \varepsilon_{t-q}$. comme

(z_t) est indépendante de ces variables, donc (z_t^2) est égale à une constante p.s.

Forcément, cette constante ne peut être que 1, d'où contradiction avec H3.

étape3 : Il découle de ce qui précède que $\frac{1}{n}X'X$ est *p.s* inversible, pour n assez grand et que *p.s* quand $n \rightarrow \infty$,

$$\hat{\theta}_n = \left(\frac{X'X}{n} \right)^{-1} \frac{X'Y}{n} \rightarrow \{E(Z_{t-1}Z'_{t-1})\}^{-1} E(Z_{t-1}\varepsilon_t^2).$$

étape4 : le processus (μ_t) est l'innovation forte de $\{\varepsilon_t^2\}$. On a donc, en particulier, les relations d'orthogonalité

$$E(\mu_t) = E(\mu_t\varepsilon_t^2 - 1) = \dots = E(\mu_t\varepsilon_t^2 - q) = 0$$

c'est-à-dire

$$E(Z_{t-1}\mu_t) = 0$$

d'où

$$E(Z_{t-1}\varepsilon_t^2) = E(Z_{t-1}Z'_{t-1})\theta_0.$$

Donc d'après les étapes 2 et 3, $\hat{\theta}_n$ converge *p.s* vers θ_0 .

La convergence forte de \hat{h}_t^2 vers h_t^2 s'en déduit.

Pour la normalité asymptotique de l'estimateur des MCO, nous devons faire l'hypothèse supplémentaire suivante :

H4 : $E(\varepsilon_t^8) < +\infty$

Introduisons les matrices carrées symétriques de taille $q + 1$

$$A = E(Z_{t-1}Z'_{t-1}), \quad I = E(h_t^4 Z_{t-1}Z'_{t-1}).$$

L'inversibilité de la matrice A ce démontre de la même façon comme précédent. celle de I sera montrée dans la preuve du résultat suivant, qui établit la normalité asymptotique de l'estimateur des MCO.

Théorème 2.2. *Sous les hypothèses H1-H4,*

$$\sqrt{n}(\hat{\theta}_n - \theta) \rightarrow^{\mathcal{L}} \mathcal{N}(0, (\mu_4 - 1)A^{-1}IA^{-1}).$$

Preuve. On a :

$$\begin{aligned} \hat{\theta}_n &= \left(\frac{1}{n} \sum_{t=1}^n Z_{t-1}Z'_{t-1} \right)^{-1} \left(\frac{1}{n} \sum_{t=1}^n Z_{t-1}\varepsilon_t^2 \right) \\ &= \left(\frac{1}{n} \sum_{t=1}^n Z_{t-1}Z'_{t-1} \right)^{-1} \left\{ \frac{1}{n} \sum_{t=1}^n Z_{t-1}(Z'_{t-1}\theta_0 + \mu_t) \right\} \\ &= \theta_0 + \left(\frac{1}{n} \sum_{t=1}^n Z_{t-1}Z'_{t-1} \right)^{-1} \left(\frac{1}{n} \sum_{t=1}^n Z_{t-1}\mu_t \right) \end{aligned}$$

Donc

$$\sqrt{n}(\hat{\theta}_n - \theta) = \left(\frac{1}{n} \sum_{t=1}^n Z_{t-1} Z'_{t-1} \right)^{-1} \left\{ \frac{1}{\sqrt{n}} \sum_{t=1}^n Z_{t-1} \mu_t \right\}. \quad (2.6)$$

Soit $\lambda \in \mathbb{R}^{q+1}$, $\lambda \neq 0$. La suite $(\lambda' Z_{t-1} \mu_t, \mathcal{F}_t)$ est une différence de martingale stationnaire, ergodique et de carré intégrable de variance

$$\begin{aligned} \text{Var}(\lambda' Z_{t-1} \mu_t) &= \lambda' E\{Z_{t-1} Z'_{t-1} \mu_t^2\} \lambda \\ &= \lambda' E\{Z_{t-1} Z'_{t-1} (z_t^2 - 1) h_t^4\} \\ &= (\mu_4 - 1) \lambda' I \lambda. \end{aligned}$$

On déduit que, pour tout $\lambda \neq 0$

$$\frac{1}{\sqrt{n}} \sum_{t=1}^n \lambda' Z_{t-1} \mu_t \xrightarrow{\mathcal{L}} \mathcal{N}(0, (\mu_4 - 1) \lambda' I \lambda).$$

Par suite, en appliquant la propriété de Cramer-Wold,

$$\frac{1}{\sqrt{n}} \sum_{t=1}^n Z_{t-1} \mu_t \xrightarrow{\mathcal{L}} \mathcal{N}(0, (\mu_4 - 1) I). \quad (2.7)$$

On montre que cette loi limite est non dégénérée, c'est-à-dire que I est inversible, par le même raisonnement que celui utilisé pour établir l'inversibilité de A dans la preuve du Théorème 2.1. Par suite, on déduit de (2.5), (2.6) et (2.7), par un raisonnement classique, que $\sqrt{n}(\hat{\theta}_n - \theta_0)$ est asymptotiquement normal, de moyenne le vecteur nul, et de variance la matrice du théorème.

2.2 Méthode d'estimation du maximum de vraisemblance

Dans cette partie nous étudions la méthode du maximum de vraisemblance conditionnelle (à des valeurs initiales). Nous présentons de façon parallèle la méthode d'estimation du maximum de vraisemblance (MV) sous l'hypothèse de normalité de la distribution conditionnelle des résidus et la méthode d'estimation du quasi maximum de vraisemblance (QMV). En effet, l'idée générale des estimateurs du PMV consiste à démontrer que si l'on commet une erreur sur la distribution conditionnelle des résidus en utilisant à tort une log-vraisemblance fondée sur une loi normale, l'estimateur du MV ainsi obtenu peut tout de même être convergent si la vraie loi des résidus appartient à la loi normale (Gourieroux, Montfort, 1989). L'estimateur sera (1) asymptotiquement convergent et (2) asymptotiquement normal. Par conséquent, la fonction de vraisemblance définissant l'estimateur du MV

sous l'hypothèse de normalité et la fonction de quasi-vraisemblance de l'estimateur du QMV sont les mêmes.

2.2.1 Quasi-vraisemblance conditionnelle

On supposera que les observations $\varepsilon_1, \dots, \varepsilon_n$ constituent une réalisation (de longueur n) d'un processus GARCH(p, q), solution strictement stationnaire non anticipative du modèle

$$\begin{aligned} \varepsilon_t &= z_t h_t \\ h_t^2 &= \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j h_{t-j}^2 \end{aligned} \quad (2.8)$$

où $\{z_t\}$ une suite de variables aléatoires *i.i.d* centrées et de variance unité, $\alpha_0 > 0$, $\alpha_i \geq 0$ ($i = 1, \dots, q$), $\beta_j \geq 0$ ($j = 1, \dots, p$).

Les ordres p et q sont supposés connus. Le vecteur des paramètres

$$\theta = (\theta_1, \dots, \theta_{p+q+1})' := (\alpha_0, \alpha_1, \dots, \alpha_q, \beta_1, \dots, \beta_p)' \quad (2.9)$$

appartient à un espace de paramètres $\Theta \subset]0, +\infty[\times]0, \infty[^{p+q}$. La vraie valeur du paramètre est inconnue et est notée $\theta_0 = (\alpha_0, \alpha_{01}, \dots, \alpha_{0q}, \beta_{01}, \dots, \beta_{0p})'$.

Pour écrire la vraisemblance du modèle, il faut spécifier une distribution particulière pour les variables z_t . On considère généralement la quasi-vraisemblance gaussienne, *i.e* la vraisemblance obtenue à partir d'une loi normale centrée réduite pour les z_t .

La spécification d'une distribution gaussienne pour les variables z_t ne permet pas d'en déduire simplement la loi de l'échantillon. On travaille avec la vraisemblance de $\varepsilon_1, \dots, \varepsilon_n$ conditionnellement à certaines valeurs initiales.

Etant données des valeurs initiales $\varepsilon_0, \dots, \varepsilon_{1-q}, \tilde{h}_0^2, \dots, \tilde{h}_{1-p}^2$ que nous allons préciser, la vraisemblance conditionnelle gaussienne $L_n(\theta)$ s'écrit

$$L_n(\theta) = L_n(\theta; \varepsilon_1, \dots, \varepsilon_n) = \prod_{t=1}^n \frac{1}{\sqrt{2\pi \tilde{h}_t^2}} \exp\left(-\frac{\varepsilon_t^2}{2\tilde{h}_t^2}\right), \quad (2.10)$$

où les \tilde{h}_t^2 sont définis récursivement, pour $t \geq 1$, par

$$\tilde{h}_t^2 = \tilde{h}_t^2(\theta) = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j \tilde{h}_{t-j}^2. \quad (2.11)$$

Pour une valeur donnée de θ , sous l'hypothèse de stationnarité au second ordre, la variance non conditionnelle (correspondant à cette valeur de θ) est un choix raisonnable pour les

valeurs initiales inconnues :

$$\varepsilon_0^2 = \dots = \varepsilon_{1-q}^2 = h_0^2 = \dots = h_{1-p}^2 = \frac{\alpha_0}{1 - \sum_{i=1}^q \alpha_i - \sum_{j=1}^p \beta_j}. \quad (2.12)$$

De telles valeurs initiales ne conviennent pas notamment pour les modèles IGARCH, pour lesquels l'hypothèse de stationnarité au second ordre est relâchée, car la constante (2.12) prendrait des valeurs négatives pour certaines valeurs de θ . On peut alors proposer de prendre comme valeurs initiales

$$\varepsilon_0^2 = \dots = \varepsilon_{1-q}^2 = \tilde{h}_0^2 = \dots = \tilde{h}_{1-p}^2 = \alpha_0 \quad (2.13)$$

où

$$\varepsilon_0^2 = \dots = \varepsilon_{1-q}^2 = \tilde{h}_0^2 = \dots = \tilde{h}_{1-p}^2 = \varepsilon_1^2. \quad (2.14)$$

Un estimateur du QMV de θ est défini comme toute quantité $\hat{\theta}_n$ vérifiant presque sûrement

$$L_n(\hat{\theta}_n) = \sup_{\theta \in \Theta} L_n(\theta). \quad (2.15)$$

On voit, en prenant le logarithme, que maximiser la vraisemblance revient à minimiser par rapport à θ

$$\tilde{I}_n(\theta) = n^{-1} \sum_{t=1}^n \tilde{\ell}_t, \quad \text{où } \tilde{\ell}_t = \tilde{\ell}_t(\theta) = \frac{\varepsilon_t^2}{\tilde{h}_t^2} + \ln \tilde{h}_t^2 \quad (2.16)$$

et \tilde{h}_t^2 est définie en (2.11). Un estimateur du quasi-maximum de vraisemblance est donc une solution mesurable de l'équation

$$\hat{\theta}_n = \arg \min_{\theta \in \Theta} \tilde{I}_n(\theta). \quad (2.17)$$

2.2.2 Equations de vraisemblance

On obtient les équations de vraisemblance en annulant la dérivée par rapport à θ du critère $\tilde{I}_n(\theta)$, ce qui donne

$$\frac{1}{n} \sum_{t=1}^n \{\varepsilon_t^2 - \tilde{h}_t^2\} \frac{1}{\tilde{h}_t^4} \frac{\partial \tilde{h}_t^2}{\partial \theta} = 0. \quad (2.18)$$

En effet, comme nous le verrons plus précisément dans la partie suivante, le terme de gauche de l'égalité précédente se comporte asymptotiquement comme

$$\frac{1}{n} \sum_{t=1}^n \{\varepsilon_t^2 - h_t^2\} \frac{1}{h_t^4} \frac{\partial h_t^2}{\partial \theta} \quad (2.19)$$

l'influence des valeurs initiales étant nulle lorsque $n \rightarrow \infty$. Or, pour la vraie valeur du paramètre, l'innovation de ε_t^2 est $\mu_t = \varepsilon_t^2 - h_t^2$. Donc sous réserve que l'espérance existe, on a

$$E_{\theta_0} \left(\mu_t \frac{1}{h_t^4(\theta_0)} \frac{\partial h_t^2(\theta_0)}{\partial \theta} \right)$$

Car $\frac{1}{h_t^4(\theta_0)} \frac{\partial h_t^2(\theta_0)}{\partial \theta}$ est une fonction mesurable des ε_{t-i} , $i > 0$. Ce résultat n'est autre que la version asymptotique de (2.16) en θ_0 , en utilisant le théorème ergodique.

2.2.3 Propriétés asymptotiques de l'estimateur du QMV

Dans ce qui suit, nous utilisons comme norme d'une matrice $A = (a_{ij})$ quelconque la norme $\|A\| = \sum |a_{ij}|$. Le rayon spectral d'une matrice A carrée sera noté $\rho(A)$. Le produit de Kronecker sera noté \otimes .

Convergence forte

Rappelons que le modèle (2.8) possède une solution strictement stationnaire si et seulement si le coefficient de Lyapounov de la suite de matrices

$$A_t = \begin{bmatrix} \alpha_1 z_t^2 & \cdots & \alpha_q z_t^2 & \beta_1 z_t^2 & \cdots & \beta_p z_t^2 \\ 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 1 & 0 & 0 & \cdots & 0 \\ \alpha_1 & \cdots & \alpha_q & \beta_1 & \cdots & \beta_p \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 1 & 0 \end{bmatrix}$$

est strictement négatif. On note $\gamma(\theta)$ ce coefficient de Lyapounov.

Notons

$$\mathcal{A}_\theta(z) = \sum_{i=1}^q \alpha_i z^i \quad \text{et} \quad \mathcal{B}_\theta(z) = 1 - \sum_{j=1}^p \beta_j z^j.$$

Par convention $\mathcal{A}_\theta(z) = 0$ si $q = 0$ et $\mathcal{B}_\theta(z) = 1$ si $p = 0$.

Pour la convergence, les hypothèses suivantes sont faites.

A1 : $\theta_0 \in \Theta$ et Θ est compact.

A2 : $\gamma(\theta_0) < 0$ et $\forall \theta \in \Theta$, $\sum_{j=1}^p \beta_j < 1$.

A3 : z_t^2 a une loi non dégénérée.

A4 : si $p > 0$, $\mathcal{A}_{\theta_0}(z)$ et $\mathcal{B}_{\theta_0}(z)$ n'ont pas de racine commune, $\mathcal{A}_{\theta_0}(1) \neq 0$, et $\alpha_{0q} + \beta_{0p} \neq 0$.

Il est pratique d'approximer la suite $(\tilde{\ell}_t(\theta))$ par une suite stationnaire ergodique. La condition de stricte stationnarité A2 implique que les racines de $\mathcal{B}_\theta(z)$ sont extérieures au disque unité. Notons donc $(h_t^2)_t = h_t^2(\theta)_t$ la solution strictement stationnaire ergodique et non anticipative de

$$h_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j h_{t-j}^2 \quad \forall t \quad (2.20)$$

et soit

$$\tilde{I}_n(\theta) = n^{-1} \sum_{t=1}^n \tilde{\ell}_t, \quad \text{où } \tilde{\ell}_t = \tilde{\ell}_t(\theta) = \frac{\varepsilon_t^2}{h_t^2} + \ln h_t^2.$$

Théorème 2.3. *Soit $(\hat{\theta}_n)$ une suite d'estimateurs du QMV satisfaisant (2.17), avec les conditions initiales (2.13) ou (2.14). Sous les hypothèses A1-A4, presque sûrement*

$$\hat{\theta}_n \rightarrow \theta_0, \quad \text{quand } n \rightarrow \infty.$$

Remarque 2.1. 1- On ne suppose pas que la vraie valeur θ_0 du paramètre appartient à l'intérieur de Θ . Le théorème permet donc de traiter les cas où certains coefficients, α_i ou β_j , sont nuls.

2- L'hypothèse A4 disparaît dans le cas ARCH. Elle permet de suridentifier l'un des ordres, p ou q , mais pas les deux.

3- L'hypothèse A4 exclut le cas où tous les α_{0i} sont nuls. Ceci est évidemment nécessaire, sinon le modèle a pour solution un bruit blanc fort qui peut s'écrire de multiples manières. Par exemple, un bruit blanc fort de variance 1 peut s'écrire sous la forme d'un GARCH(1,1) avec $h_t^2 = \alpha_0 + 0 \times \varepsilon_{t-1}^2 + \beta h_{t-1}^2$, pour tous α_0 et β positifs tels que $\alpha_0 = 1 - \beta$.

4- L'hypothèse d'absence de racines communes, dans A4, n'est restrictive que si $p > 1$ et $q > 1$. En effet, si $q = 1$ la seule racine de $\mathcal{A}_{\theta_0}(z)$ est 0 et $\mathcal{B}_{\theta_0}(0) \neq 0$. Si $p = 1$ et $\beta_{01} \neq 0$, la seule racine de $\mathcal{B}_{\theta_0}(z)$ est $1/\beta_{01} > 0$ (si $\beta_{01} = 0$, le polynôme n'admet pas de racine). En raison de la positivité des coefficients α_{0i} , cette valeur ne peut annuler $\mathcal{A}_{\theta_0}(z)$.

Normalité asymptotique

Pour montrer la normalité asymptotique les hypothèses supplémentaires suivantes sont nécessaires.

A5 : $\theta_0 \in \dot{\Theta}$, où $\dot{\Theta}$ est l'intérieur de Θ .

A6 : $\kappa_z = E z_t^4 < \infty$.

Théorème 2.4 (Normalité asymptotique des estimateurs du QMV). *Sous les hypothèses A1-A6, $\sqrt{n}(\hat{\theta}_n - \theta_0)$ tend en loi vers une $\mathcal{N}(0, (\kappa_z - 1)J^{-1})$, où*

$$J := E_{\theta_0} \left(\frac{\partial^2 \ell_t(\theta_0)}{\partial \theta \partial \theta'} \right) = E_{\theta_0} \left(\frac{1}{h_t^4(\theta_0)} \frac{\partial h_t^2(\theta_0)}{\partial \theta} \frac{\partial h_t^2(\theta_0)}{\partial \theta'} \right) \quad (2.21)$$

Remarque 2.2. 1- L'hypothèse A5 est classique car elle permet d'utiliser le fait que les conditions du premier ordre sont valides, au moins asymptotiquement. En effet si $\hat{\theta}_n$ est convergent, il appartient également à l'intérieur de Θ pour n grand. En tant que maximum, il doit donc annuler la dérivée de la fonction critère. Cette hypothèse est cependant restrictive car elle exclut par exemple le cas $\alpha_{01} = 0$ (il est cependant clair que dans ce cas, $\sqrt{n}(\hat{\alpha}_1 - \alpha_{01})$ est concentrée sur $[0, \infty[$ et ne peut donc être asymptotiquement normale). Ce type de problèmes, dits "de bord", doit faire l'objet d'une étude spécifique.

2- L'hypothèse A6 ne porte pas sur ε_t^2 , et n'exclut bien sûr pas le cas IGARCH. Seule une hypothèse d'existence du moment d'ordre 4 est imposée sur la suite (z_t) . Cette hypothèse est clairement nécessaire pour l'existence de la variance du vecteur du score $\partial \ell_t(\theta_0) / \partial \theta$.

Chapitre 3

Représentation ARCH(∞)

3.1 Introduction

Les modèles autoregréssifs conditionnellement hétéroscédastique infinis (ARCH(∞)) constitue une généralisation de la classe ARCH(q), introduite par Engle(1982), et leurs extensions GARCH (ARCH généralisés) proposé par Bollerslev (1986), pour permettre une forme flexible de la décroissance des coefficients dans l'équation de volatilité. En particulier, il est possible dans la classe de modèle ARCH(∞) d'inclure un effet de longue mémoire dans le carré des observations, ce qui est souvent le cas dans les rendements financiers. Des exemples de modèles parcimonieux ayant une représentation ARCH(∞), dont les coefficients sont à décroissance hyperbolique, sont les modèles GARCH Fractionnement Intégré notés (FIGARCH). Ces modèles ont été introduits par Baillie *et al*(1996). Davidson (2004) s'est intéressé aux propriétés des moments du modèles FIGARCH et propose une extension, le modèle GARCH hyperbolique avec une transition naturelle, l'auteur a présenté une classe plus large incluant les modèles GARCH noté IGARCH.

Berkes *et al* (2003) ont montré l'existence et l'unicité de la représentation ARCH(∞) d'un modèle GARCH(p,q) et ont montré, sous certains conditions, la consistance et la normalité asymptotique de l'estimateur de quasi-maximum de vraisemblance (QMV). Ces processus ont une mémoire courte avec des coefficients à décroissance exponentielle. Douc *et al* (2008) donnent des conditions suffisantes pour l'existence d'une solution causale stationnaire d'un processus ARCH(∞), qui permet aux coefficients la décroissance et inclut donc les modèles FIGARCH. Dans le cas d'un modèle FIGARCH, cette solution implique nécessairement une variance infinie du processus. Les résultats théoriques sur les modèles ARCH et les propriétés associées ont joué un rôle particulier dans les travaux empiriques d'analyse des données sur les taux de change, les cours des actions.

3.2 Représentation ARCH(∞)

Considérons le processus GARCH(p,q) défini par les équations suivantes :

$$\varepsilon_t = h_t z_t \quad (3.1)$$

et

$$h_t^2 = \alpha_0 + \sum_{i=1}^p \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^q \beta_j h_{t-j}^2 \quad (3.2)$$

où $\alpha_0 > 0$, $\alpha_i \geq 0$, $0 \leq i \leq p$, et $\beta_j \geq 0$, $0 \leq j \leq q$ sont des constantes. Supposons aussi que $\{z_i, i \in \mathbb{Z}\}$ sont des variables aléatoires *i.i.d* centrées et de variance 1.

Nelson (1990) montre que dans le cas d'un GARCH(1,1), les équations (3.1) et (3.2) ont une unique solution stationnaire si et seulement si $E \log(\beta_1 + \alpha_1 z_0^2) < 0$. Dans le cas d'un GARCH(p,q), Bougerol et Picard (1992) donnent des conditions nécessaires et suffisantes pour l'existence d'une unique solution stationnaire des équations (3.1) et (3.2).

Afin de caractériser ces conditions, il est nécessaire d'introduire les notations suivantes utiles pour certaines démonstrations :

$$\tau_n = (\beta_1 + \alpha_1 z_n^2, \beta_2, \dots, \beta_{q-1}) \in \mathbb{R}^{q-1}$$

$$\xi_n = (z_n^2, 0, \dots, 0)' \in \mathbb{R}^{q-1}$$

et

$$\alpha = (\alpha_2, \dots, \alpha_{p-1})' \in \mathbb{R}^{p-2}$$

Introduisons maintenant la matrice A_n de dimension $(p+q-1) \times (p+q-1)$, définie par :

$$A_n = \begin{bmatrix} \tau_n & \beta_q & \alpha & \alpha_p \\ I_{q-1} & 0 & 0 & 0 \\ \xi_n & 0 & 0 & 0 \\ 0 & 0 & I_{p-2} & 0 \end{bmatrix}$$

où I_{q-1} et I_{p-2} sont les matrices d'identités de dimension $q-1$ et $p-2$, respectivement.

Le plus grand exposant de Lyapunov γ_L associé à la suite de matrices $\{A_n, n \in \mathbb{Z}\}$ est :

$$\gamma_L = \inf_{0 \leq n \leq \infty} \frac{1}{n+1} E \log \|A_0 A_1 \cdots A_n\|,$$

en admettant que

$$E(\log \|A_0\|) < \infty. \quad (3.3)$$

En utilisant la condition (3.3), Kingman (1973) montre que

$$\lim_{n \rightarrow \infty} \frac{1}{n+1} \log \|A_0 A_1 \cdots A_n\| = \gamma_L \text{ p.s.}$$

Bougerol et Picard (1992) montrent que si la condition (3.3) est vérifiée, alors (3.1) et (3.2) ont une solution unique strictement stationnaire si et seulement si

$$\gamma_L < 0. \quad (3.4)$$

Soit maintenant $X_n = (h_n^2, \dots, h_{n-q-1}^2, \varepsilon_{n-1}^2, \dots, \varepsilon_{n-p-1}^2)^T \in \mathbb{R}^{p+q-1}$ et $D = (\alpha_0, 0, \dots, 0)^T \in \mathbb{R}^{p+q-1}$. Alors les équations (3.1) et (3.2) peuvent être écrites de manières équivalentes comme suit :

$$X_{n+1} = A_n X_n + D.$$

Bougerol et Picard (1992a; 1992b) montrent que si la condition (3.4) est vérifiée, alors :

$$X_n = D + \sum_{k=0}^{\infty} A_n \cdots A_{n-k} D. \quad (3.5)$$

Dans ce qui suit, nous supposons que les conditions (3.1)-(3.5) sont vérifiées. Il existe un ensemble de conditions minimales d'un modèle GARCH(p,q) stationnaire. Notons par $\theta = (\alpha_0, \alpha_1, \dots, \alpha_q, \beta_1, \dots, \beta_p)'$ le vecteur des paramètres. Supposons que $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ ont été observées. Lumsdaine (1996) a étudié l'estimation du paramètre inconnu θ dans le cas de $p = q = 1$. Il recommande l'estimateur de quasi-maximum de vraisemblance et montre la consistance et la normalité asymptotique pour un modèle GARCH(1,1). Cependant, certaines conditions semblent être inutilement restrictives et devraient être relâchées.

3.2.1 Représentation d'un GARCH(p,q)

Considérons le modèle GARCH(p,q) strictement stationnaire vérifiant (3.1) et (3.2) dont le vecteur des paramètres est : $(\alpha_0, \alpha_1, \dots, \alpha_p, \beta_1, \dots, \beta_q)'$. Posons :

$$\mathcal{A}(x) = \alpha_1 x + \alpha_2 x^2 + \cdots + \alpha_p x^p$$

et

$$\mathcal{B}(x) = 1 - \beta_1 x - \beta_2 x^2 - \cdots - \beta_q x^q.$$

On suppose que l'ordre de $\mathcal{B}(x)$ est exactement q, *i.e.*

$$\beta_q \neq 0. \quad (3.6)$$

Bougerol et Picard (1992) montrent que $\gamma_L < 0$ implique

$$\beta_1 + \beta_2 + \cdots + \beta_q < 1. \quad (3.7)$$

La relation (3.7) implique que toutes les racines de $\mathcal{B}(x) = 0$ se trouvent en dehors du cercle unitaire. Plus précisément, nous avons le lemme suivant.

Lemme 3.1 (Berkès et al (2003)). *La relation (3.7) est équivalente à :*

$$|\gamma_i| > 1 \text{ pour tout } 1 \leq i \leq \ell, \text{ où } \gamma_1, \dots, \gamma_\ell \text{ représentent toutes les solutions de } \mathcal{B}(x) = 0 \quad (3.8)$$

avec les multiplicités v_1, \dots, v_ℓ .

Démonstration. Supposons que $\beta_1 + \beta_2 + \cdots + \beta_q \geq 1$. Comme $\mathcal{B}(0) = 1$ et $\mathcal{B}(1) = 1 - (\beta_1 + \beta_2 + \cdots + \beta_q) \leq 0$, on a au moins une solution de $\mathcal{B}(x) = 0$ dans l'intervalle $[0, 1]$, ce qui contredit (3.8).

Supposons que l'équation (3.7) est vérifiée. Pour tout $|z| \leq 1$, nous avons $\mathcal{B}(z) \geq 1 - (\beta_1|z| + \beta_2|z|^2 + \cdots + \beta_q|z|^q) \geq 1 - (\beta_1 + \beta_2 + \cdots + \beta_q) > 0$, donc (3.8) est vérifiée. \square

Le lemme qui suit est utile pour la représentation d'un ARCH(∞). Soit $\log^+ x = \log x$ si $x > 1$, et 0 sinon.

Lemme 3.2 (Berkès et al (2003)). *Soit $\{\varepsilon_t\}$ un processus GARCH(p, q) vérifiant (3.1) et (3.2) avec $\xi_n = (z_n^2, 0, \dots, 0) \in \mathbb{R}^{q-1}$.*

Si $\{\xi_k, k \in \mathbb{N}\}$ est une suite de variables aléatoires identiquement distribuées satisfaisant

$$E \log^+ |\xi_0| < \infty, \quad (3.9)$$

alors $\sum_{k=0}^{\infty} \xi_k z^k$ converge en probabilité 1 pour tout $|z| < 1$.

Démonstration. D'après le lemme de Borel-Cantelli, il est suffisant de montrer que, pour tout $\zeta > 1$,

$$\sum_{k=1}^{\infty} P\{|\xi_k| > \zeta^k\} < \infty. \quad (3.10)$$

la distribution de ξ_k ne dépend pas de k , alors

$$\begin{aligned} \sum_{k=1}^{\infty} P\{|\xi_k| > \zeta^k\} &= \sum_{k=1}^{\infty} P\{\log^+ |\xi_k| > k \log \zeta\} \\ &= \sum_{k=1}^{\infty} P\{\log^+ |\xi_0| > k \log \zeta\} \\ &\leq E \log^+ |\xi_0| / \log \zeta, \end{aligned}$$

et donc l'équation (3.9) implique (3.10). \square

Nous établissons maintenant une représentation de h_k^2 en termes de ε_{k-i}^2 , $i \geq 1$. Comme on a d'après, le lemme 3.1, $\mathcal{B}(x)$ a toutes les racines en dehors du cercle unitaire, alors nous avons

$$\sum_{j=0}^{\infty} d_j x^j = \frac{1}{\mathcal{B}(x)}, \quad |x| \leq 1, \quad (3.11)$$

et les coefficients d_0, d_1, d_2, \dots ont une décroissance exponentielle. Soit

$$c_0 = \alpha_0 \sum_{m=0}^{\infty} d_m \quad (3.12)$$

et

$$c_j = \alpha_1 d_{j-1} + \alpha_2 d_{j-2} + \dots + \alpha_p d_{j-p}, \quad 1 \leq j < \infty, \quad c_0 = \alpha_0 / \mathcal{B}(1). \quad (3.13)$$

Et aussi,

$$\frac{\mathcal{A}(x)}{\mathcal{B}(x)} = \sum_{i=1}^{\infty} c_i x^i, \quad |x| \leq 1. \quad (3.14)$$

Nous avons par (3.13) les coefficients c_1, c_2, \dots ont une décroissance exponentielle.

Le théorème suivant nous donne une condition importante pour l'existence d'une représentation d'un ARCH(∞)

Théorème 3.1 (Berkes *etal*, (2003)). *Soit $\{\varepsilon_t\}$ un GARCH(p, q) vérifiant les équations (3.1) et (3.2). Si*

$$E \log h_0^2 < \infty, \quad (3.15)$$

alors

$$h_k^2 = c_0 + \sum_{i=1}^{\infty} c_i \varepsilon_{k-i}^2, \quad \text{pour tous } k, \quad (3.16)$$

avec la probabilité 1.

Démonstration. Comme $\|A_0\| \geq \|\xi_0\| = z_0^2$, (3.3) donne que $E \log^+ z_0^2 \leq E \log^+ \|A_0\| < \infty$. De l'équation (3.1) nous avons $E \log^+ \varepsilon_0^2 \leq E \log^+ h_0^2 + E \log^+ z_0^2$, et de (3.15) nous obtenons

$$E \log^+ \varepsilon_0^2 < \infty. \quad (3.17)$$

Comme la suite c_1, c_2, \dots est à une décroissance exponentielle, le lemme 3.2 donne que la série (3.16) est absolument convergent en probabilité 1. Donc on a

$$\xi_k = \alpha_0 + \sum_{i=1}^{i=p} \alpha_i \varepsilon_{k-i}^2 \quad (3.18)$$

satisfait l'équation (3.9).

Nous devons montrer que

$$h_k^2 = \sum_{0 \leq m < \infty} d_m \xi_{k-m}. \quad (3.19)$$

La série (3.19) est absolument convergente en probabilité 1 par le lemme 3.2 et la décroissance exponentielle de d_i . Il en résulte de l'équation (3.11) que

$$\begin{aligned} d_0 &= 1, \\ d_1 &= \beta_1, \\ d_2 &= d_1\beta_1 + \beta_2, \\ &\vdots \\ d_q &= d_{q-1}\beta_1 + \cdots + d_1\beta_{q-1} + \beta_q \end{aligned} \quad (3.20)$$

et

$$d_i = d_{i-1}\beta_1 + \cdots + d_{i-q}\beta_q, \text{ pour } i > q. \quad (3.21)$$

En utilisant les équations (3.20) et (3.21), il n'est pas difficile de vérifier que, pour $j \geq q$,

$$\xi_k + d_1\xi_{k-1} + d_2\xi_{k-2} + \cdots + d_j\xi_{k-j} = h_k^2 - \sum_{i=1}^{i=q} (d_{i+j-q}\beta_q + \cdots + d_j\beta_i)h_{k-i-j}^2. \quad (3.22)$$

Par le lemme 3.2, comme $j \rightarrow \infty$, le côté gauche de l'équation (3.22) est vérifiée *p.s.* que le côté droit de l'équation (3.19). En utilisant la décroissance exponentielle de d_j et (3.15), on obtient

$$\sum_{1 \leq j < \infty} P\left\{ \left| \sum_{1 \leq i \leq q} (d_{i+j-q}\beta_q + \cdots + d_j\beta_i)h_{k-i-j}^2 \right| > \delta \right\} < \infty, \text{ pour tous } \delta > 0,$$

et donc, en appliquant le lemme de Borel-Cantelli, nous concluons (3.16). \square

Remarque 3.1. En utilisant l'opérateur de retard L , on peut écrire alors la variance définie dans l'équation (3.16) comme suit

$$h_k^2 = \frac{\mathcal{A}(L)}{\mathcal{B}(L)} \left(\frac{\alpha_0}{\mathcal{A}(1)} + \varepsilon_k^2 \right)$$

L'unicité de la représentation d'un GARCH(∞) est établie par le théorème suivant :

Théorème 3.2 (Berkès et al, (2003)). Soit $\{\varepsilon_t\}$ un GARCH(p, q) vérifiant les équations (3.1) et (3.2). Nous supposons que

$$z_0^2 \text{ est une variable aléatoire non-dégénérée.} \quad (3.23)$$

Si pour un certain k ,

$$h_k^2 = c_0 + \sum_{1 \leq i < \infty} c_i \varepsilon_{k-i}^2 \text{ p.s.} \quad \text{et} \quad h_k^2 = c_0^* + \sum_{1 \leq i < \infty} c_i^* \varepsilon_{k-i}^2 \text{ p.s.}, \quad (3.24)$$

alors $c_i = c_i^*$ pour tous $1 \leq i < \infty$.

Démonstration. Montrons ce résultat par l'absurde. Soit $m > 0$ le plus petit entier satisfaisant $c_m \neq c_m^*$. Si $c_j = c_j^*$ pour tous $j > 0$, alors $c_0 = c_0^*$ doit être aussi vérifiée.

Par la définition de m , nous avons

$$(c_m^* - c_m) \varepsilon_{k-m}^2 = c_0 - c_0^* + \sum_{m < j < \infty} (c_j - c_j^*) \varepsilon_{k-j}^2,$$

et (3.1) nous donne

$$z_{k-m}^2 = \frac{1}{(c_m^* - c_m) h_{k-m}^2} \{c_0 - c_0^* + \sum_{m < j < \infty} (c_j - c_j^*) \varepsilon_{k-j}^2\}.$$

Comme $h_{k-m}^2 \geq \alpha_0 > 0$, alors z_{k-m}^2 est bien défini. Soit \mathcal{F}_j le σ -algèbre engendrée par $\{z_i\}$ avec $-\infty < i \leq j$. La relation donnée par (3.5) montre que ε_j est \mathcal{F}_j -mesurable et donc z_{k-m}^2 défini ci-dessus est une variable aléatoire à valeur réelle, mesurable par rapport à \mathcal{F}_{k-m-1} . Comme les z_i sont indépendent, cela implique que z_{k-m}^2 est une constante p.s., ce qui contredit (3.23). \square

En combinant les deux résultats relatifs à l'existence et l'unicité de la représentation d'un ARCH(∞), nous avons le théorème et le lemme suivant :

Théorème 3.3 (Berkes et al, (2003)). *Si la condition (3.15) est satisfaite, alors le modèle GARCH(p, q) vérifiant (3.1) et (3.2) admet une représentation ARCH(∞) donnée par l'équation (3.16), avec une décroissance exponentielle des c_i .*

Si de plus, l'équation (3.23) est aussi satisfaite, alors la représentation (3.16) est unique.

La preuve de ce théorème peut se faire en combinant les deux preuves précédente.

Lemme 3.3 (Berkes et al, (2003)). *Soit $\{\varepsilon_t\}$ un GARCH(p, q) vérifiant (3.1) et (3.2). Si*

$$E|z_0^2|^\delta < \infty \quad \text{pour un certain } \delta > 0, \quad (3.25)$$

alors il existe un $\delta^ > 0$ tel que*

$$E|\varepsilon_0^2|^{\delta^*} < \infty \quad \text{et} \quad E|h_0^2|^{\delta^*} < \infty.$$

Démonstration. Par l'équation (3.4) et la définition de γ_L , il existe un entier $m \geq 1$ tel que

$$E \log \|A_0 A_1 \cdots A_{m-1}\| < 0, \quad (3.26)$$

et $\|A_0\| \leq C(1 + z_0^2)$. De (3.25) nous avons

$$E\|A_0 A_1 \cdots A_{m-1}\|^\delta \leq (E\|A_0\|^\delta)^m < \infty. \quad (3.27)$$

Introduisons la fonction $s(t) = E\|A_0 A_1 \cdots A_{m-1}\|^t$. Comme $s'(t) = E \log \|A_0 A_1 \cdots A_{m-1}\| < 0$, $s(t)$ décroît au voisinage de 0, et comme $s(0) = 1$, il existe $0 < \delta^* < 1$ tel que

$$E\|A_0 A_1 \cdots A_{m-1}\|^{\delta^*} < 1.$$

En utilisant (3.5), nous concluons que

$$\|X_0\| \leq \|D\| + \sum_{0 \leq k < \infty} \|A_0 \cdots A_{-k}\| \|D\|,$$

et comme $0 < \delta^* < 1$ on obtient que

$$\|X_0\|^{\delta^*} \leq \|D\|^{\delta^*} + \sum_{0 \leq k < \infty} \|A_0 \cdots A_{-k}\|^{\delta^*} \|D\|^{\delta^*}.$$

En utilisant aussi $E\|A_0 A_1 \cdots A_{m-1}\|^{\delta^*} < 1$, il en résulte facilement qu'il existe $0 < \tilde{c} < \infty$ et $0 < \tilde{\rho} < 1$ tel que

$$E\|A_0 A_1 \cdots A_k\|^{\delta^*} \leq \tilde{c} \tilde{\rho}^k,$$

ce qui justifie que $E\|X_0\|^{\delta^*} < \infty$, d'où la conclusion du lemme 3.3 □

Les théorèmes 3.1 et 3.2 montrent qu'il existe une identification un par un entre la séquence $h_k^2, k \in \mathbb{Z}$ et les coefficients $c_i, i \in \mathbb{Z}$. Cependant, cela n'est pas suffisant pour estimer les paramètres d'un GARCH(p,q), alors nous avons également besoin que la définition minimale dans le sens où il n'y a pas de couple (p^*, q^*) tel que $p^* < p$ ou $q^* < q$ et

$$h_k^2 = \alpha_0^* + \sum_{1 \leq i \leq p^*} \alpha_i^* \varepsilon_{k-i}^2 + \sum_{1 \leq j \leq q^*} \beta_j^* h_{k-j}^2 \quad (3.28)$$

pour un certain (pas nécessairement non-négatif) $\alpha_0^*, \alpha_i^* (1 \leq i \leq p^*)$ et $\beta_j^* (1 \leq j \leq q^*)$.

Théorème 3.4 (Berkes et al, (2003)). *Nous supposons que les équations (3.15) et (3.23) sont satisfaites. Alors le modèle (3.1)-(3.2) est minimale si et seulement si*

les polynômes $\mathcal{A}(x)$ et $\mathcal{B}(x)$ sont des coprimés dans l'ensemble des polynômes à coefficients réels (3.29)

Démonstration. Supposons que (3.29) est vérifiée et supposons qu'il existe (p^*, q^*) , tels que $p^* < p$ ou $q^* < q$, et α_0^*, α_i^* ($1 \leq i \leq p^*$) et β_j^* ($1 \leq j \leq q^*$) tels que (3.29) est vérifiée. Soit $\mathcal{A}^*(x) = x\alpha_1^* + \dots + x^{p^*}\alpha_{p^*}^*$ et $\mathcal{B}^*(x) = 1 - (\beta_1^*x + \dots + \beta_{q^*}^*x^{q^*})$. Comme l'équation (3.14), nous avons

$$\sum_{i=1}^{\infty} c_i x^i = \frac{\mathcal{A}^*(x)}{\mathcal{B}^*(x)}, \quad (3.30)$$

par conséquent, $\mathcal{A}(x)/\mathcal{B}(x) = \mathcal{A}^*(x)/\mathcal{B}^*(x)$. Par (3.29), $\mathcal{A}(x)$ et $\mathcal{B}(x)$ sont coprimés, et donc nous concluons qu'il existe un polynôme $\mathcal{P}(x)$ tel que $\mathcal{A}^*(x) = \mathcal{A}(x)\mathcal{P}(x)$ et $\mathcal{B}^*(x) = \mathcal{B}(x)\mathcal{P}(x)$, et par conséquent $p^* \geq p, q^* \geq q$, est une contradiction.

Supposons maintenant, que la définition (3.1)-(3.2) est minimale mais (3.29) n'est pas vérifiée, *i.e* il existe des polynômes $\mathcal{A}^*, \mathcal{B}^*$ et \mathcal{P} tel que \mathcal{P} n'est pas constant et $\mathcal{A}^*(x) = \mathcal{A}(x)\mathcal{P}(x)$ et $\mathcal{B}^*(x) = \mathcal{B}(x)\mathcal{P}(x)$. Nous avons montrer que avec $\tilde{\alpha}_0 = \alpha_0\mathcal{B}^*(1)/\mathcal{B}(1)$, on a :

$$h_k^2 = \tilde{\alpha}_0 + \sum_{1 \leq i \leq p^*} \alpha_i^* \varepsilon_{k-i}^2 + \sum_{1 \leq j \leq q^*} \beta_j^* h_{k-j}^2. \quad (3.31)$$

par les deux équations (3.23) et (3.30) on obtient

$$h_k^2 = \frac{\alpha_0}{\mathcal{B}(1)} + \frac{\mathcal{A}^*(L)}{\mathcal{B}^*(L)} \varepsilon_k^2 = \frac{\tilde{\alpha}_0}{\mathcal{B}^*(1)} + \frac{\mathcal{A}^*(L)}{\mathcal{B}^*(L)} \varepsilon_k^2.$$

Donc

$$\mathcal{B}^*(L)h_k^2 = \mathcal{A}^*(L) \left(\frac{\tilde{\alpha}_0}{\mathcal{A}^*(1)} + \varepsilon_k^2 \right).$$

En remarquant que les degrés de \mathcal{A}^* et \mathcal{B}^* sont moins que \mathcal{A} et \mathcal{B} , (3.31) contredit la supposition que la définition (3.1) et (3.2) est minimale. \square

Le corollaire suivant montre l'unicité du vecteur des paramètres.

Corollaire 3.1 (Berkes et al (2003)). *Nous supposons que les conditions (3.15), (3.23) et (3.29) sont satisfait. Alors, il n'y a pas de vecteur*

$(\alpha_0^*, \alpha_1^*, \dots, \alpha_p^*, \beta_1^*, \dots, \beta_q^*) \neq (\alpha_0, \alpha_1, \dots, \alpha_p, \beta_1, \dots, \beta_q)$ tel que

$$h_k^2 = \alpha_0^* + \sum_{1 \leq i \leq p} \alpha_i^* \varepsilon_{k-i}^2 + \sum_{1 \leq j \leq q} \beta_j^* \varepsilon_{k-j}^2. \quad (3.32)$$

Démonstration. Soit le vecteur des paramètres $(\alpha_0^*, \alpha_1^*, \dots, \alpha_p^*, \beta_1^*, \dots, \beta_q^*)'$ qui satisfait l'équation (3.32) et soit $\mathcal{A}^*, \mathcal{B}^*$ les polynômes analogues de \mathcal{A} et \mathcal{B} pour le vecteur des paramètres $(\alpha_0^*, \alpha_1^*, \dots, \alpha_p^*, \beta_1^*, \dots, \beta_q^*)'$. En suivant la première partie de la preuve du théorème 3.4, il découle qu'il existe un polynôme \mathcal{P} tel que $\mathcal{A}^*(x) = \mathcal{A}(x)\mathcal{P}(x)$ et $\mathcal{B}^*(x) = \mathcal{B}(x)\mathcal{P}(x)$. Puisque \mathcal{B} et \mathcal{B}^* ont le même degré q et le même terme constant 1, nous obtenons $\mathcal{P}(x) = 1$ et le corollaire est prouvé. \square

La section suivante contient des résultats préliminaires utilisés dans la construction de l'estimateur quasi-maximum de vraisemblance.

3.2.2 Récurrences liées à la représentation infinie de GARCH(p,q)

Nous avons déjà donné la formule explicite pour les coefficients c_i dans l'équation (3.16) en terms des racines $\gamma_1, \dots, \gamma_\ell$ du polynôme \mathcal{B} . Cependant, la solution de $\mathcal{B}(x) = 0$ n'est pas facile à calculer, en particulier si q est grand, et ainsi les formules sont numériquement peu pratiques. On a

$$c_n = \frac{d^n}{dx^n} \left(\frac{\mathcal{A}(x)}{\mathcal{B}(x)} \right)_{x=0} \quad 1 \leq n < \infty,$$

En calculant les dérivés et en utilisant $\mathcal{B}(0) = 1$, nous constatons que les c_n , $1 \leq n < \infty$, sont en fait des polynômes de $\alpha_1, \dots, \alpha_p, \beta_1, \dots, \beta_q$ avec des coefficients entiers. Posons, $C(x) = \sum_{n=1}^{\infty} c_n x^n$, l'équation (3.14) montre que $\mathcal{A}(x) = \mathcal{B}(x)C(x)$. En effectuant la multiplication à droite et en identifiant les coefficients, on obtient

$$\alpha_1 = c_1, \quad \alpha_2 = c_2 - \beta_1 c_1, \quad \alpha_3 = c_3 - \beta_1 c_2 - \beta_2 c_1, \dots$$

à partir duquel les coefficients c_1, c_2, \dots peuvent être calculés de manière récursive.

Dans le problème d'estimation, le vecteur des paramètres $\theta = (\alpha_0, \alpha_1, \dots, \alpha_p, \beta_1, \dots, \beta_q)'$ est fixe. Nous avons affaire à une classe de processus dont le vecteur de paramètres sera noté par $u = (x, s_1, \dots, s_p, t_1, \dots, t_q)'$. Par conséquent, les coefficients c_i , $0 \leq i < \infty$, seront des fonctions de u et la formule de récurrence obtenue ci-dessus prend la forme suivante.

Si $q \geq p$, alors

$$\begin{aligned} c_0(u) &= \frac{x}{1 - (t_1 + \dots + t_q)}, \\ c_1(u) &= s_1, \\ c_2(u) &= s_2 + t_1 c_1(u), \\ &\vdots \\ c_p(u) &= s_p + t_1 c_{p-1}(u) + \dots + t_{p-1} c_1(u), \\ c_{p+1}(u) &= t_1 c_p(u) + \dots + t_p c_1(u), \\ &\vdots \\ c_q(u) &= t_1 c_{q-1}(u) + \dots + t_{q-1} c_1(u). \end{aligned}$$

Si $q < p$ l'équation ci-dessus est remplacé par

$$\begin{aligned} c_0(u) &= \frac{x}{1 - (t_1 + \dots + t_q)}, \\ c_1(u) &= s_1, \\ c_2(u) &= s_2 + t_1 c_1(u), \\ &\vdots \\ c_{q+1}(u) &= s_{q+1} + t_1 c_q(u) + \dots + t_q c_1(u), \\ &\vdots \\ c_p(u) &= s_p + t_1 c_{p-1}(u) + \dots + t_q c_{p-q}(u). \end{aligned}$$

Si $i > R = \max(p, q)$, alors

$$c_i(u) = t_1 + c_{i-1}(u) + t_2 c_{i-2}(u) + \dots + t_q c_{i-q}(u). \quad (3.33)$$

Soit $0 < \underline{u} < \bar{u}$, $0 < \rho_0 < 1$, $q\underline{u} < \rho_0$ et définie

$$U = \{u : t_1 + t_2 + \dots + t_q \leq \rho_0 \text{ et } \underline{u} \leq \min(x, s_1, \dots, s_p, t_1, \dots, t_q) \leq \max(x, s_1, \dots, s_p, t_1, \dots, t_q) \leq \bar{u}\}.$$

Nous supposons que θ est à l'intérieur de U .

Nous utilisons $|\cdot|$ pour noté $\max \|\cdot\|$ des vecteurs et des matrices et $x \vee y = \max(x, y)$.

Lemme 3.4 (Berkes et al (2003)). *Pour tous $u = (x, s_1, \dots, s_p, t_1, \dots, t_q) \in U$ et $u^* = (x^*, s_1^*, \dots, s_p^*, t_1^*, \dots, t_q^*) \in U$, nous avons*

$$C_1 u^i \leq c_i(u), \quad 0 \leq i < \infty, \quad (3.34)$$

$$c_i(u) \leq C_2 \rho_0^{i/q}, \quad 0 \leq i < \infty, \quad (3.35)$$

et

$$\frac{c_i(u^*)}{c_i(u)} \leq C_3 \left(\max_{1 \leq j \leq q} \frac{t_j^*}{t_j} \vee 1 \right)^i, \quad 0 \leq i < \infty, \quad (3.36)$$

pour les constantes $0 < C_1, C_2, C_3 < \infty$.

Démonstration. Les résultats sont triviaux pour $i = 0$. Nous utilisons l'induction pour $i \geq 1$. Comme on a $c_i(u)$, $i \geq 1$, sont des polynômes des coordonnées de u avec des coefficients positifs. Cependent, les équations (3.34)-(3.36) sont vérifiées pour tous $1 \leq i \leq R = \max(p, q)$ si C_1 est petit et C_2, C_3 sont assez grand.

Soit $j > R$ et supposons que (3.34)-(3.36) sont valables pour tous $i < j$. Comme $\underline{u} < 1$, par (3.33) nous avons

$$c_j(u) \geq \underline{u} \min_{1 \leq k \leq q} c_{j-k}(u) \geq C_1 \underline{u}^i.$$

et on aura aussi

$$c_j(u) \leq (t_1 + t_2 + \dots + t_q) \max_{1 \leq k \leq q} c_{j-k}(u) \leq \rho_0 C_2 \rho_0^{(j-q)/q} \leq C_2 \rho_0^{j/q}.$$

En utilisant (3.33) et l'hypothèse d'induction nous avons, pour $j > R$,

$$\begin{aligned} 0 &< \frac{c_j(u^*)}{c_j(u)} \\ &= t_1^* \frac{c_{j-1}(u^*)}{c_j(u)} + t_2^* \frac{c_{j-2}(u^*)}{c_j(u)} + \dots + t_q^* \frac{c_{j-q}(u^*)}{c_j(u)} \\ &= \frac{t_1^* c_{j-1}(u^*)}{t_1 c_{j-1}(u)} t_1 \frac{c_{j-1}(u)}{c_j(u)} + \frac{t_2^* c_{j-2}(u^*)}{t_2 c_{j-2}(u)} t_2 \frac{c_{j-2}(u)}{c_j(u)} + \dots + \frac{t_q^* c_{j-q}(u^*)}{t_q c_{j-q}(u)} t_q \frac{c_{j-q}(u)}{c_j(u)} \\ &\leq C_3 \left(\max_{1 \leq i \leq q} \frac{t_i^*}{t_i} \vee 1 \right)^{j-1} \frac{1}{c_j(u)} \{t_1 c_{j-1}(u) + t_2 c_{j-2}(u) + \dots + t_q c_{j-q}(u)\} \left(\max_{1 \leq i \leq q} \frac{t_i^*}{t_i} \right) \\ &\leq C_3 \left(\max_{1 \leq i \leq q} \frac{t_i^*}{t_i} \vee 1 \right)^j. \end{aligned}$$

Donc (3.34)-(3.36) sont aussi vérifiées pour j . □

Par la suite, nous montrons des résultats similaires pour le vecteur suivant :

$$c'_i(u) = \left(\frac{\partial c_i(u)}{\partial x}, \frac{\partial c_i(u)}{\partial s_1}, \dots, \frac{\partial c_i(u)}{\partial s_p}, \frac{\partial c_i(u)}{\partial t_1}, \dots, \frac{\partial c_i(u)}{\partial t_q} \right).$$

Lemme 3.5 (Berkes et al (2003)). *Supposons que $u \in U$. Alors*

$$\left| \frac{\partial c_0(u)}{\partial x} \right| \leq \frac{1}{1 - \rho_0}, \quad (3.37)$$

$$\frac{\partial c_0(u)}{\partial s_j} = 0, \quad 1 \leq j \leq p, \quad (3.38)$$

$$\left| \frac{\partial c_0(u)}{\partial t_j} \right| \leq \frac{\bar{u}}{(1 - \rho_0)^2}, \quad 1 \leq j \leq q, \quad (3.39)$$

et

$$|c'_i(u)|/c_i(u) \leq C_4 i, \quad 1 \leq i < \infty, \quad (3.40)$$

pour une constante C_4 .

Démonstration. Comme on a $c_0(u) = x/(1 - (t_1 + \dots + t_q))$, alors les équations (3.37)-(3.39) sont évidentes. Il reste à prouver l'équation (3.40), nous commençons à observer que

$$\frac{\partial c_i(u)}{\partial x} = 0, \quad 1 \leq i < \infty.$$

Ensuite nous montrons que

$$\left| \frac{\partial c_i(u)}{\partial s_j} / c_i(u) \right| \leq C_5, \quad 1 \leq i < \infty, 1 \leq j \leq p, \quad (3.41)$$

pour une constante C_5 . En utilisant à nouveau l'induction, (3.41) est vérifiée pour $i \leq R = \max(p, q)$, à condition que C_5 soit assez grand. Par (3.33) nous avons pour $i > R$,

$$\frac{\partial c_i(u)}{\partial s_j} = t_1 \frac{\partial c_{i-1}(u)}{\partial s_j} + t_2 \frac{\partial c_{i-2}(u)}{\partial s_j} + \dots + \frac{\partial c_{i-q}(u)}{\partial s_j},$$

et comme $c_\ell > 0$, $\partial c_\ell / \partial s_j \geq 0$ (en rappelant que c_ℓ est un polynôme de $s_1, \dots, s_p, t_1, \dots, t_q$ avec des coefficients positifs), on conclut que

$$\begin{aligned} 0 &< \frac{1}{c_i(u)} \frac{\partial c_i(u)}{\partial s_j} \\ &\leq \max_{1 \leq m \leq q} \frac{1}{c_{i-m}(u)} \frac{\partial c_{i-m}(u)}{\partial s_j} \{t_1 c_{i-1}(u) + \dots + t_q c_{i-q}(u)\} \frac{1}{c_i(u)} \\ &= \max_{1 \leq m \leq q} \frac{1}{c_{i-m}(u)} \frac{\partial c_{i-m}(u)}{\partial s_j}. \end{aligned}$$

Donc si $i > R$ et (3.41) sont vérifiées pour tous les indices inférieurs à i , alors c'est aussi vérifiée pour i .

Nous devons maintenant montrer que

$$\left| \frac{\partial c_i(u)}{\partial t_j} / c_i(u) \right| \leq C_6 i, \quad 1 \leq i < \infty, 1 \leq j \leq q \quad (3.42)$$

pour un certain C_6 . Par (3.33), on a pour $i > R$,

$$\frac{\partial c_i(u)}{\partial t_j} = c_{i-j}(u) + t_1 \frac{\partial c_{i-1}(u)}{\partial t_j} + \dots + t_q \frac{\partial c_{i-q}(u)}{\partial t_j}$$

et donc

$$\frac{1}{c_i(u)} \frac{\partial c_i(u)}{\partial t_j} = \frac{c_{i-j}(u)}{c_i(u)} + \frac{1}{c_{i-1}(u)} \frac{\partial c_{i-1}(u)}{\partial t_j} t_1 \frac{c_{i-1}(u)}{c_i(u)} + \dots + \frac{1}{c_{i-q}(u)} \frac{\partial c_{i-q}(u)}{\partial t_j} t_q \frac{c_{i-q}(u)}{c_i(u)}.$$

Comme $c_i(u) \geq t_j c_{i-j}(u)$, on obtient que

$$\begin{aligned} \left| \frac{1}{c_i(u)} \frac{\partial c_i(u)}{\partial t_j} \right| &\leq \frac{1}{t_j} + \max_{1 \leq m \leq q} \frac{1}{c_{i-m}(u)} \frac{\partial c_{i-m}(u)}{\partial t_j} \frac{1}{c_i(u)} \{t_1 c_{i-1}(u) + \dots + t_q c_{i-q}(u)\} \\ &\leq \frac{1}{u} + \max_{1 \leq m \leq q} \frac{1}{c_{i-m}(u)} \frac{\partial c_{i-m}(u)}{\partial t_j} \end{aligned} \quad ,$$

et donc (3.42) est vérifiée par induction, en supposant que C_6 est choisi plus grand que $\frac{1}{u}$. (3.40) ce déduit par (3.41)-(3.42). \square

Lemme 3.6 (Berkès et al (2003)). *Pour tous $u \in U$, nous avons*

$$|c_0''(u)| \leq C_7,$$

$$|c_i''(u)| \leq C_8 i^2 c_i(u), \quad 1 \leq i < \infty,$$

$$|c_0^{(3)}(u)| \leq C_9,$$

et

$$|c_0^{(3)}(u)| \leq C_{10} i^3 c_i(u), \quad 1 \leq i < \infty,$$

pour certaines constantes C_7, C_8, C_9 et C_{10} .

3.3 Méthode d'estimation et la normalité asymptotique

Le logarithme de la fonction quasi-maximum de vraisemblance dans les GARCH(p,q) est défini comme suit

$$L_n(u) = \sum_{1 \leq k \leq n} -\frac{1}{2} \left\{ \log w_k(u) + \frac{\varepsilon_k^2}{w_k(u)} \right\}, \quad (3.43)$$

où

$$w_k(u) = c_0(u) + \sum_{i=1}^{\infty} c_i(u) \varepsilon_{k-i}^2. \quad (3.44)$$

Les fonctions $c_i(u)$, $i \in \mathbb{N}$, sont déjà définis. En assemblant le lemme 3.2, et les équations (3.17) et (3.35), on obtient que $w_k(u)$ existe en probabilité 1. Clairement, $w_k(\theta) = h_k^2$. Si nous devons supposer que z_0 est standard normal, conditionnellement $\mathcal{F}_{k-1} = \sigma\{z_i, -\infty < i \leq k-1\}$, $\varepsilon_k/h_k(\theta)$ est aussi standard normal. La vraisemblance dans l'équation (3.43) est dérivé sous cette hypothèse. Cependant, nous montrons que la méthode d'estimation de quasi-maximum de vraisemblance $\hat{\theta}_n$, est défini comme suit :

$$\hat{\theta}_n = \arg \max_{u \in U} L_n(u), \quad (3.45)$$

sera consistant et asymptotiquement normal sans supposer que z_0 est standard normale.

Théorème 3.5 (Berkes et al (2003)). *Supposons que θ est à l'intérieur de U et (3.23), (3.29) sont vérifiées et*

$$E|z_0^2|^{1+\delta} < \infty, \quad \text{pour certains } \delta > 0, \quad (3.46)$$

$$\lim_{t \rightarrow 0} t^{-\mu} P\{z_0^2 \leq t\} = 0, \quad \text{pour certains } \mu > 0, \quad (3.47)$$

et

$$Ez_0^2 = 1. \quad (3.48)$$

Donc

$$\hat{\theta}_n \rightarrow \theta \text{ p.s. lorsque } n \rightarrow \infty. \quad (3.49)$$

Les lemmes suivant seront nécessaire pour les preuves des théorèmes suivant :

Lemme 3.7 (Berkes et al (2003)). *Supposons que θ est à l'intérieur de U , (3.46), (3.47) sont vérifiées. Alors*

$$\sup_{u \in U} \left| \frac{1}{n} L_n(u) - L(u) \right| \rightarrow 0, \text{ p.s. pour } n \rightarrow \infty,$$

où

$$L(u) = -\frac{1}{2} E \left\{ \log w_0(u) + \frac{\varepsilon_0^2}{w_0(u)} \right\}.$$

Lemme 3.8 (Berkes et al (2003)). *Si les conditions du théorème 3.5 sont satisfaites, alors $L(u)$, avec $u \in U$ a un unique maximum en θ .*

Lemme 3.9 (Berkes et al (2003)). *Si les conditions du théorème 3.6 sont satisfaites, alors*

$$\sup_{u \in U} \left| \frac{1}{n} L'_n(u) - L'(u) \right| \rightarrow 0, \text{ p.s.},$$

et

$$\sup_{u \in U} \left| \frac{1}{n} L''_n(u) - L''(u) \right| \rightarrow 0, \text{ p.s.}$$

lorsque $n \rightarrow \infty$

Lemme 3.10 (Berkes et al (2003)). *Si les conditions du théorème 3.6 sont satisfaites, alors A_0 est non-singulier.*

Lemme 3.11 (Berkes et al (2003)). *Si les conditions du théorème 3.5 sont satisfaites, alors*

$$\sup_{u \in U} \left| \sum_{1 \leq k \leq n} (\log w_k(u) - \log \tilde{w}_k(u)) \right| = O(1), \text{ p.s}$$

et

$$\sup_{u \in U} \left| \sum_{1 \leq k \leq n} \left(\frac{w'_k(u)}{w_k(u)} - \frac{\tilde{w}'_k(u)}{\tilde{w}_k(u)} \right) \right| = O(1), \text{ p.s}$$

lorsque $n \rightarrow \infty$

Lemme 3.12 (Berkes et al (2003)). *Si les conditions du théorème 3.5 sont satisfaites, alors*

$$\sup_{u \in U} \left| \sum_{1 \leq k \leq n} \left(\frac{\varepsilon_k^2}{w_k(u)} - \frac{\varepsilon_k^2}{\tilde{w}_k(u)} \right) \right| = O(1), \text{ p.s}$$

et

$$\sup_{u \in U} \left| \sum_{1 \leq k \leq n} \varepsilon_k^2 \left(\frac{w'_k(u)}{w_k^2(u)} - \frac{\tilde{w}'_k(u)}{\tilde{w}_k^2(u)} \right) \right| = O(1), \text{ p.s}$$

lorsque $n \rightarrow \infty$

Démonstration. On a U est un ensemble compact. Nous avons déjà par le lemme 3.7 que $L_n(u)/n$ converge uniformément vers $L(u)$ dans U avec une probabilité 1 et le lemme 3.8 montre que $L(u)$ a un maximum unique lorsque $u = \theta$. Donc les arguments standards montrent que les emplacement du maximum de $L_n(u)/n$ converge p.s pour $L(u)$. \square

Nous allons dans ce qui suit discuter la normalité asymptotique de $n^{1/2}(\hat{\theta}_n - \theta)$. Soit

$$\ell_k(u) = -\frac{1}{2}(\log w_k(u) + \varepsilon_k^2/w_k(u)) \tag{3.50}$$

et introduisant la matrice $A_0 = Cov(\ell'_0(\theta))$ et $B_0 = E(\ell''_0(\theta))$.

Théorème 3.6 (Berkes et al (2003)). *Nous supposons que θ est à l'intérieur de U et (3.23), (3.29), (3.47), (3.48) sont vérifiées et*

$$E|z_0^2|^{2+\delta} < \infty, \text{ pour certains } \delta > 0. \tag{3.51}$$

Alors

$$A_0 \text{ et } B_0 \text{ sont non singulier,} \tag{3.52}$$

$$\hat{\theta}_n - \theta = \frac{1}{n} \sum_{1 \leq k \leq n} \frac{1}{2} (1 - z_k^2) \frac{w'_k(\theta)}{w_k(\theta)} B_0^{-1} + op(\sqrt{n}), \text{ pour } n \rightarrow \infty, \tag{3.53}$$

et

$$n^{1/2}(\hat{\theta}_n - \theta) \rightarrow^{\mathcal{D}} \mathcal{N}(0, B_0^{-1}A_0B_0^{-1}), \quad \text{pour } n \rightarrow \infty, \quad (3.54)$$

où $\mathcal{N}(0, C)$ représente des variables aléatoires normales multivariées avec une moyenne 0 et une matrice de covariance C .

Démonstration. par (3.43) et (3.44), $L_n(u)$ est deux fois continuellement différentiable sur U et atteint son maximum à $\hat{\theta}_n$, pour lequel n est suffisamment grand, est un point intérieur de U par (3.49) et θ à l'intérieur de U . Donc il existe un indice aléatoire n_0 tel que

$$L'_n(\hat{\theta}_n) = 0, \quad \text{si } n \geq n_0. \quad (3.55)$$

Par conséquent pour $n \geq n_0$ nous avons $L'_n(\hat{\theta}_n) - L'_n(\theta) = -L'_n(\theta)$, donc par le théorème de la valeur moyenne pour les coordonnées de $L'_n = (L'_{n,0}, \dots, L'_{n,p+q})$, nous obtenons

$$(\hat{\theta}_n - \theta)L''_{n,i}(\xi_{n,i}) = -L'_{n,i}(\theta), \quad 0 \leq i \leq p+q,$$

où $\xi_{n,i}$ est entre $\hat{\theta}_n$ et θ . Alors, en utilisant le lemme 3.9 et la continuité de $L''(u)$, on obtient

$$(\hat{\theta}_n - \theta)(B_0 + o(1)) = -\frac{1}{n}L'_n(\theta) \text{ p.s.} \quad (3.56)$$

nous avons

$$\ell'_k(\theta) = \frac{1}{2}(z_k^2 - 1)\frac{w'_k(\theta)}{w_k(\theta)} \quad (3.57)$$

En vue de (3.48), $E((\ell'_k(\theta))^T \ell'_j(\theta))$, $j \neq k$, et la matrice nulle, et donc nous concluons que

$$\text{cov}(n^{-1/2}L'_n(\theta)) = A_0. \quad (3.58)$$

Par le lemme 3.10, A_0 est non singulier et donc (3.56) implique B_0 est non singulier.

Si n est grand, alors $(B_0 + o(1))^{-1}$ existe et égale à $B_0^{-1} + o(1)$. Par (3.58) et l'inégalité de Markov nous avons $L'_n(\theta)/n^{1/2} = Op(1)$, et donc (3.53) vient des équations (3.56) et (3.57). Comme $\ell'_k(\theta) = \frac{1}{2}(z_k^2 - 1)f(z_{k-1}^2, z_{k-2}^2, \dots)$ pour une certaine fonction f mesurable, $\ell'_k(\theta)$ est une différence de martingale stationnaire. \square

En pratique, nous observons seulement $\varepsilon_1, \dots, \varepsilon_n$ et le logarithme de la fonction de quasi-maximum dans (3.43) ne peut pas être calculé à partir de ces données. Donc nous remplaçons $L_n(u)$ par :

$$\tilde{L}_n(u) = \sum_{1 < k \leq n} -\frac{1}{2} \left\{ \log \tilde{w}_k(u) + \frac{\varepsilon_k^2}{\tilde{w}_k(u)} \right\},$$

où

$$\tilde{w}_k(u) = c_0(u) + \sum_{1 \leq i \leq k-1} c_i(u) \varepsilon_{k-i}^2.$$

Semblable à (3.45), nous le définissons comme suit

$$\tilde{\theta}_n = \arg \max_{u \in U} \tilde{L}_n(u).$$

Les deux théorèmes suivants, montrent que les théorèmes de limites resteront vrais pour $\tilde{\theta}_n$.

Théorème 3.7 (Berkes et al (2003)). *Sous les conditions du théorème (3.5) nous avons*

$$\tilde{\theta}_n \rightarrow \theta \text{ p.s. avec } n \rightarrow \infty.$$

Démonstration. Les lemmes 3.11 et 3.12 implique que

$$\sup_{u \in U} \left| \frac{1}{n} L_n(u) - \frac{1}{n} \tilde{L}_n(u) \right| = o(1), \text{ p.s.},$$

et donc, par le lemme 3.7 nous avons

$$\sup_{u \in U} \left| \frac{1}{n} \tilde{L}_n(u) - L(u) \right| = o(1), \text{ p.s.}$$

alors $\tilde{\theta}_n \rightarrow \theta$ p.s comme dans la preuve du théorème 3.5. □

Théorème 3.8 (Berkes et al (2003)).

$$\hat{\theta}_n - \theta = \frac{1}{n} \sum_{1 \leq k \leq n} \frac{1}{2} (1 - z_k^2) \frac{w'_k(\theta)}{w_k(\theta)} B_0^{-1} + op(\sqrt{n}), \quad \text{comme } n \rightarrow \infty,$$

et

$$n^{1/2}(\hat{\theta}_n - \theta) \rightarrow^D \mathcal{N}(0, B_0^{-1} A_0 B_0^{-1}), \quad \text{comme } n \rightarrow \infty,$$

Démonstration. En observant que $\sup_{u \in U} |L'_n(u) - \tilde{L}'_n(u)|$ est borné par la somme du membre à gauche des équations du lemme 3.11, et en utilisant le lemme 3.12, on obtient sous les conditions du théorème 3.8 que

$$\sup_{u \in U} \left| \frac{1}{n} L'_n(u) - \frac{1}{n} \tilde{L}'_n(u) \right| = O\left(\frac{1}{n}\right), \text{ p.s.} \tag{3.59}$$

Comme la preuve du théorème 3.6, il existe une variable aléatoire n_0 , tel que

$$\tilde{L}'_n(\tilde{\theta}_n) = 0, \quad \text{si } n \geq n_0. \tag{3.60}$$

En utilisant (3.55), (3.59) et (3.60), on obtient

$$\frac{1}{n}L'_n(\hat{\theta}_n) - \frac{1}{n}L'_n(\tilde{\theta}_n) = \frac{1}{n}L'_n(\hat{\theta}_n) - \frac{1}{n}\tilde{L}'_n(\tilde{\theta}_n) + O\left(\frac{1}{n}\right) = O\left(\frac{1}{n}\right) \text{ p.s.} \quad (3.61)$$

par l'équation de lemme 3.9 et une application de coordonnée du théorème de la valeur moyenne, on obtient

$$\frac{1}{n}L'_n(\hat{\theta}_n) - \frac{1}{n}L'_n(\tilde{\theta}_n) = (\hat{\theta}_n - \tilde{\theta}_n)L''(\theta)(1 + o(1)) \text{ p.s.},$$

et donc (3.61) implique

$$|\hat{\theta}_n - \tilde{\theta}_n| = O\left(\frac{1}{n}\right) \text{ p.s.}$$

d'où le théorème 3.8 est la conséquence immédiate du théorème 3.6. \square

Conclusion

Dans ce travail, nous nous sommes intéressés aux modèles ARCH introduits initialement par Engle en 1982 et leurs extensions GARCH, IGARCH et FIGARCH.

Plus précisément, nous avons traité les questions liées à la stationnarité faible et forte ainsi qu'à l'estimation des paramètres, et ce, en étudiant la consistance et la normalité asymptotique.

Aussi, nous avons abordé la représentation $ARCH(\infty)$, en donnant des conditions d'existence et d'unicité d'une telle représentation.

Comme perspectives futures, il serait intéressant d'élargir notre étude aux modèles GARCH multivarié. Il y'a lieu de développer des relations de récurrences plus faciles à mettre en oeuvre en pratique pour les représentation $ARCH(\infty)$.

Bibliographie

Bollerslev, T. (1986). *Generalized autoregressive conditional heteroskedasticity*. J. Econometrics, **31**, 307-327.

Baillie, R.T., Bollerslev T. et Mikkelsen H.O., (1996), *Fractionally integrated generalized autoregressive conditional heteroscedasticity*, Journal of Econometrics, **74**, 3-30.

Baillie, R.T., Bollerslev T. et Mikkelsen H.O., (1996), *Fractionally integrated generalized autoregressive conditional heteroscedasticity*, Journal of Econometrics, **74**, 3-30.

Berkes, I., Horvath, L., & Kokoszka, P. (2003). GARCH processes : Structure and estimation, Bernoulli 9, 201–227.

Bougerol, N. et Picard, N.M., (1992), *Stationarity of GARCH processes and of some nonnegative time series*, Journal of Econometrics, Vol. **52**, 115-127

Brockwell, P. J., & Davis, R. A. (2013). Time series : theory and methods. Springer Science & Business Media.

Engle, R.F. (1982), *AutoRegressive Conditional Heteroskedasticity with Estimates of the Variance of U.K. Inflation*, Econometrica, **50**, 987-1008.

Engle, R.F., et T. Bollerslev, (1986), *Modelling the persistence of conditional variances*, *Econometric reviews*, Vol. **5**, 1-50.

Francq, C., & Zakoian, J. M. (2011). GARCH models : structure, statistical inference and financial applications. John Wiley & Sons.

Francq, C., & Zakoian, J. M. (2004). Maximum likelihood estimation of pure GARCH and ARMA-GARCH processes. Bernoulli, 10(4), 605-637.

Francq, C., & Zakoian, J. M. (2007). Quasi-maximum likelihood estimation in GARCH processes when some coefficients are equal to zero. Stochastic Processes and their Applications, 117(9), 1265-1284.