

République Algérienne Démocratique et Populaire.
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique.
Université de Mouloud Mammeri, Tizi-Ouzou

Faculté des Sciences
Département des Mathématiques

Mémoire de Master en

Mathématiques
Spécialité
Modélisation mathématique

Thème

Sur les modèles M-GARCH

Présenté par:

MESLOUB Ali

Devant le jury d'examen composé de :

BEDOUHENE-KHELLAS Fazia	Professeur	UMMTO	Président
HAMAZ Abdelghani	Maître de conférences B	UMMTO	Rapporteur
ACHEMINE-MERAKEB Farida	Maître assistante A	UMMTO	Examinatrice

Soutenu le - / - / 2017

Rremerciements

je remercie, avant tout, le bon Dieu de m'avoir donné le courage et la volonté, pour l'élaboration de ce modeste travail.

je tiens à exprimer ma profonde gratitude et mes sincères remerciements à mon chère promoteur, *M^r* HAMAZ.G qui ma a fait l'honneur de diriger ce travail, sa gentillesse et ses précieux conseils furent d'un apport considérable.

Mes remerciements vont également au président ainsi qu'aux membres de jury pour l'honneur qu'ils me font de juger ce travail.

Sans oublier de remercier tous les enseignants ayant contribué à ma formation de près ou de loin.

Sans oublier de remercier tous les amis et camarades, qui ont contribué ne serait ce que par un sourire à mon égard.

Merci infiniment à tous.

Dédicace

C'est avec une pensée pleine de reconnaissance inspirée par la Générosité et la gentillesse
que je dédie ce modeste travail:

A mes chères parent pour leurs grand et généreux amour, leurs sacrifice, leur
compréhension et leur soutien;

A mes deux frère bien evidement;

A mes oncles et mes tantes sans oublier les cousins et les cousines;

A tous mes amis qui m'ont soutenu moralement;

A tous ceux qui m'ont aide, conseille, et a tous ceux que j'aime et qui m'aime .

Ali

Table des matières

1	Introduction	3
2	Processus ARCH et GARCH	5
2.1	Présentation du modèle ARCH et GARCH	5
2.1.1	Propriétés d'un processus ARCH	6
2.1.2	Étude de la stationnarité	9
3	processus GARCH multivarié	19
3.1	Introduction au modèle M-GARCH	19
3.2	Les différents modèles M-GARCH	20
3.2.1	Modèles GARCH vectoriels	20
3.2.2	Modèle BEKK	22
3.2.3	Comparaison entre le modèle BEKK et le modèle vectoriel	29
3.2.4	Etude de la stationnarité	30
3.2.5	Modèles à corrélations conditionnelles constantes(CCC)	33
3.2.6	Etude de la stationnarité des modèles CCC	37
4	Estimation des paramètres du GARCH multivarié	39
4.1	Estimation par le quasi-maximum de vraisemblance QMV	39
4.2	Estimation par la methode EbE (Equation-by-Equation)	40
4.3	consistance et normalité asymptotique de l'estimateur EbE	42
4.4	comparaison entre l'estimateur EbE et l'estimateur QMV	44
5	Etude par simulation sous le langage R	46
5.1	Etude de Monte-Carlo	46
5.2	Estimation des volatilités individuelles d'un DCC par EbEE	56
	Conclusion	63

Annexe	64
5.3 Martingale	64
5.4 Ergodicité	64
5.5 Théorème de Chung-Fuchs	65
5.6 Critère de Cauchy	65
5.7 La Régression Linéaire Multiple	65
5.7.1 Notations	66
Bibliographie	67

Chapitre 1

Introduction

Les modèles GARCH ont été introduits pour rendre compte des propriétés des séries financières, et en particulier de leur hétéroscédasticité. Engle (1982) a proposé le modèle ARCH(q). Ce modèle est celui qui a rencontré le plus de succès dans la littérature. Bollerslev(1986) a ensuite généralisé le modèle ARCH(q) : c'est le modèle ARCH(p,q) généralisé ou GARCH((p,q)). Le modèle GARCH(p,q) univarié consiste à écrire la variance conditionnelle du processus GARCH en fonction de son passé.

Bien que dans le cas univarié le modèle GARCH ait été l'objet de nombreuses études, le cas multivarié reste encore peu étudié. Dans le cas multivarié, on considère un vecteur $\varepsilon_t = (\varepsilon_{1,t}, \dots, \varepsilon_{n,t})'$ et la matrice de covariance conditionnelle H_t du processus $\{\varepsilon_t\}_{t \in \mathbb{N}}$ à la date t .

Dans les modèles GARCH multivariés, on exprime cette matrice de covariance conditionnelle en fonction du passé du processus $\{\varepsilon_t\}_{t \in \mathbb{N}}$ et des propres valeurs passées de H_t . Une première approche consiste à considérer un processus GARCH multivarié $\{\varepsilon_t\}_{t \in \mathbb{N}}$ ayant pour composantes des processus GARCH univariés indépendants entre eux. Mais cette façon de traiter le problème est restrictive. En effet, l'étude des séries financières a mis en évidence une corrélation non nulle entre les composantes des processus GARCH multivariés. Ainsi, différentes formulations tenant compte des corrélations entre les composantes d'un processus GARCH multivarié (ou MGARCH) ont été introduites. Le but de ce mémoire est de présenter les processus ARCH et GARCH-univarié, ensuite présenter les différents modèles GARCH-multivariés. A travers ces approches pour représenter l'hétéroscédasticité conditionnelle des séries financières.

Ce travail est organisé comme suit. Dans la première partie, nous soulignons l'intérêt des processus ARCH et nous donnons les principales définitions et propriétés statistiques de ces processus (existence des solutions stationnaires, des représentations, la notion de station-

narité faible et forte). Dans la seconde partie nous allons présenter les différents modèles GARCH-multivariés et leurs propriétés statistiques. Ensuite, nous estimerons ces processus avec l'estimateur QMV (quasi-maximum de vraisemblance) et l'estimateur EbE (Equation-by-Equation) la méthode étape par étape et nous comparerons les deux estimateurs suivis de quelques exemples simulés avec le langage R.

Chapitre 2

Processus ARCH et GARCH

2.1 Présentation du modèle ARCH et GARCH

Dans un premier temps, nous donnons la définition du processus ARCH et GARCH fondée sur les deux premiers moments de ε_t conditionnels à son passé, et quelques propriétés relatives aux moments conditionnels et non conditionnels.

Dans la suite de ce chapitre, nous désignons par $(\Omega, \mathcal{F}, P)$ un espace probabilisé, et $\mathcal{F}_{t-1}$ la σ -algèbre engendrée par tout le passé du processus ε_s , pour $s < t$, i.e $\mathcal{F}_{t-1} = \sigma(\varepsilon_s, s < t)$.

Définition 2.1.1. Un processus $(\varepsilon_t)_{t \in \mathbb{Z}}$ est défini comme étant un processus ARCH(q) s'il vérifie l'équation suivante :

$$\begin{cases} \varepsilon_t = \sigma_t \eta_t \\ \sigma_t^2 = \omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 = \omega + \alpha(B) \varepsilon_t^2 \end{cases} \quad (2.1)$$

où $\omega > 0$, $\alpha_i \geq 0$, $i = 1, \dots, q$, et B est l'opérateur retard tel que $\alpha(B) = \sum_{i=1}^q \alpha_i B^i$ avec

$$B^i \varepsilon_t^2 = \varepsilon_{t-i}^2.$$

$(\eta_t)_t$ désigne une suite de variables aléatoires indépendantes, identiquement distribuées (*iid*), centrée de variance unité.

$\{\sigma_t\}_{t \in \mathbb{Z}}$ désigne une suite de variables telles que :

- σ_t est mesurable par rapport à une tribu, notée $\mathcal{F}_{t-1}$ engendrée par le passé de ε_t .
- η_t est indépendant de $\mathcal{F}_{t-1}$.
- $\sigma_t > 0$.

Le processus d'innovation pour ε_t^2 est par définition $\nu_t = \varepsilon_t^2 - \mathbb{E}\{\varepsilon_t^2 | \varepsilon_j, j \leq t-1\} = \varepsilon_t^2 - \sigma_t^2$,

qui vérifie $\mathbb{E}(\nu_t | \mathcal{F}_{t-1}) = 0$, on remplace σ_t^2 dans l'équation (2.1) par $\varepsilon_t^2 - \nu_t$ on aura :

$$\begin{aligned} \varepsilon_t^2 - \nu_t &= \omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2, \\ \Rightarrow \\ \varepsilon_t^2 &= \omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \nu_t, \end{aligned} \quad (2.2)$$

on obtient alors la représentation autorégressive AR(q) pour ε_t^2 .

2.1.1 Propriétés d'un processus ARCH

Le processus ε_t ARCH défini par l'équation (2.1) possède les propriétés statistiques suivantes :

Propriétés 2.1.1. *Le processus $(\varepsilon_t)_{t \in \mathbb{Z}}$ ARCH(q) défini par l'équation (2.1) est une différence de martingale homoscedastique :*

$$\mathbb{E}(\varepsilon_t | \mathcal{F}_{t-1}) = 0 \quad , \quad \mathbf{V}(\varepsilon_t) = \frac{w}{1 - \sum_{i=1}^q \alpha_i}.$$

Cette propriété signifie que le processus ARCH ε_t qui peut s'apparenter à un processus de bruit blanc (faible), ce qui explique notamment que l'on spécifiera des erreurs de modèles sous la forme ARCH. On retrouve alors toutes les propriétés de modèles établies sous la propriété de bruit blanc des erreurs. Mais cette propriété signifie en outre que le processus ARCH ε_t est non conditionnellement homoscedastique.

Preuve. Pour démontrer cela, nous utilisons l'équation précédente (2.1)

1. $\mathbb{E}(\varepsilon_t | \mathcal{F}_{t-1}) = \mathbb{E}(\sigma_t \eta_t | \mathcal{F}_{t-1}) = \sigma_t \mathbb{E}(\eta_t | \mathcal{F}_{t-1}) = \sigma_t \mathbb{E}(\eta_t) = 0$,
car σ_t est mesurable par rapport à la tribu $\mathcal{F}_{t-1}$, et η_t est indépendante de $\mathcal{F}_{t-1}$, et $\mathbb{E}(\eta_t) = 0$.

$$\begin{aligned} 2. \quad \mathbf{V}(\varepsilon_t) &= \mathbb{E}(\varepsilon_t^2) = \mathbb{E}(\sigma_t^2 \eta_t^2) = \mathbb{E}(\sigma_t^2) = w + \sum_{i=1}^q \alpha_i \mathbb{E}(\varepsilon_{t-i}^2) = w + \sum_{i=1}^q \alpha_i \mathbb{E}(\varepsilon_t^2), \\ \mathbb{E}(\varepsilon_t^2) \left(1 - \sum_{i=1}^q \alpha_i\right) &= w \quad \Rightarrow \quad \mathbb{E}(\varepsilon_t^2) = \frac{w}{1 - \sum_{i=1}^q \alpha_i}. \end{aligned}$$

□

Propriétés 2.1.2. [La variance conditionnelle] C'est la propriété centrale des processus ARCH. Le processus $(\varepsilon_t)_{t \in \mathbb{Z}}$ défini par l'équation (2.1) a une variance conditionnelle dépend du temps et vérifie :

$$\mathbb{E}(\varepsilon_t^2 | \mathcal{F}_{t-1}) = \sigma_t^2 \quad \forall t.$$

Preuve. $\mathbb{E}(\varepsilon_t^2 | \mathcal{F}_{t-1}) = \sigma_t^2 \mathbb{E}(\eta_t^2 | \mathcal{F}_{t-1}) = w + \sum_{i=1}^q \alpha_i \varepsilon_{t-1}^2$, (car $\mathbb{E}(\eta_t^2 | \mathcal{F}_{t-1}) = \mathbb{E}(\eta_t^2) = 1$) . □

Propriétés 2.1.3. [Les auto-covariances conditionnelles] Les auto-covariances conditionnelles du processus $(\varepsilon_t)_{t \in \mathbb{Z}}$ sont nulles. C'est à dire que :

$$\text{cov}(\varepsilon_t, \varepsilon_{t+k} | \mathcal{F}_{t-h}) = 0 \quad \forall h \geq 1, \quad \forall k \geq 1.$$

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

$$\begin{aligned} \text{cov}(\varepsilon_t, \varepsilon_{t+k} | \mathcal{F}_{t-h}) &= \mathbb{E}(\varepsilon_t \varepsilon_{t+k} | \mathcal{F}_{t-h}) - \mathbb{E}(\varepsilon_t | \mathcal{F}_{t-h}) \mathbb{E}(\varepsilon_{t+k} | \mathcal{F}_{t-h}) = \mathbb{E}(\varepsilon_t \varepsilon_{t+k} | \mathcal{F}_{t-h}) \\ &= \mathbb{E}[\mathbb{E}(\varepsilon_t \varepsilon_{t+k} | \mathcal{F}_{t+k-1}) | \mathcal{F}_{t-h}] \\ &= \mathbb{E}[\varepsilon_t \mathbb{E}(\varepsilon_{t+k} | \mathcal{F}_{t+k-1}) | \mathcal{F}_{t-h}] = \mathbb{E}(\varepsilon_t \times 0 | \mathcal{F}_{t-h}) = 0. \end{aligned}$$

□

Propriétés 2.1.4. [Moment centré d'ordre quatre] Les modèles ARCH permettent d'avoir des processus avec des queues de distribution plus épaisses. Le moment conditionnel centré d'ordre 4 du processus $(\varepsilon_t)_{t \in \mathbb{Z}}$ vérifie :

$$\mathbb{E}(\varepsilon_t^4 | \mathcal{F}_{t-1}) = 3(w + \alpha \varepsilon_{t-1}^2)^2,$$

sous l'hypothèse $3\alpha^2 < 1$ le moment non conditionnel centré d'ordre 4 du processus $(\varepsilon_t)_{t \in \mathbb{Z}}$ est égal à :

$$\mathbb{E}(\varepsilon_t^4) = \frac{3w^2(1+\alpha)}{(1-3\alpha^2)(1-\alpha)},$$

le kurtosis non conditionnelle associée au processus ARCH(1) est égale à :

$$k = \frac{\mathbb{E}(\varepsilon_t^4)}{\mathbb{E}(\varepsilon_t^2)^2} = 3 \left(\frac{1-\alpha^2}{1-3\alpha^2} \right) > 3.$$

Preuve. On a : $\varepsilon_t = \eta_t \sigma_t$,

et on sait que si une variable X centré et suit une loi normale alors : $\mathbb{E}(X^4) = 3\mathbb{E}(X^2)^2$,

$$\begin{aligned}
\mathbb{E}(\varepsilon_t^4) &= \mathbb{E}[\mathbb{E}(\varepsilon_t^4 | \mathcal{F}_{t-1})] \\
&= \mathbb{E}\left[3(w + \alpha\varepsilon_{t-1}^2)^2\right] \\
&= 3\mathbb{E}(w^2 + \alpha^2\varepsilon_{t-1}^4 + 2w\alpha\varepsilon_{t-1}^2) \\
&= 3[w^2 + \alpha^2\mathbb{E}(\varepsilon_{t-1}^4) + 2w\alpha\mathbb{E}(\varepsilon_{t-1}^2)] \\
&= 3\left[w^2 + \frac{2w^2\alpha}{1-\alpha} + \alpha^2\mathbb{E}(\varepsilon_{t-1}^4)\right]
\end{aligned}$$

$$\mathbb{E}(\varepsilon_t^4) = \frac{3w^2(1+\alpha)}{(1-3\alpha^2)(1-\alpha)}.$$

□

Remarque 2.1.1. Le kurtosis d'un processus ARCH est toujours supérieur à 3, la loi non conditionnelle d'un processus ARCH est donc une loi de distribution à queue épaisse, est donc plus aplatie qu'une gaussienne, on dit que cette distribution est leptokurtique .

Remarque 2.1.2. La moyenne non conditionnelle et les autocovariances d'un processus ARCH étant nulles, ce qui signifie que ce processus peut être caractérisé comme étant un processus bruit blanc faible.

Nous avons vu que le processus ARCH décrit précédemment dans la définition (2.1) admet une représentation autorégressive (AR) de la variance conditionnelle $(\sigma_t^2)_t$, donnée par l'équation (2.2). De manière similaire aux processus ARMA, il semble naturel de généraliser l'expression (2.2) en permettant à $(\sigma_t^2)_t$ de présenter une partie moyenne-mobile. Bollerslev (1986) a donc introduit les processus GARCH(p,q).

Définition 2.1.2. [Processus GARCH(p,q) semi-fort] On dit que $(\varepsilon_t)_{t \in \mathbb{Z}}$ est un processus GARCH(p,q) si ses deux premiers moments conditionnels existent et vérifient :

1. $\mathbb{E}(\varepsilon_t | \varepsilon_s, s < t) = 0, \quad t \in \mathbb{Z};$
2. Il existe des constantes $\omega, \alpha_i, i = 1, \dots, q$ et $\beta_j, j = 1, \dots, p$ telles que :

$$\sigma_t^2 = \mathbf{V}(\varepsilon_t | \varepsilon_s, s < t) = \omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2, \quad t \in \mathbb{Z}. \quad (2.3)$$

L'équation (2.3) est équivalente à :

$$\sigma_t^2 = \omega + \alpha(B) \varepsilon_t^2 + \beta(B) \sigma_t^2, \quad t \in \mathbb{Z}. \quad (2.4)$$

où B est l'opérateur retard, $\alpha(B) = \sum_{i=1}^q \alpha_i B^i$ avec $B^i \varepsilon_t^2 = \varepsilon_{t-i}^2$ et $\beta(B) = \sum_{j=1}^p \beta_j B^j$ avec $B^j \sigma_t^2 = \sigma_{t-j}^2$.

Le processus d'innovation pour ε_t^2 est par définition $\nu_t = \varepsilon_t^2 - \mathbb{E}\{\varepsilon_t^2 | \varepsilon_j, j \leq t-1\} = \varepsilon_t^2 - \sigma_t^2$, qui vérifie $\mathbb{E}(\nu_t | \mathcal{F}_{t-1}) = 0$, on remplace σ_t^2 dans l'équation (2.3) par $\varepsilon_t^2 - \nu_t$ on aura la structure linéaire d'un modèle ARMA :

$$\begin{aligned} \varepsilon_t^2 - \nu_t &= \omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j (\varepsilon_{t-j}^2 - \nu_{t-j}), \\ \Rightarrow \\ \varepsilon_t^2 &= \omega + \nu_t + \sum_{i=1}^r (\alpha_i + \beta_i) \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j \nu_{t-j}, \end{aligned}$$

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

Définition 2.1.3. [Processus GARCH(p, q) fort] Soit (η_t) une suite de variables *iid* de loi η , on dit que $(\varepsilon_t)_{t \in \mathbb{Z}}$ est un GARCH(p,q) au sens fort s'il vérifie l'équation suivante :

$$\begin{cases} \varepsilon_t = \sigma_t \eta_t \\ \sigma_t^2 = \omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2 \end{cases} \quad (2.5)$$

où $\omega > 0$, $\alpha_i \geq 0$, $i = 1, \dots, q$, $\beta_j \geq 0$, $j = 1, \dots, p$.

Maintenant en remplaçant ε_{t-i} par $\sigma_{t-i} \eta_{t-i}$ dans l'équation (2.5), on obtient une représentation autorégressive de σ_t^2 à coefficients aléatoires :

$$\begin{aligned} \sigma_t^2 &= \omega + \sum_{i=1}^q \alpha_i \sigma_{t-i}^2 \eta_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2, \\ \Leftrightarrow \\ \sigma_t^2 &= \omega + \sum_{i=1}^r (\alpha_i \eta_{t-i}^2 + \beta_i) \sigma_{t-i}^2 = \omega + \sum_{i=1}^r \alpha_i (\eta_{t-i}) \sigma_{t-i}^2, \end{aligned}$$

avec $\alpha_i(z) = \alpha_i z^2 + \beta_i$, et $r = \max(p, q)$.

2.1.2 Étude de la stationnarité

Dans cette partie, notre objectif est de trouver sous quelles conditions, il existe des processus stationnaires au sens strict et au second-ordre vérifiant les définitions (2.3) et/ou (2.5). On s'intéresse plus particulièrement aux solutions non anticipatives du processus

$(\varepsilon_t)_{t \in \mathbb{Z}}$ tel que ε_t soit une fonction mesurable des variables η_{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. Cas d'un GARCH(1,1)

Dans le cas où $p = q = 1$, le modèle (2.5) s'écrit :

$$\begin{cases} \varepsilon_t = \sigma_t \eta_t \\ \sigma_t^2 = \omega + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2 \end{cases} \quad (2.6)$$

en utilisant la représentation autorégressive on obtient :

$$\begin{cases} \varepsilon_t = \sigma_t \eta_t \\ \sigma_t^2 = \omega + \alpha(\eta_{t-1}) \sigma_{t-1}^2 \end{cases} \quad (2.7)$$

où $\omega > 0$, $\alpha \geq 0$, $\beta \geq 0$, et $\alpha(z) = \alpha z^2 + \beta$.

Théorème 2.1.1. [La stationnarité stricte du modèle GARCH(1,1)] *Si*

$$-\infty \leq \gamma = \mathbb{E} \{ \log (\alpha \eta_t^2 + \beta) \} < 0, \quad (2.8)$$

la série

$$h_t = \omega + \sum_{i=1}^{\infty} \alpha(\eta_{t-1}) \cdots \alpha(\eta_{t-i}) \omega, \quad (2.9)$$

converge presque sûrement et le processus (ε_t) défini par $\varepsilon_t = \sqrt{h_t} \eta_t$ est l'unique solution strictement stationnaire du modèle (2.6), cette solution est non anticipative et ergodique.

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

Remarque 2.1.3.

1. $\gamma = \mathbb{E} \{ \log \alpha(\eta_t) \}$ existe toujours dans $[-\infty, +\infty[$, car $\mathbb{E} (\log^+ \{ \alpha(\eta_t) \}) \leq \mathbb{E} \{ \alpha(\eta_t) \} = \alpha + \beta$.
2. Si $\alpha + \beta < 1 \Rightarrow \gamma < 0$, inversement si $\gamma < 0 \Rightarrow \alpha + \beta < 1$,
car si $\alpha + \beta < 1$, alors $\mathbb{E} \{ \log (\alpha \eta_t^2 + \beta) \} \leq \log \{ \mathbb{E} (\alpha \eta_t^2 + \beta) \} = \log (\alpha + \beta) < 0$.
Inversement si $\gamma < 0$, alors par absurde supposons que $\alpha + \beta \geq 1$ donc :
 $\mathbb{E} \{ \log [\alpha \eta_t^2 + \beta] \} = \mathbb{E} \left\{ \log \left[\beta \left(\frac{\alpha}{\beta} \eta_t^2 + 1 \right) \right] \right\} = \mathbb{E} \left\{ \log (\beta) + \log \left(\frac{\alpha}{\beta} \eta_t^2 + 1 \right) \right\} > 0$, d'où la contradiction.
3. Dans le cas ARCH(1) ($\beta = 0$), la contrainte de stationnarité stricte peut s'écrire comme suit

$$0 \leq \alpha < \exp \{ -\mathbb{E} (\log \eta_t^2) \}, \quad (2.10)$$

En effet,

$$\begin{aligned} -\infty \leq \gamma &= \{\mathbb{E}(\log \alpha + \log \eta_t^2)\} < 0 \\ \Leftrightarrow -\infty \leq \log \alpha &< -\mathbb{E}(\log \eta_t^2) \\ \Leftrightarrow 0 \leq \alpha &< \exp \{-\mathbb{E}(\log \eta_t^2)\}. \end{aligned}$$

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

4. Dans le cas où $\omega = 0$ et $\gamma < 0$, il est clair d'après (2.9) que la seule solution strictement stationnaire du modèle est $\varepsilon_t = 0$.

Preuve. En utilisant la deuxième équation du modèle (2.7), et par itération, on aura pour $N \geq 1$:

$$\begin{aligned} \sigma_t^2 &= \omega + \alpha(\eta_{t-1})\sigma_{t-1}^2 \\ &= \omega \left\{ 1 + \sum_{n=1}^N \alpha(\eta_{t-1}) \cdots \alpha(\eta_{t-n}) \right\} + \alpha(\eta_{t-1}) \cdots \alpha(\eta_{t-N-1}) \sigma_{t-N-1}^2 \\ &:= h_t(N) + \alpha(\eta_{t-1}) \cdots \alpha(\eta_{t-N-1}) \sigma_{t-N-1}^2. \end{aligned} \quad (2.11)$$

Soit $h_t = \lim_{N \rightarrow +\infty} h_t(N) \in [0, +\infty]$, puisque les termes sont positifs.

De plus $h_t(N) = \omega + \alpha(\eta_{t-1})h_{t-1}(N-1)$, donc quand $N \rightarrow +\infty$, on obtient :

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

Maintenant montrons que le processus limite (h_t) est à valeurs finies si et seulement si $\gamma < 0$.

En utilisant la règle de Cauchy pour les séries à termes positifs, on a :

$$h_t = \omega \{1 + \alpha(\eta_{t-1}) \cdots \alpha(\eta_{t-n})\},$$

donc

$$\begin{aligned} \{\alpha(\eta_{t-1}) \cdots \alpha(\eta_{t-n})\}^{1/n} &= \exp \left\{ \frac{1}{n} \log (\alpha(\eta_{t-1}) \cdots \alpha(\eta_{t-n})) \right\} \\ &= \exp \left\{ \frac{1}{n} \sum_{i=1}^n \log [\alpha(\eta_{t-i})] \right\} \\ &\rightarrow \exp \{ \mathbb{E} [\log (\alpha \eta_{t-i}^2)] \} = \exp(\gamma) \quad p.s. \end{aligned} \quad (2.12)$$

par l'application de la loi forte des grands nombres, quand $n \rightarrow +\infty$, à la suite *iid* $(\log \{\alpha(\eta_{t-i})\})$, la série h_t converge presque sûrement dans $\mathbb{R}$, par application de la règle de Cauchy, et le processus limite, (h_t) , est à valeurs réelles positives. Par suite le processus (ε_t)

défini par $\varepsilon_t = \sqrt{h_t}\eta_t$, est strictement stationnaire et ergodique (théorème d'ergodicité), et est non anticipatif car il s'écrit comme fonction mesurable des variables η_{t-i} , $i \geq 0$, de plus (ε_t) vérifie le modèle (2.6).

Montrons l'unicité :

Soit $\tilde{\varepsilon}_t = \sigma_t \eta_t$ une autre solution strictement stationnaire. D'après (1.11), on a :

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

Par suite

$$\sigma_t^2 - h_t = \{h_t(N) - h_t\} + \alpha(\eta_{t-1}) \cdots \alpha(\eta_{t-N-1}) \sigma_{t-N-1}^2, \quad \forall N.$$

Lorsque $\gamma < 0$, le premier terme qui est entre accolade tend vers 0 *p.s* quand $N \rightarrow \infty$, donc le second terme (qui ne dépend pas de N) est nul, car la série h_t converge *p.s* et de plus la loi de σ_{t-N-1}^2 ne dépend pas aussi de N par stationnarité. On a montré alors que $\sigma_t^2 = h_t$ *p.s*.

Si $\gamma > 0$, d'après (1.12) et la règle de Cauchy, $\sum_{n=1}^N \alpha(\eta_{t-1}) \cdots \alpha(\eta_{t-n}) \rightarrow +\infty$, *p.s* lorsque $N \rightarrow \infty$. Donc si $\omega > 0$, $h_t = +\infty$, *p.s*. D'après (1.11) il est clair que alors $\sigma_t^2 = +\infty$ *p.s*. Par suite il n'existe pas de solution finie de (2.6).

Dans le cas $\gamma = 0$, nous procéderons par l'absurde. Supposons qu'il existe une solution strictement stationnaire $(\varepsilon_t, \sigma_t^2)$ de (2.6). Nous avons pour $n > 0$,

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

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

$$\sum_{i=1}^n \log \alpha(\eta_i) + \log \omega \rightarrow -\infty \quad \text{p.s.} \quad \text{quand} \quad n \rightarrow \infty. \quad (2.13)$$

d'après le théorème de Chung-Fuchs nous avons $\limsup \sum_{i=1}^n \log \alpha(\eta_i) = +\infty$ avec probabilité 1, ce qui contredit (2.13).

□

Théorème 2.1.2. [La stationnarité au second ordre du modèle GARCH(1,1)] *Si $\omega > 0$ et $\alpha + \beta < 1$, le processus $(\varepsilon_t)_{t \in \mathbb{Z}}$ admet une solution stationnaire au second ordre, de plus elle est unique, plus précisément (ε_t) est un bruit blanc.*

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

Preuve. Si ε_t est stationnaire au second-ordre et non anticipatif,

$$\mathbb{E}(\varepsilon_t^2) = \mathbb{E}(\sigma_t^2 \eta_t^2) = \mathbb{E}(\sigma_t^2) = \omega + \alpha \mathbb{E}(\varepsilon_{t-1}^2) + \beta \mathbb{E}(\sigma_{t-1}^2).$$

soit

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

Il faut donc $\alpha + \beta < 1$. On obtient de plus : $\mathbb{E}(\varepsilon_t^2) > 0$.

Inversement si $\alpha + \beta < 1$ on a $\gamma < 0$, alors la solution strictement stationnaire vérifie :

$$\begin{aligned} \mathbb{E}(\varepsilon_t^2) &= \mathbb{E}(h_t) \\ &= \left\{ 1 + \sum_{n=1}^{\infty} \mathbb{E}\{\alpha(\eta_{t-1}) \cdots \alpha(\eta_{t-n})\} \right\} \omega \\ &= \left\{ 1 + \sum_{n=1}^{\infty} \{\mathbb{E}\alpha(\eta_t)\}^n \right\} \omega \\ &= \left\{ 1 + \sum_{n=1}^{\infty} \{\alpha + \beta\}^n \right\} \omega \\ &= \frac{\omega}{1 - \alpha - \beta}. \end{aligned}$$

□

2. Cas d'un GARCH(p,q)

Dans le cas général de GARCH(p,q) fort, la représentation vectorielle sera très utile.

On a

$$Z_t = b_t + A_t Z_{t-1}, \quad (2.14)$$

où

$$Z'_t = (\varepsilon_t^2, \varepsilon_{t-1}^2, \dots, \varepsilon_{t-q+1}^2, \sigma_t^2, \dots, \sigma_{t-p+1}^2) \in \mathbb{R}^{p+q},$$

$$b'_t = (\omega \eta_t^2, 0, \dots, 0, \omega, 0, \dots, 0) \in \mathbb{R}^{p+q},$$

$$A_t = \begin{pmatrix} \alpha_1 \eta_t^2 & \cdots & \alpha_q \eta_t^2 & \beta_1 \eta_t^2 & \cdots & \beta_p \eta_t^2 \\ & I_{q-1} & 0 & & 0 & \\ \alpha_1 & \cdots & \alpha_q & \beta_1 & \cdots & \beta_p \\ & 0 & & I_{p-1} & 0 & \end{pmatrix}. \quad (2.15)$$

A_t est une matrice de dimension $(p+q) \times (p+q)$. Dans le cas ARCH(q), Z_t ne contient que ε_t^2 et ses $q-1$ premières valeurs passées, et A_t se limite au bloc supérieur gauche de

la matrice ci-dessus. L'équation (2.14) constitue un modèle vectoriel autorégressif d'ordre un, avec coefficients positifs et *iid*.

Si on déroule le modèle (2.14) on obtient :

$$Z_t = b_t + \sum_{k=1}^{\infty} A_t A_{t-1} \cdots A_{t-k+1} b_{t-k}. \quad (2.16)$$

sous réserve que la série existe au sens presque sûr. L'objet de ce qui suit est de trouver des conditions justifiant l'existence de cette série. Lorsque le membre de droite de l'équation (2.16) a un sens, cela n'assure pas pour autant que les composantes de ce vecteur sont positives. Une condition suffisante pour que, presque sûrement,

$$b_t + \sum_{k=1}^{\infty} A_t A_{t-1} \cdots A_{t-k+1} b_{t-k} > 0, \quad (2.17)$$

au sens où toutes les composantes de ce vecteur sont strictement positives (éventuellement infinies), est évidemment

$$\omega > 0, \quad \alpha_i \geq 0 \quad (i = 1, \dots, q), \quad \beta_j \geq 0 \quad (j = 1, \dots, p). \quad (2.18)$$

La stationnarité stricte

L'outil principale pour l'étude de la stationnarité stricte est le concept d'exposant de Lyapounov. Soit A une matrice $(p+q) \times (q+q)$. Son rayon spectral, noté $\rho(A)$, est le plus grand module de ses valeurs propres. Soit $\| \cdot \|$ une norme quelconque sur l'espace des matrices $(p+q) \times (q+q)$. On a le résultat d'algèbre suivant :

$$\lim_{t \rightarrow \infty} \frac{1}{t} \log \| A^t \| = \log \rho(A), \quad t \in \mathbb{N}. \quad (2.19)$$

Théorème 2.1.3. [Exposant de Lyapounov] *Soit $\{A_t\}_{t \in \mathbb{Z}}$ une suite de matrices aléatoires, strictement stationnaire et ergodique, telle que $\mathbb{E} \log^+ \|A_t\| < \infty$. On a :*

$$\lim_{t \rightarrow +\infty} \frac{1}{t} \mathbb{E} (\log \|A_t A_{t-1} \dots A_1\|) = \gamma = \inf_{t \in \mathbb{N}^*} \frac{1}{t} \mathbb{E} (\log \|A_t A_{t-1} \dots A_1\|), \quad (2.20)$$

et γ (resp. $\exp(\gamma)$) s'appelle plus grand exposant de Lyapounov (resp. rayon spectral) de la suite de matrices $\{A_t\}_{t \in \mathbb{Z}}$. De plus

$$\gamma = \lim_{t \rightarrow +\infty} p.s. \frac{1}{t} \log \|A_t A_{t-1} \dots A_1\|. \quad (2.21)$$

γ : l'exposant de Lyapounov.

$\exp\{\gamma\}$: Rayon spectral de la suite $\{A_t\}_{t \in \mathbb{Z}}$.

Remarque 2.1.4.

1. On a toujours $\gamma \leq \mathbb{E}(\log \|A_1\|)$, avec égalité en dimension 1.
2. Si $A_t = A$ pour tout $t \in \mathbb{Z}$, on a $\gamma = \log \rho(A)$ d'après (2.19).
3. Toutes les normes étant équivalentes sur un espace de dimension fini, il est facile de voir que γ est indépendant du choix de la norme.

Le lemme général suivant est très utile pour l'étude du produit de matrices aléatoires.

Lemme 2.1.1. *Soit $\{A_t\}_{t \in \mathbb{Z}}$ une suite de matrices aléatoires iid telle que $\mathbb{E}[\log^+ \|A_t\|] < \infty$, et de plus grand exposant de Lyapounov γ . Alors :*

$$\lim_{t \rightarrow \infty} p.s. \|A_0 \dots A_{-t}\| = 0 \quad \Rightarrow \quad \gamma < 0. \quad (2.22)$$

Comme pour les modèles ARMA, nous nous intéressons plus particulièrement aux solutions (ε_t) non anticipatives du modèle (2.5), c'est à dire telles que ε_t appartient à la tribu engendrée par $\{\eta_t, \eta_{t-1}, \dots\}$.

Théorème 2.1.4. [La stationnarité stricte du modèle GARCH(p,q)] *Le modèle GARCH (p,q) a une unique solution strictement stationnaire non anticipative et ergodique si et seulement si $\gamma < 0$, où γ est le plus grand exposant de Lyapounov de la suite $\{A_t\}_{t \in \mathbb{Z}}$.*

Preuve. Nous utiliserons la norme définie par $\|A_t\| = (\sum |a_{ij}|)$. Par commodité la norme sera notée de manière identique quelle que soit la dimension de A . Avec cette convention, la norme est clairement multiplicative: $\|AB\| \leq \|A\| \|B\|$ pour toutes matrices A et B telles que AB existe.

Remarquons que, les variables η_t étant de variance finie, tous les termes de la matrice A_t sont intégrables. On a donc

$$\mathbb{E}[\log^+ \|A_t\|] \leq \mathbb{E}\|A_t\| = \mathbb{E}\left[\sum |a_{ij}|\right] < \infty.$$

Supposons $\gamma < 0$. Alors, l'égalité (2.21) implique que la série :

$$Z_t = b_t + \sum_{n=1}^{\infty} A_t A_{t-1} \dots A_{t-n} b_{t-n-1},$$

converge presque sûrement pour tout t . En effet, en utilisant la multiplicativité de la norme,

$$\|Z_t\| \leq \|b_t\| + \sum_{n=1}^{\infty} \|A_t A_{t-1} \cdots A_{t-n}\| \|b_{t-n-1}\| \quad (2.23)$$

et

$$\begin{aligned} (\|A_t A_{t-1} \cdots A_{t-n}\| \|b_{t-n-1}\|)^{1/n} &= \exp \left\{ \frac{1}{n} \log(\|A_t A_{t-1} \cdots A_{t-n}\|) + \frac{1}{n} \log \|b_{t-n-1}\| \right\} \\ &\xrightarrow{p.s.} \exp(\gamma) < 1 \quad (\text{car } \mathbb{E}|\log \|b_{t-n-1}\|| < \infty). \end{aligned}$$

Par application de la règle de Cauchy, $\gamma < 0$ implique Z_t est bien défini. Soit $Z_{q+1,t}$ la $q+1$ -ème composante de Z_t . En posant $\varepsilon_t = \sqrt{Z_{q+1,t}} \eta_t$ on définit une solution strictement stationnaire du modèle (2.5). D'après (2.16) ε_t s'exprime comme fonction mesurable de $\eta_t, \eta_{t-1}, \dots$. La solution est donc non anticipative et ergodique puisque (η_t) est ergodique.

L'unicité se démontre par le même raisonnement que dans le cas $p = q = 1$.

Supposons qu'il existe une autre solution strictement stationnaire Z_t^* du modèle (2.14).

Alors, pour tout $n \geq 0$,

$$\begin{aligned} \|Z_t - Z_t^*\| &= \|b_t + A_t Z_{t-1} - (b_t + A_t Z_{t-1}^*)\| \\ &\leq \|A_t A_{t-1} \cdots A_{t-n}\| \|Z_{t-n-1} - Z_{t-n-1}^*\|. \end{aligned}$$

Par absurde, $P\{\|Z_t - Z_t^*\| \neq 0\} > 0$, or on sait que

$$\|A_t A_{t-1} \cdots A_{t-n}\| \xrightarrow{p.s.} 0,$$

par suite

$$P\{\|Z_{t-n-1} - Z_{t-n-1}^*\| \longrightarrow \infty\} > 0,$$

ce qui implique que $\|Z_{t-n-1}\| \longrightarrow \infty$ ou $\|Z_{t-n-1}^*\| \longrightarrow \infty$ avec une probabilité positive. Ceci est impossible car les suites $(Z_t)_t$ et $(Z_t^*)_t$ sont stationnaires. On en conclut que $Z_t = Z_t^*$ pour tout t , *p.s.*

Finalement Nous montrons la partie nécessaire du théorème. D'après le lemme (2.1.1), il suffit d'établir (2.22). Nous allons montrer que, pour $1 \leq i \leq p + q$

$$\lim_{t \rightarrow \infty} A_0 \cdots A_{-t} e_i = 0, \quad p.s. \quad (2.24)$$

où e_i est le i -ème élément de la base canonique de $\mathbb{R}^{p+q}$. Soit (ε_t) une solution strictement

stationnaire de (2.5) et soit (Z_t) défini par (2.14). On a pour $t > 0$

$$\begin{aligned} Z_0 &= b_0 + A_0 Z_{-1} \\ &= b_0 + \sum_{k=0}^{t-1} A_0 \cdots A_{-k} b_{-k-1} + A_0 \cdots A_{-t} Z_{-t-1}, \\ &\geq \sum_{k=0}^{t-1} A_0 \cdots A_{-k} b_{-k-1} \end{aligned}$$

car les coefficients des matrices A_t , b_0 et z_t sont positifs. La série $A_0 \cdots A_{-k} b_{-k-1}$ tend presque sûrement vers 0 quand $k \rightarrow \infty$. Or $b_{-k-1} = \omega \eta_{-k-1}^2 e_1 + \omega e_{q+1}$, donc $A_0 \cdots A_{-k} b_{-k-1}$ se décompose en deux termes positifs et on a :

$$\lim_{t \rightarrow \infty} A_0 \cdots A_{-k} \omega \eta_{-k-1}^2 e_1 = 0, \quad \lim_{t \rightarrow \infty} A_0 \cdots A_{-k} \omega e_{q+1} = 0, \quad p.s. \quad (2.25)$$

Puisque $\omega \neq 0$, (2.24) est vraie pour $i = q + 1$. En utilisant la relation

$$A_{-k} e_{q+i} = \beta_i \eta_{-k}^2 e_1 + \beta_i e_{q+1} + e_{q+i+1}, \quad i = 1, \dots, p \quad (2.26)$$

avec par convention $e_{p+q+1} = 0$, pour $i = 1$ on obtient :

$$0 = \lim_{t \rightarrow \infty} A_0 \cdots A_{-k} e_{q+1} \geq \lim_{k \rightarrow \infty} A_0 \cdots A_{-k+1} e_{q+2} \geq 0,$$

donc (2.24) est vraie pour $i = q + 2$, et par récurrence, pour $i = q + j$, $j = 1, \dots, p$, en utilisant (2.26). Par ailleurs, on remarque que $A_{-k} e_q = \alpha_q \eta_{-k}^2 e_1 + \alpha_q e_{q+1}$, ce qui permet de voir, d'après (2.25), que (2.24) est vérifiée pour $i = q$. On conclut pour les autres valeurs de i en utilisant

$$A_{-k} e_i = \alpha_i \eta_{-k}^2 e_1 + \alpha_i e_{i+1} + e_{i+1}, \quad i = 1, \dots, q - 1,$$

et une récurrence ascendante. Le théorème (2.1.4) est donc démontré. □

Corollaire 2.1.1. [Conséquences de stationnarité stricte] Soit γ le plus grand exposant de Lyapounov de la suite $\{A_t, t \in \mathbb{Z}\}$ définie par (2.15). Si $\gamma < 0$ nous avons les propriétés équivalentes suivantes :

- (i) $\sum_{j=1}^p \beta_j < 1$,
- (ii) $1 - \beta_1 z - \dots - \beta_p z^p = 0 \Rightarrow |z| > 1$,

(iii) $\rho(B) < 1$, où B est la sous-matrice de A_t définie par :

$$B = \begin{pmatrix} \beta_1 & \beta_2 & \cdots & \beta_p \\ 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & & \cdots & 1 & 0 \end{pmatrix}.$$

Preuve. Comme tous les termes des matrices A_t sont positifs, il est clair que γ est supérieur au coefficient de Lyapounov de la suite obtenue en remplaçant les coefficients des q premières lignes et des q premières colonnes par 0 dans les matrices A_t . En utilisant la Remarque 2 du Théorème (2.1.3) on voit que

$$\gamma \geq \log \rho(B).$$

Par suite $\gamma < 0 \Rightarrow$ (iii). Il est facile de montrer (par récurrence sur p et en développant par rapport à la dernière colonne) que, pour $\lambda \neq 0$,

$$\det(B - \lambda I_p) = (-1)^p \{ \lambda^p - \lambda^{p-1} \beta_1 - \cdots - \lambda \beta_{p-1} - \beta_p \} = (-\lambda)^p \mathcal{B}\left(\frac{1}{\lambda}\right),$$

où $\mathcal{B}(z) = 1 - \beta_1 z - \cdots - \beta_p z^p$. On en déduit que si $\gamma < 0$ alors $\mathcal{B}(z) = 0$ a toutes ses racines en dehors du cercle unité, d'où l'équivalence entre (ii) et (iii).

A présent, montrons que (i) $\Leftrightarrow$ (ii). On a $\mathcal{B}(0) = 1$ et $\mathcal{B}(1) = 1 - \sum_{j=1}^p \beta_j$, donc si $\sum_{j=1}^p \beta_j \geq 1$, alors $\mathcal{B}(1) \leq 0$ et, par continuité, il existe une racine dans $]0,1]$. Ainsi (ii) $\Rightarrow$ (i)

Inversement si $\sum_{j=1}^p \beta_j < 1$ et si $\mathcal{B}(z_0) = 0$ pour un z_0 de module inférieur ou égal à 1 alors

$$1 = \sum_{j=1}^p \beta_j z_0^j = \left| \sum_{j=1}^p \beta_j z_0^j \right| \leq \sum_{j=1}^p \beta_j |z_0|^j \leq \sum_{j=1}^p \beta_j, \text{ ce qui est impossible, par suite (i) } \Rightarrow \text{(ii).}$$

□

Chapitre 3

processus GARCH multivarié

3.1 Introduction au modèle M-GARCH

Bien que dans le cas univarié le modèle GARCH ait été l'objet de nombreuses études, le cas multivarié reste encore peu étudié. Dans le cas multivarié, on considère un vecteur $\varepsilon_t = (\varepsilon_{1,t}, \dots, \varepsilon_{n,t})'$ et la matrice de covariance conditionnelle H_t du processus $\{\varepsilon_t\}_{t \in \mathbb{N}}$ à la date t de terme général $h_{ij,t}$:

$$H_t = E_{t-1}(\varepsilon_t \varepsilon_t').$$

Dans les modèles GARCH multivariés, on exprime cette matrice de covariance conditionnelle en fonction du passé du processus $\{\varepsilon_t\}_{t \in \mathbb{N}}$ et des propres valeurs passées de H_t . Une première approche consiste à considérer un processus GARCH multivarié $\{\varepsilon_t\}_{t \in \mathbb{N}}$ ayant pour composantes des processus GARCH univariés indépendants entre eux. Mais cette façon de traiter le problème est restrictive. En effet, l'étude des séries financières a mis en évidence une corrélation non nulle entre les composantes des processus GARCH multivariés. Ainsi, différentes formulations tenant compte des corrélations entre les composantes d'un processus GARCH multivarié (ou MGARCH) ont été introduites. De la même manière que dans le cas univarié, on dit que $\{\varepsilon_t\}_{t \in \mathbb{N}}$ est un processus MGARCH(p,q) si:

$$\begin{cases} \varepsilon_t = H_t^{1/2} \eta_t \\ g(H_t) = f(H_{t-1}, \dots, H_{t-p}, \varepsilon_{t-1}, \dots, \varepsilon_{t-p}) \end{cases}$$

avec $H_t^{1/2}$ une matrice telle que $H_t^{1/2} H_t^{1/2} = H_t$. Le processus $\{\eta_t\}_{t \in \mathbb{N}}$ est un processus de dimension n , iid tel que pour tout $t \in \mathbb{N}$ on a: $E(\eta_t) = 0$ et $E(\eta_t \eta_t') = \Gamma$, la matrice

Γ étant une matrice carrée de dimension n . Notons enfin que f est une transformation I_{t-1} mesurable et g une transformation I_t mesurable. On voit que la notion de GARCH fort n'a de sens que si $H_t^{1/2}$ existe et donc que H_t est définie positive. Ainsi f est une transformation qui doit assurer la définie positivité de la matrice H_t . Le but de ce chapitre est de présenter les différents modèles GARCH-multivariés. A travers ces approches pour représenter l'hétéroscédasticité conditionnelle des séries financières, nous étudierons les restrictions intrinsèques a ces modèles, et les restrictions motivées par l'identifiabilité de ces modèles et la définie positivité de la matrice H_t .

3.2 Les différents modèles M-GARCH

3.2.1 Modèles GARCH vectoriels

Le modèle vectoriel a été introduit par Engle et Kroner (1995), ce modèle est le plus simple pour représenter H_t par rapport à son passé. En effet, cette représentation présente une expression de H_t similaire à l'expression de la volatilité (variance conditionnelle) dans le cas univarié :

$$h_t = C_0 + \sum_{i=1}^q A_i e_{t-i} + \sum_{j=1}^p B_j h_{t-j}. \quad (3.1)$$

où $e_t = \text{vec}(\varepsilon_t \varepsilon_t')$, $h_t = \text{vec}(H_t)$, A_i, B_i matrices de dimension $n \times n$, et C_0 vecteur de dimension n . L'opérateur vec appliqué a une matrice consiste à mettre en une seule colonne les colonnes d'une matrice. Cependant la matrice H_t étant clairement symétrique certaines équations du modèle précédent sont redondantes, il est donc préférable d'utiliser l'opérateur $\text{vech}(\cdot)$ dans (3.1) qui met en seule colonne la partie triangulaire supérieure d'une matrice. Ainsi les dimensions des matrices A_i, B_i , sont de dimension $(n(n+1)) \setminus 2 \times (n(n+1)) \setminus 2$, et C_0 est un vecteur de dimension $(n(n+1)) \setminus 2$, le nombre de paramètres est ainsi considérablement réduit. On peut écrire le modèle en regroupant le vecteur C_0 et les matrices de paramètres dans une même matrice :

$$h_t = F z_t = (z_t' \otimes I) \text{vec}(F) = Z_t \alpha. \quad (3.2)$$

avec $z_t' = (1, \varepsilon_{t-1}', \dots, \varepsilon_{t-q}', h_{t-1}', \dots, h_{t-p}')$, $\alpha = \text{vec}(F)$, $F = (C_0 A_1 \dots A_q B_1 \dots B_p)$ et $Z_t = (z_t' \otimes I)$.

Pour illustrer le modèle vectoriel considérons le modèle $GARCH(1,1)$ avec deux équations ($n = 2$) en utilisant l'opérateur $vech(\cdot)$:

$$\begin{aligned}
 h_t = \begin{pmatrix} h_{11,t} \\ h_{21,t} \\ h_{22,t} \end{pmatrix} &= \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} + \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} \varepsilon_{1,t-1}^2 \\ \varepsilon_{1,t-1}\varepsilon_{2,t-1} \\ \varepsilon_{2,t-1}^2 \end{pmatrix} \\
 &+ \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix} \begin{pmatrix} h_{11,t-1} \\ h_{21,t-1} \\ h_{31,t-1} \end{pmatrix} \quad (3.3)
 \end{aligned}$$

Modèle Diagonal:

Dans certaines séries financières, on peut supposer que les volatilités et les covariances ne dépendent que de leurs valeurs passées dans le modèle précédent. Ainsi, $h_{rs,t}$ ne dépend que des valeurs passées de $\varepsilon_{r,t}\varepsilon_{s,t}$. Dans ce modèle les matrices A_i et B_j sont des matrices diagonales ce qui réduit le nombre de paramètres. Dans le cas de l'exemple précédent on obtient :

$$h_t = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} + \begin{pmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{pmatrix} \begin{pmatrix} \varepsilon_{1,t-1}^2 \\ \varepsilon_{1,t-1}\varepsilon_{2,t-1} \\ \varepsilon_{2,t-1}^2 \end{pmatrix} + \begin{pmatrix} b_{11} & 0 & 0 \\ 0 & b_{22} & 0 \\ 0 & 0 & b_{33} \end{pmatrix} \begin{pmatrix} h_{11,t-1} \\ h_{21,t-1} \\ h_{31,t-1} \end{pmatrix}.$$

On peut cependant remarquer que le modèle diagonal ne correspond pas au modèle où il n'existe pas de corrélation entre les composantes du processus $GARCH_{multivarié}$ car H_t n'est pas diagonale.

Remarque 3.1 La représentation vectorielle (et donc à priori le modèle diagonal) telle qu'elle est définie dans la partie précédente ne tient pas compte de la définie positivité nécessaire de H_t . Les hypothèses établissant la définie positivité pour H_t à chaque date t pour le modèle vectoriel ne sont pas claires. Ainsi, il est difficile de considérer un $GARCH$ fort pour les processus multivariés dans le cadre de ce modèle.

Généralité du modèle vectoriel

Définition 3.1 Un modèle est général quand tout les termes $\varepsilon_{r,t-i}\varepsilon_{s,t-i}$ et $h_{rs,t-i}$ peuvent apparaitre dans l'expression de chaque terme de la matrice de covariance H_t , de plus, aucune combinaison de valeurs entre les coefficients de chaque terme ne doit être écartée par l'expression de ce modèle.

D'après la définition précédente le modèle vectoriel est général. En effet chaque élément de la matrice H_t est une fonction linéaire des valeurs passées de $\varepsilon_{r,t-i}\varepsilon_{s,t-i}$ et de $h_{rs,t-i}$ dans le modèle vectoriel. Ainsi, grâce à sa linéarité la notion de généralité est facilement vérifiable pour le modèle vectoriel.

3.2.2 Modèle BEKK

Le modèle BEKK suivant a été introduit par Engle et Kroner (1995) pour imposer la restriction de définie positivité à la matrice H_t sous des conditions très faibles.

L'expression de la matrice H_t à chaque instant t dans le modèle BEKK nous est donnée par la relation suivante :

$$H_t = C'C + \sum_{k=1}^K \sum_{i=1}^q A'_{ik} \varepsilon_{t-i} \varepsilon'_{t-i} A_{ik} + \sum_{k=1}^K \sum_{i=1}^p B'_{jk} h_{t-j} B_{jk}. \quad (3.4)$$

où A_{ik} , B_{jk} , $k \in 1, \dots, K$ sont des matrices de dimension $n \times n$ et C est une matrice constante de dimension $n \times n$.

L'introduction de la somme sur l'entier k est motivée par un problème de généralité du modèle que nous aborderons plus loin.

Exemple 3.2 Considérons le modèle GARCH(1,1) avec $K = 1$. L'expression du modèle est la suivante :

$$H_t = C'C + A' \varepsilon_{t-1} \varepsilon'_{t-1} A + B' H_{t-1} B$$

Dans le cas où $n = 2$, la relation précédente devient :

$$\begin{aligned} H_t = C'C + & \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}' \begin{pmatrix} \varepsilon_{1,t-1}^2 & \varepsilon_{1,t-1}\varepsilon_{2,t-1} \\ \varepsilon_{2,t-1}\varepsilon_{1,t-1} & \varepsilon_{2,t-1}^2 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \\ & + \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}' \begin{pmatrix} h_{11,t-1} & h_{12,t-1} \\ h_{21,t-1} & h_{22,t-1} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \end{aligned} \quad (3.5)$$

On remarque que ce modèle nécessite moins de paramètres que le modèle GARCH(1,1) proposé en exemple pour le modèle vectoriel. Cependant, le nombre de paramètres croît quadratiquement avec n . Ainsi, il est préférable de réduire ce nombre, tout en imposant des restrictions nous assurant l'identifiabilité et la généralité du modèle. Engle et Kroner(1995) ont établi la proposition suivante qui nous donne des conditions que doit remplir le modèle pour que H_t soit définie positive $\forall t \in \mathbb{N}$.

Proposition 3.2 Si $H_0, H_{-1}, \dots, H_{-p+1}$ sont définies positives, alors H_t est définie positive dans le modèle BEKK si :

$$\ker\{C\} \bigcap \bigcap_{j=1}^p \bigcap_{k=1}^K \ker\{B_{jk}\} = \{0\}$$

On voit que les conditions pour obtenir une matrice H_t définie positive quelque soit $t \in \mathbb{N}$ sont faibles : la condition portant sur les noyaux des matrices C et B_{jk} $j \in \{1, \dots, n\}$ est vérifiée si la matrice C ou l'une des matrices B_{jk} est de plein rang. De plus on remarque dans cette proposition que les conditions portent sur les termes GARCH et la partie constante seulement de l'expression de la représentation BEKK. En effet la quantité $A'_{ik} \varepsilon_{t-1} \varepsilon'_{t-1} A_{ik}$ est semi-définie positive car $\varepsilon_{t-1} \varepsilon'_{t-1}$ est semi-définie positive. Ainsi, pour assurer la définie positivité de H_t il suffit que $C'C + \sum_{k=1}^K \sum_{i=1}^p B'_{jk} h_{t-j} B_{jk}$ soit définie positive. Une conséquence de cette proposition est que pour le modèle GARCH la condition nécessaire à la définie positivité de H_t , $t \in \mathbb{N}$, devient $\ker\{C\} = \{0\}$, c'est à dire que C soit de plein rang.

Remarque 3.2 Une condition suffisante de la proposition précédente serait de prendre la matrice C triangulaire avec des éléments diagonaux strictement positifs. En effet une telle matrice est de plein rang. Nous allons voir que cette hypothèse nous permet d'identifier le modèle.

Identifiabilité du modèle BEKK pour GARCH(1,1) quand $K = 1$:

Dans ce paragraphe, nous supposons que la matrice C est triangulaire.

Proposition 3.3 Pour un modèle GARCH(1,1) défini par (3.5), supposons que les éléments de la diagonale de C sont positifs. Si de plus a_{11} et b_{11} positifs, alors il n'existe pas d'autre matrice C , A ou B donnant une représentation équivalente du modèle. (C'est à dire une même matrice de covariance H_t pour un même processus donné $\forall t \in \mathbb{Z}$)

La condition de positivité des éléments diagonaux vient du fait que la décomposition d'une matrice en produit de matrice triangulaire et sa transposée existe toujours et est unique si les éléments sur la diagonale de la matrice triangulaire sont positifs (Proposition 58, Remarque 34 Dhrymes 1984, pp 68-69). Ecrivons maintenant l'élément $h_{lm,t}$ de la matrice H_t :

$$h_{lm,t} = c_{lm} + \sum_{i=1}^n \sum_{j=1}^n a_{il} a_{jm} \varepsilon_{i,t-1} \varepsilon_{j,t-1} + \sum_{i=1}^n \sum_{j=1}^n b_{il} b_{jm} h_{ij,t-1}. \quad (3.6)$$

Si on écrit le terme $h_{11,t}$ en ignorant les termes GARCH, on obtient :

$$h_{11,t} = \sum_{i=1}^n \sum_{j=1}^n a_{i1} a_{j1} \varepsilon_{i,t-1} \varepsilon_{j,t-1}$$

On remarque que le coefficient correspondant à $\varepsilon_{1,t-1}^2$ dans l'équation de $h_{11,t}$ est a_{11}^2 . Ainsi a_{11} est identifié à un signe près qu'on fixe comme étant positif. Le coefficient correspondant au terme $\varepsilon_{i,t-1} \varepsilon_{j,t-1}$ dans cette équation est $(a_{i1} a_{j1} + a_{j1} a_{i1}) = 2a_{i1} a_{j1}$. On en déduit que a_{j1} est identifié car a_{11} est déjà identifié. Ainsi la première colonne de la matrice A est identifiée. Ecrivons maintenant le terme $h_{12,t}$ de la matrice H_t en ignorant toujours les termes GARCH :

$$h_{12,t} = \sum_{i=1}^n \sum_{j=1}^n a_{i1} a_{j2} \varepsilon_{i,t-1} \varepsilon_{j,t-1}$$

Dans cette équation, le coefficient correspondant au terme $\varepsilon_{1,t-1}^2$ est $(a_{11} a_{12} + a_{12} a_{11}) = 2a_{11} a_{12}$. Comme a_{11} est identifié, alors on peut identifier a_{12} . De la même manière, en considérant les termes $\varepsilon_{1,t-1} \varepsilon_{j,t-1}$ dans l'équation précédente, on identifie les coefficients a_{j2} . Nous avons donc identifié la deuxième colonne de la matrice A . Un raisonnement similaire nous permet d'identifier toutes les colonnes de la matrice.

Remarque 3.3 La matrice C a été choisie comme étant triangulaire par souci de simplicité, cependant on peut prendre n'importe quelle factorisation qui nous assure le fait

que la partie constante est une matrice définie positive.

Généralité de la représentation BEKK pour le modèle GARCH(1,1)

Pour assurer la généralité du modèle BEKK et rendre possible certaines combinaisons de paramètres écartées par les contraintes de définie positivité et d'identifiabilité, on doit prendre $K > 1$. Cependant si K est grand, le nombre de paramètres à estimer devient important, ainsi on doit choisir un entier K minimal assurant l'élimination de toutes les restrictions du modèle. Engle et Kroner (1995) nous donnent deux conditions nécessaires à vérifier pour que le modèle GARCH(1,1) soit général :

Proposition 3.4 Pour qu'il soit général le modèle BEKK $GARCH(1,1)$ doit vérifier les deux conditions suivantes :

- a- Soit $s = n(n+1)/2$, alors K doit être assez grand pour qu'il y ait un total d'au moins s^2 éléments dans les matrices .
- b- Soit $a_{ij,k}$ le i,j ème élément d'une matrice A_{1k} . Alors, il doit exister une matrice A_{1k} qui contient ou bien la paire $a_{il,k}$ et $a_{jm,k}$ ou bien la paire $a_{jl,k}$ et $a_{im,k}$ pour tous les i,j,l,m entre 1 et n .

Des restrictions similaires doivent être étendues aux matrices B_{1k} .

La première condition impose qu'il y ait plus de paramètres dans le modèle BEKK que dans le modèle vectoriel, s'il a moins de s^2 paramètres cela revient à imposer des conditions sur le modèle. La deuxième condition impose que certaines paires de paramètres doivent apparaître ensemble dans une matrice A_{1k} . En effet, en écrivant l'expression de $h_{lm,t}$:

$$h_{lm,t} = c + \sum_{k=1}^K \sum_{i=1}^n \sum_{j=1}^n a_{il,k} a_{jm,k} \varepsilon_{i,t-1} \varepsilon_{j,t-1} + \sum_{k=1}^K \sum_{i=1}^n \sum_{j=1}^n b_{il,k} b_{jm,k} h_{ij,t-1}.$$

On remarque que le coefficient correspondant à $\varepsilon_{i,t-1} \varepsilon_{j,t-1}$ est :

$$\sum_{k=1}^K a_{il,k} a_{jm,k} + a_{jl,k} a_{im,k} \quad \text{si } i \neq j$$

$$\sum_{k=1}^K a_{il,k} a_{im,k} \quad \text{si } i = j$$

Ainsi, il doit exister $K \in \{1, \dots, N\}$ tel que $a_{il,k}$ et $a_{jm,k}$ ou $a_{il,k}$ et différents de zéro pour que le terme $\varepsilon_{i,t-1}\varepsilon_{j,t-1}$ puisse apparaître dans l'expression de $h_{lm,t}$.
Ce qui nous donne la deuxième condition.

Exemple 3.3: cas $n=2$

1- L'ensemble de matrices suivantes :

$$A_{11} = \begin{pmatrix} a_{11,1} & a_{12,1} \\ 0 & a_{22,1} \end{pmatrix}, A_{12} = \begin{pmatrix} a_{11,2} & a_{12,2} \\ a_{21,2} & 0 \end{pmatrix}, A_{13} = \begin{pmatrix} a_{11,3} & 0 \\ a_{21,3} & a_{22,3} \end{pmatrix}$$

satisfait les deux conditions énoncées : d'une part le nombre de paramètres est égal à neuf qui correspond à s^2 quand $n = 2$. D'autre part on remarque que $\forall a_{ij,k}$ un élément de A_{1k} , il existe une matrice $A_{1k'}$ telle que $a_{ij,k} \times a_{lm,k'} \neq 0 \forall i, j \in \{1, 2\}$, et $k, k' \in \{1, 2, 3\}$.

2- Considérons maintenant l'ensemble de matrices suivantes :

$$A_{11} = \begin{pmatrix} a_{11,1} & a_{12,1} \\ 0 & a_{22,1} \end{pmatrix}, A_{12} = \begin{pmatrix} a_{11,2} & a_{12,2} \\ 0 & 0 \end{pmatrix},$$

$$A_{13} = \begin{pmatrix} a_{11,3} & 0 \\ a_{21,3} & 0 \end{pmatrix}, A_{14} = \begin{pmatrix} 0 & a_{12,4} \\ 0 & a_{22,4} \end{pmatrix}.$$

Cet ensemble de matrices n'engendre pas une représentation générale, car il ne respecte pas la deuxième condition de la proposition. En effet, les éléments a_{21} et a_{22} n'apparaissent jamais ensemble dans une matrice A_{1k} .

Si on écrit les équations exprimant $h_{12,t}$ ou $h_{21,t}$ on obtient :

$$h_{12,t} = c_1 + (a_{11,1}a_{12,1} + a_{11,2}a_{12,2}\varepsilon_{1,t-1}^2 + 2a_{11,1}a_{22,1}\varepsilon_{1,t-1}\varepsilon_{2,t-1})$$

$$h_{21,t} = c_2 + (a_{11,1}a_{12,1} + a_{11,2}a_{12,2}\varepsilon_{1,t-1}^2 + 2a_{22,1}a_{11,1}\varepsilon_{1,t-1}\varepsilon_{2,t-1})$$

le coefficient correspondant à $\varepsilon_{2,t-1}^2$, qui est $\sum_{k=1}^K a_{21,k}a_{22,k}$ n'apparaît pas. Ainsi le terme $\varepsilon_{2,t-1}^2$ sera éliminé des équations exprimant $h_{12,t}$ et $h_{21,t}$, ce qui constitue une restriction. Cependant les conditions énoncées ci-dessus sont des conditions nécessaires, mais non suffisantes pour la généralité du modèle BEKK. Si on considère l'ensemble de matrices suivantes :

$$A_{11} = \begin{pmatrix} a_{11,1} & a_{12,1} \\ a_{21,1} & a_{22,1} \end{pmatrix}, A_{12} = \begin{pmatrix} 0 & a_{12,2} \\ 0 & a_{22,2} \end{pmatrix},$$

$$A_{13} = \begin{pmatrix} 0 & a_{12,3} \\ 0 & a_{22,3} \end{pmatrix}, A_{14} = \begin{pmatrix} 0 & 0 \\ 0 & a_{22,4} \end{pmatrix}.$$

Cet ensemble de matrices satisfait les conditions de la proposition, mais n'est cependant pas pleinement général, car si l'on écrit l'expression du terme $h_{11,t}$ on obtient :

$$h_{11,t} = c + (a_{11,1}^2 \varepsilon_{1,t-1}^2 + 2a_{11,1}a_{21,1}\varepsilon_{1,t-1}\varepsilon_{2,t-1} + a_{21,1}^2 \varepsilon_{2,t-1}^2$$

On peut remarquer que le paramètre $a_{11,1}$ correspond au terme $\varepsilon_{1,t-1}^2$, le paramètre $a_{11,1}a_{21,1}$ correspond au terme $\varepsilon_{1,t-1}\varepsilon_{2,t-1}$, les paramètres $a_{11,1}$ et $a_{21,1}$ sont fixés ce qui ne laisse aucune liberté au paramètre exprimant la quantité $\varepsilon_{2,t-1}^2$ dans l'équation de $h_{11,t}$, ce qui est une restriction pour le modèle. Dans notre exemple, la restriction est due au fait qu'il n'y a que deux paramètres pour trois termes dans la première équation. En fait, il est possible de trouver des conditions suffisantes dont un cas extrême serait de prendre $K = s$ et des matrices qui n'auraient aucune restriction. Cependant dans ce type de cas le nombre de paramètres à estimer sont introduits sans qu'ils soient forcément nécessaires pour lever les restrictions mises en évidence dans ce dernier exemple.

Parallèlement à ces problèmes se pose un problème d'identification que nous avons discuté pour le cas $K = 1$ seulement. Par exemple pour le cas $K > 1$ le fait d'interchanger les matrices A_{11} et A_{12} nous donne deux expressions différentes d'un même modèle. C'est pourquoi on doit imposer certaines restrictions aux matrices A_{1k} et B_{1k} . Dans la proposition suivante Engle et Kroner proposent une restriction particulièrement pratique car ce modèle est identifiable mais aussi général :

Proposition 3.5 Supposons que les éléments diagonaux de C soient positifs. Considérons la classe des modèles $BEKK$ dans lesquels les matrices A_{1k_r} sont obtenues en annulant les $r - 1$ premières colonnes et les $k_r - n(r - 1) - 1$ premières lignes, où $k_r = n(r - 1) + 1, \dots, nr$ et $r = 1, \dots, n$. Si on suppose aussi que $a_{nn,k_r} > 0; \forall k_r$ et que des restrictions similaires sont imposées sur les matrices B_{1k_r} , alors un modèle $BEKK$ général sans autre représentation

équivalente est obtenu dans cette classe.

Remarque 3.4 Dans les modèles définis par la proposition précédente, on a: $K = n^2$

Exemple 3.4 cas $n = 2$, Calculons les indices des matrices A_{1k_r} : les valeurs possibles de k_r sont $k_r = 2(r - 1) + 1$, $2(r - 1) + 2$, et $r = 1, 2$.

	r=1	r=2
$k_r = 2(r - 1) + 1$	1	3
$k_r = 2r$	2	4

On obtient les matrices suivantes dans le cas $ARCH(1)$:

$$A_{11} = \begin{pmatrix} a_{11,1} & a_{12,1} \\ a_{21,1} & a_{22,1} \end{pmatrix}, A_{12} = \begin{pmatrix} 0 & 0 \\ a_{21,2} & a_{22,2} \end{pmatrix},$$

$$A_{13} = \begin{pmatrix} 0 & a_{12,3} \\ 0 & a_{22,3} \end{pmatrix}, A_{14} = \begin{pmatrix} 0 & 0 \\ 0 & a_{22,4} \end{pmatrix}.$$

On remarque effectivement que cet exemple satisfait les conditions nécessaires de généralité du modèle BEKK. De plus si on écrit l'expression des éléments de $vech(H_t)$ on a:

$$\left\{ \begin{array}{l} h_{11,t} = c_{11} + \underline{a_{11,1}^2} \varepsilon_{1,t-1}^2 + 2a_{11,1} \underline{a_{21,1}} \varepsilon_{1,t-1} \varepsilon_{2,t-1} + (\underline{a_{21,1}^2} + \underline{a_{21,2}^2}) \varepsilon_{2,t-1}^2 \\ h_{21,t} = c_{21} + a_{11,1} \underline{a_{12,1}} \varepsilon_{1,t-1}^2 + (a_{21,1} a_{12,1} + a_{11,1} \underline{a_{22,1}}) \varepsilon_{1,t-1} \varepsilon_{2,t-1} + \\ (a_{21,1} a_{22,1} + a_{21,2} \underline{a_{22,2}}) \varepsilon_{2,t-1}^2 \\ h_{22,t} = c_{22} + (\underline{a_{12,1}^2} + \underline{a_{12,3}^2}) \varepsilon_{1,t-1}^2 + (a_{12,1} a_{22,1} + a_{12,3} \underline{a_{22,3}}) \varepsilon_{1,t-1} \varepsilon_{2,t-1} + \\ (\underline{a_{22,1}^2} + \underline{a_{22,2}^2} + \underline{a_{22,4}^2}) \varepsilon_{2,t-1}^2 \end{array} \right.$$

On remarque que chaque terme a un paramètre libre (on peut considérer les paramètres soulignés par exemple). En ce sens, ce modèle est général.

Cependant on peut remarquer que pour certains termes correspondent une somme de carrés de paramètres induisant des restrictions: par exemple le coefficient du terme $\varepsilon_{2,t-1}^2$ sera supérieur à $a_{21,1}^2$ ce paramètre étant déjà fixé par le terme $\varepsilon_{1,t-1} \varepsilon_{2,t-1}$.

3.2.3 Comparaison entre le modèle BEKK et le modèle vectoriel

En écrivant l'expression du modèle BEKK pour GARCH(1,1) et en utilisant la relation $vec(ABC) = (C' \otimes A)vec(B)$ on a :

$$h_t = (C \otimes C')vec(\mathbf{I}_n) + \sum_{k=1}^K (a_{1k} \otimes a_{1k})'vec(\varepsilon_{t-1}\varepsilon_{t-1}') \\ + \sum_{k=1}^K (B_{1k} \otimes B_{1k})'vec(H_{t-1})$$

.

Si on prend $A_1 = \sum_{k=1}^K (a_{1k} \otimes a_{1k})'$ et $B_1 = \sum_{k=1}^K (B_{1k} \otimes B_{1k})'$ on obtient l'expression du modèle vectoriel. Ce qui nous amène à formuler la proposition suivante :

Proposition 3.6 Un modèle vectoriel appartient à la classe des modèles BEKK si et seulement si il existe C, A_{ik} et B_{ik} tels que :

$$C_0 = (C \otimes C')vec(\mathbf{I}_n),$$

$$A_i = \sum_{k=1}^K (a_{ik} \otimes a_{ik})'$$

$$B_j = \sum_{k=1}^K (B_{jk} \otimes B_{jk})'$$

Remarques 3.5

- 1- Un modèle vectoriel obtenu à partir d'un modèle BEKK est unique d'après la proposition précédente. Cependant l'inverse n'est pas vraie car $(a_{1k} \otimes a_{1k}) = ((-a_{1k}) \otimes (-a_{1k}))$ ou encore en changeant les indices k on obtient le même modèle vectoriel. Donc le choix des matrices A_{ik} n'est pas unique.

- 2- D'après les équations de la proposition précédente, il est clair qu'un modèle diagonal vectoriel peut être obtenu à partir d'un modèle BEKK si et seulement si les matrices A_{ik} et B_{jk} sont diagonales.
- 3- Les représentations BEKK étant restreintes de manière à ce que H_t soit définie positive, on voit que toutes les représentations vectorielles dont la matrice de covariance conditionnelle H_t n'est pas définie positive pour tout t n'admettent pas de représentation BEKK.

Proposition 3.7 Dans le modèle vectoriel, si on suppose que $C_0 = \text{vec}(\Omega)$ avec Ω matrice définie positive, si les matrices A_i et B_i sont diagonales. Alors, il existe une matrice triangulaire C et des matrices diagonales A_{ik} et B_{ik} , $k = 1, \dots, K$ telles que :

$$C_0 = \text{vec}(C' C),$$

$$A_i = \sum_{k=1}^n (a_{1k} \otimes a_{1k})'$$

$$B_i = \sum_{k=1}^n (B_{ik} \otimes B_{ik})'$$

La proposition précédente montre que le modèle BEKK inclut comme cas particuliers tout les modèles linéaires diagonaux définis positifs. De plus, on peut restreindre les $(n - k)$ premiers éléments des matrices A_{ik} et B_{ik} à être nuls. D'après la proposition (3.6), ce modèle est général sans autre représentation équivalente dans cette classe. D'une manière générale, on voit qu'on ne peut considérer le modèle vectoriel uniquement quand cette représentation est équivalente au modèle BEKK qui garantit la définie positivité de H_t .

3.2.4 Etude de la stationnarité

Nous allons étudier la stationnarité au second ordre des modèles vectoriels et BEKK. Soit L l'opérateur retard défini comme suit $L^k \omega_t \equiv \omega_{t-k}$, on considère le processus MGARCH suivant :

$$h_t = C + A(L)e_t + B(L)h_t \quad (3.7)$$

on obtient le processus écrit sous le modèle vectoriel en prenant :

$$A(L) = A_1L + \dots + A_qL^q$$

$$B(L) = B_1L + \dots + B_pL^p$$

où les matrices $A_1, \dots, A_q, B_1, \dots, B_p$ sont de dimension $n \times n$. En utilisant la relation $vec(ABC) = (C' \otimes A)vec(B)$, on obtient le modèle BEKK auquel on a appliqué l'opérateur $vec(\cdot)$ en prenant :

$$A(L) = \sum_{k=1}^K (A_{ik} \otimes A_{ik})L + \dots + \sum_{k=1}^K (A_{qk} \otimes A_{qk})L^q$$

$$B(L) = \sum_{k=1}^K (B_{ik} \otimes B_{ik})L + \dots + \sum_{k=1}^K (B_{pk} \otimes B_{pk})L^p$$

où les matrices $A_{1k}, \dots, A_{qk}, B_{1k}, \dots, B_{pk}$ sont de dimension $n \times n$

Développons l'équation (3.7) :

$$\begin{aligned} h_t &= C + A(L)e_t + B(L)h_t = C + A(L)e_t + \sum_{j=1}^p B_j h_{t-j} \\ &= C + A(L)e_t + \sum_{j=1}^p B_j (C + A(L)\varepsilon_{t-j} + B(L)h_{t-j}) \\ &= C + A(L)e_t + \sum_{j=1}^p B_j C + A(L) \sum_{j=1}^p B_j \varepsilon_{t-j} + B(L) \sum_{j=1}^p B_j h_{t-j} \\ &= C + A(L)e_t + B(L)(C + A(L)\varepsilon_t) + B^2(L)h_t \end{aligned} \tag{3.8}$$

si on réitère $(n - 1)$ fois l'opération obtient :

$$h_t = \sum_{i=1}^n B^{i-1}(C + A(L)e_t) + B^n(L)h_t$$

Ainsi, on remarque que la quantité $h_t = \sum_{i=1}^{\infty} B^{i-1}(C + A(L)e_t)$ est solution du processus GARCH. En effet :

$$\begin{aligned}
h_t &= C + A(L)e_t + \sum_{i=2}^{\infty} B^{i-1}(L)(C + A(L)e_t) \\
&= C + A(L)e_t + B(L) \sum_{i=1}^{\infty} B^{i-1}(L)(C + A(L)e_t) \\
&= C + A(L)e_t + B(L)h_t
\end{aligned} \tag{3.9}$$

Notons que la somme $h_t = \sum_{i=1}^{\infty} B^{i-1}(L)(C + A(L)e_t)$ existe si le rayon spectral de B est inférieur à 1. Cette écriture va nous permettre d'exprimer la condition de stationnarité pour les modèles vectoriels et BEKK :

Proposition 3.8 Soit $\{\varepsilon_t\}_{t \in \mathbb{N}}$ un processus, si l'équation (3.7) définit un processus GARCH, alors $\{\varepsilon_t\}_{t \in \mathbb{N}}$ est stationnaire au second ordre si et seulement si $|\rho(A(1) + B(1))| < 1$.

$\rho(A)$ étant le rayon spectral de la matrice A .

Démonstration

Nous allons considérer la stationnarité dans le cas GARCH(1,1) uniquement. L'extension à des modèles plus généraux suit le même raisonnement, mais demande des calculs trop amples pour être exposée ici. Exprimons l'espérance conditionnelle de e_t aux dates $t-1, t-2, \dots$:

$$E_{t-1}(e_t) = \sum_{i=1}^{\infty} B_1^{i-1}(C + A_1 e_{t-i}),$$

$$E_{t-2}(e_t) = E_{t-1}E_{t-2}(e_t)$$

$$= E_{t-2} \left(\sum_{i=1}^{\infty} B_1^{i-1}(C + A_1 e_{t-i}) \right)$$

$$\begin{aligned}
&= \sum_{i=1}^{\infty} B_1^{i-1} (C + A_1 E_{t-2} e_{t-i}) \\
&= C + A_1 E_{t-2} e_{t-i} + \sum_{i=2}^{\infty} B_1^{i-1} (C + A_1 E_{t-2} e_{t-i}) \\
&= C + A_1 + \sum_{i=1}^{\infty} B_1^{i-1} (C + A_1 e_{t-i-1}) \\
&\quad + B_1 \sum_{i=1}^{\infty} B_1^{i-1} (C + A_1 e_{t-i-1}) \\
&= C + (A_1 + B_1) \sum_{i=1}^{\infty} B_1^{i-1} (C + A_1 e_{t-i-1}). \quad (3.10)
\end{aligned}$$

En réitérant l fois cette opération, on obtient :

$$\begin{aligned}
E_{t-l}(e_t) &= (I + (A_1 + B_1) + \dots + (A_1 + B_1)^{l-2})C \\
&\quad + (A_1 + B_1)^{l-1} \sum_{i=1}^{\infty} B_1^{i-1} (C + A_1 e_{t-l+1}).
\end{aligned}$$

Soit Z une matrice carrée, $Z^l \longrightarrow 0$ quand $l \longrightarrow \infty$ si et seulement si $\rho(Z) < 1$ et si et seulement si $(I + Z + \dots) \longrightarrow (I - Z)^{-1}$. Ainsi, $E_{t-l}(e_t)$ converge vers $(I - A_1 - B_1)C$ quand $l \longrightarrow \infty$ si et seulement si les valeurs propres de A_1 et B_1 sont inférieures à 1.

Remarque 3.6 Pour le modèle diagonal, la condition de stationnarité au second ordre pour le modèle GARCH(1,1) : $a_{ii} + b_{ii} < 1 \ \forall i \in \{1, \dots, N\}$ dans le cas du modèle vectoriel, et $\sum_{k=1}^n (a_{ii,k} + b_{ii,k})$ dans le modèle BEKK.

3.2.5 Modèles à corrélations conditionnelles constantes (CCC)

Les modèles à corrélations conditionnelles constantes sont des modèles pour lesquels les variances des composantes du processus GARCH multivarié varient en fonction du temps et les corrélations conditionnelles sont quant à elles constantes. Cette hypothèse énoncée par

Bollerslev(1990) est justifiée par l'étude empirique des processus multivariés. La corrélation conditionnelle entre ε_{it} et ε_{jt} à l'instant $t - 1$ est donnée par la relation suivante :

$$\rho_{ij,t} = h_{ij,t}/(h_{ii,t}h_{jj,t})^{1/2}.$$

En général cette corrélation varie au cours du temps, mais dans certains cas on peut la supposer constante : les covariances conditionnelles sont proportionnelles à la racine carrée du produit des variances :

$$\rho_{ij,t} = h_{ij,t}/(h_{ii,t}h_{jj,t})^{1/2}. \quad (3.11)$$

avec $j = 1, \dots, n$ et $i = j + 1, \dots, n$

Remarque 3.7 La validité de l'hypothèse nous donnant l'équation précédente relève d'une étude empirique de certaines séries de processus MGARCH.

Il est cependant important de souligner que les variances conditionnelles varient au cours du temps dans ce modèle : $h_{ii,t}$ est une fonction du passé. Ainsi, on tire une décomposition de H_t :

$$H_t = \Delta_t \Gamma \Delta_t \quad (3.12)$$

avec Δ_t une matrice diagonale de dimension $n \times n$ ayant pour éléments $\sqrt{h_{11,t}}, \dots, \sqrt{h_{nn,t}}$. La matrice R est constante par rapport au temps, de dimension $n \times n$ et qui a pour éléments $\rho_{i,j}$ et 1 sur la première diagonale. D'après l'équation (3.12), il est clair que H_t est définie positive dès que les variances conditionnelles sont bien définies et R définie positive car :

$$x' H_t x > 0 \quad \forall x \in \mathbb{R}^n, x \neq 0 \Leftrightarrow x' D_t R D_t > 0 \quad \forall x \in \mathbb{R}^n x \neq 0.$$

De plus on peut remarquer que D_t et $\Gamma^{1/2}$ commutent, en effet, on a :

$$H_t^{1/2} = \Delta_t \Gamma^{1/2} = \Gamma^{1/2} \Delta_t$$

Modèles à corrélations constantes dynamiques (DCC):

D'une manière générale faire l'hypothèse que les corrélations conditionnelles sont constantes peut être trop restrictive pour l'étude de certaines séries. Engle (2002) a introduit le modèle à corrélations conditionnelles dynamiques (Dynamic Constant Correlation). Dans ce modèle R varie au cours du temps, l'expression de H_t est la suivante :

$$H_t = D_t R D_t \quad (3.13)$$

où la matrice D_t a les mêmes spécifications que dans le modèle à corrélation constante et la matrice R_t de dimension $n \times n$ a pour éléments $r_{ij,t} = \rho_{ij,t}$ quand $i \neq j$ et 1 quand $i = j$. Les éléments $r_{ij,t}$ sont fonctions du passé, Engle a donné deux spécifications GARCH de ces termes. Cependant aucune étude n'a encore été faite sur ces spécifications.

extension du modèle à corrélation constante:

Jeantheau (1998) a proposé une extension du modèle à corrélation constante de Bollerslev qui admet la même décomposition de la matrice de covariance conditionnelle H_t définie en (2.12). Les covariances conditionnelles étant constantes dans ce modèle, ce sont uniquement les variances conditionnelles qui sont exprimées en fonction du passé. Soit une série $\{\varepsilon_t, t \in \mathbb{Z}\}$ de variables aléatoires à valeurs dans $\mathbb{R}^n$ qui suit un processus GARCH multivarié, on a :

$$\varepsilon_t = \Delta_t \eta_t \quad (3.14)$$

où Δ_t est une matrice diagonale et les éléments $\Delta_{t,ii}$ satisfont la relation suivante :

$$\begin{pmatrix} \Delta_{t,11}^2 \\ \cdot \\ \cdot \\ \cdot \\ \Delta_{t,nn}^2 \end{pmatrix} = C + \sum_{i=1}^q A_i \begin{pmatrix} \varepsilon_{t-i,1}^2 \\ \cdot \\ \cdot \\ \cdot \\ \varepsilon_{t-i,n}^2 \end{pmatrix} + \sum_{j=1}^p B_j \begin{pmatrix} \Delta_{t-i,11}^2 \\ \cdot \\ \cdot \\ \cdot \\ \Delta_{t-i,nn}^2 \end{pmatrix} \quad (3.15)$$

avec $C \in \mathbb{R}^n$, A_i et $B_j \in \mathbb{R}^N \times \mathbb{R}^N$, nous supposons que les coefficients des matrices de la relation précédente sont tous positifs. De plus, η_t est une série de variables aléatoires

indépendantes identiquement distribuées vérifiant les conditions suivantes :

- 1- L'espérance de η_t est nulle, et la matrice de covariance Γ est telle que :

$$\begin{pmatrix} 1 & \rho_{12} & \cdots & \rho_{1n} \\ \rho_{12} & \ddots & & \vdots \\ \vdots & & \ddots & \rho_{(n-1)n} \\ \rho_{1n} & \cdots & \rho_{(n-1)n} & 1 \end{pmatrix}$$

- 2- La variable aléatoire η_t est indépendant de la tribu I_{t-1} engendrée par les valeurs passées du processus.
- 3- La loi de η_t est telle qu'il n'existe pas de forme quadratique q telle que $q(\eta_t) = \gamma$ p.s, avec $\gamma \in \mathbb{R}$. Enfin θ est le vecteur des paramètres.

De cette définition, on obtient l'expression des éléments $h_{ij,t}$ de la matrice de covariance conditionnelle du processus ε_t :

$$\begin{cases} h_{ij,t} = (\Delta_{t,ii})^2 \\ h_{ij,t} = \rho_{ij}\Delta_{t,ii}\Delta_{t,jj} \quad \text{pour } i \neq j. \end{cases}$$

En effet, si on exprime H_t en fonction des matrices Δ_t et Γ on a :

$$H_t = E_{t-1}(\varepsilon_t \varepsilon_t') = E_{t-1}(\Delta_t \eta_t \eta_t' \Delta_t') = \Delta_t \Gamma \Delta_t$$

Ainsi, on vérifie bien que les corrélations conditionnelles sont constantes.

Remarques 3.8

- 1- Si on impose que les composantes de η_t sont indépendantes entre elles alors H_t est diagonale.
- 2- Si A_i et B_j sont diagonales, $h_{ii,t}$ ne dépend que de ses propres valeurs passées et des valeurs passées de ε_t , alors que dans ce cas $h_{ij,t}$ dépend des valeurs passées de $h_{ii,t}$ et $h_{jj,t}$ et des valeurs passées de $\varepsilon_{i,t}$ et $\varepsilon_{j,t}$.

Nous allons maintenant spécifier les conditions à imposer sur les paramètres pour que le modèle soit stationnaire.

3.2.6 Etude de la stationarité des modèles CCC

A partir des équations (3.14) et (3.15) on peut écrire le modèle à corrélation constante de la manière suivante :

$$\begin{pmatrix} h_{11,t} \\ \vdots \\ h_{nn,t} \end{pmatrix} = C + \sum_{i=1}^s (A_i(\eta_t^2) + B_i) \begin{pmatrix} h_{11,t-i} \\ \vdots \\ h_{nn,t-i} \end{pmatrix} \quad (3.16)$$

où $s = \sup(p,q)$, $A_i(\eta_t^2) = 0$ pour $i > q$ et $B_j = 0$ pour $j > p$.

La proposition suivante nous donne la condition de stationnarité du modèle à corrélation constante :

Proposition 3.9 Si le vecteur des paramètres θ est tel que $\det(I_n - \sum_{i=1}^s (A_i + B_i)\lambda)$ a ses racines en dehors du cercle unité alors le modèle MGARCH(p,q) à corrélation constante admet une solution stationnaire au second ordre. De plus cette solution est unique et strictement stationnaire ergodique.

En effet, exprimons le modèle sous forme matricielle :

Soit

$$\begin{pmatrix} H_{11,t} & \cdots & H_{nn,t} & H_{11,t-1} & \cdots & H_{dd,t-n+1} \end{pmatrix} \quad \text{ona} :$$

$$V_t = \begin{pmatrix} A_1(\eta_{t-1}^2) + B_1 & A_2(\eta_{t-2}^2) + B_2 & \cdots & A_n(\eta_{t-n}^2) + B_n \\ I_n & & & 0 \\ & \ddots & & \vdots \\ & & I_n & 0 \end{pmatrix} V_{t-1} + \begin{pmatrix} C \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

on obtient:

$$V_t = F(\zeta)V_{t-1} + G$$

avec $\zeta'_t = (\eta_{t-1}, \eta_{t-2}, \dots, \eta_{t-s})$ et, $G' = (C, 0, \dots, 0)$ et $F(\zeta_t)$ la matrice de la relation précédente. On en déduit que :

$$V_t = F(\zeta_t) \dots F(\zeta_{t-k+1}) V_{t-1} + \sum_{t=0}^{k-1} F(\zeta_t) \dots F(\zeta_{t-k+1}) G$$

On remarque que le second terme de l'égalité précédente est une série de termes positifs, de plus on a :

$$E(F(\zeta_t) \dots F(\zeta_{t-k+1})) = F^i G,$$

$$\text{avec } F = \begin{pmatrix} A_1(\eta_{t-1}^2) + B_1 & A_2(\eta_{t-2}^2) + B_2 & \dots & A_n(\eta_{t-n}^2) + B_n \\ I_n & & & 0 \\ & \ddots & & \vdots \\ & & I_n & 0 \end{pmatrix}.$$

Par conséquent, $\det(-F)$ et $\det(I_n - \sum_{i=1}^s (A_i + B_i) \lambda^{-1})$ ont les mêmes racines. Ainsi, la série converge dans L^1 si $\det(I_n - \sum_{i=1}^s (A_i + B_i) \lambda)$ a ses racines en dehors du cercle unité.

$$\text{soit } \bar{V}_{t,\theta} = \sum_{t=1}^{\infty} F(\zeta_t) \dots F(\zeta_{t-k+1}) G,$$

comme les variables aléatoires η_t sont indépendantes, alors le processus (ζ_t) , et par conséquent le processus $(\bar{V}_{t,\theta})$ sont strictement stationnaires ergodiques. On en déduit que $\bar{\varepsilon}_{t,i} = \sqrt{(h_{ii,t})} \eta_{t,i}$ est un processus GARCH(p,q) strictement stationnaire et converge dans L^2 .

De plus cette solution est unique car si l'on suppose que ε_t est une autre solution, alors V_t satisfait :

$$V_t = F(\zeta_t) \dots F(\zeta_{t-k+1}) V_{t-k} + \sum_{t=0}^{k-1} F(\zeta_t) \dots F(\zeta_{t-k+1}) G$$

Or, on a :

$$F(\zeta_t) \dots F(\zeta_{t-k+1}) V_{t-k} = F^k E(V_{t-k}) = c F^k \longrightarrow 0$$

avec $c > 0$. Ainsi, $\bar{V}_t = V_t$ et donc $\bar{\varepsilon}_t = \varepsilon_t$.

Chapitre 4

Estimation des paramètres du GARCH multivarié

4.1 Estimation par le quasi-maximum de vraisemblance QMV

Un avantage de ce modèle est la simplicité du calcul du quasi-maximum de vraisemblance au moment de procéder à l'estimation des paramètres. Exprimons l'expression de la log vraisemblance dans le cas du modèle BEKK :

$$\begin{aligned}
 L(\theta) &= \sum_{t=1}^T \log[1/(2\pi)^{n/2} \times 1/|H_t|^{-1/2} \times \exp(-1/2\varepsilon_t' H_t^{-1} \varepsilon_t)] \\
 &= \sum_{t=1}^T -n/2 \log(2\pi) - 1/2(\log(|H_t|) + \varepsilon_t' H_t^{-1} \varepsilon_t) \\
 &= -Tn/2 \log(2\pi) - 1/2 \sum_{t=1}^T (\log(|H_t|) + \varepsilon_t' H_t^{-1} \varepsilon_t)
 \end{aligned}$$

où T est le nombre d'observations de la série.

Dans cette expression on doit inverser la matrice H_t de dimension $(n \times n)$ à chaque instant t ce qui demande beaucoup de calculs sachant que le nombre d'observations pour les séries financières est en général grand. Dans le modèle à corrélations constantes le fait que l'on puisse décomposer la matrice H_t nous permet de ne pas inverser H_t à chaque date t car $H_t^{-1} = \Delta_t^{-1} \Gamma^{-1} \Delta_t^{-1}$ on obtient une expression du quasi-maximum de vraisemblance qui est la suivante :

$$\begin{aligned}
 L(\theta) &= -Tn/2\log(2\pi) - 1/2 \sum_{t=1}^T (\log(|\Delta_t R \Delta_t|) + \varepsilon_t' \Delta_t^{-1} \Gamma^{-1} \Delta_t^{-1} \varepsilon_t) \\
 &= -Tn/2\log(2\pi) - T/2\log(|R|) - \sum_{t=1}^T (\log(|\Delta_t|) - 1/2 \sum_{t=1}^T \tilde{\varepsilon}_t' \Gamma^{-1} \tilde{\varepsilon}_t) \\
 \text{où } \tilde{\varepsilon}_t &= \Delta_t^{-1} \varepsilon_t
 \end{aligned}$$

4.2 Estimation par la methode EbE (Equation-by-Equation)

soit $\varepsilon_t = (\varepsilon_{1t}, \dots, \varepsilon_{nt})'$ un processus a valeurs dans $\mathbb{R}^n$, et $\mathcal{F}_{t-1}^\varepsilon$ la tribu engendrée par $\{\varepsilon_u, u < t\}$, nous supposons que:

$$E(\varepsilon_t \setminus \mathcal{F}_{t-1}^\varepsilon) = 0, \quad \text{var}(\varepsilon_t \setminus \mathcal{F}_{t-1}^\varepsilon) = H_t \quad \text{existe et elle est définie positive.} \quad (4.1)$$

avec σ_{kt}^2 les éléments de la diagonale de la matrice H_t , qui sont les variances conditionnelles par rapport a $\mathcal{F}_{t-1}^\varepsilon$ des composants de ε_t .

nous introduisant un vecteur

$$\eta_{kt}^* = D_t^{-1} \varepsilon_t = (\varepsilon_{1t}/\sigma_{1t}, \dots, \varepsilon_{mt}/\sigma_{mt})' \quad \text{avec} \quad D_t = \text{diag}(\sigma_{1t}, \dots, \sigma_{mt})$$

par (4.1) nous avons $E(\eta_{kt}^* \setminus F_{t-1}^\varepsilon) = 0$ et la matrice des correlations conditionnelles est données par

$$R_t = \text{var}(\eta_{kt}^* \setminus F_{t-1}^\varepsilon) = D_t^{-1} H_t D_t^{-1}.$$

qui suit que pour $k = 1, \dots, m$, on a:

$$E(\eta_{kt}^* \setminus F_{t-1}^\varepsilon) = 0, \quad \text{var}(\eta_{kt}^* \setminus F_{t-1}^\varepsilon) = 1. \quad (4.2)$$

Nous nous intéressons à l'estimation de la variances conditionnelle de chaque composante de ε_t qui satisfait:

$$\begin{cases} \varepsilon_t = H_t^{1/2} \eta_t, & E(\eta_t \setminus F_{t-1}^\varepsilon) = 0, \quad \text{var}(\eta_t \setminus F_{t-1}^\varepsilon) = I_m, \\ H_t = H(\varepsilon_{t-1}, \dots, \varepsilon_{t-p}) = D_t R_t D_t \end{cases} \quad (4.3)$$

Où H_t la matrice des variances conditionnelle qui est définie positive, $D_t = \{\text{diag}(H_t)\}^{1/2}$ et $R_t = \text{Corr}(\varepsilon_t, \varepsilon_t \setminus F_{t-1}^\varepsilon)$ la matrice de corrélation conditionnelle.

Nous supposons que la variance conditionnelle de la k -ème composante de ε_t est paramétrée Par un paramètre $\theta_0^{(k)} \in \mathbb{R}^{d_k}$, de sorte que

$$\begin{cases} \varepsilon_{kt} = \sigma_{kt} \eta_{kt}^*, \\ \sigma_{kt} = \sigma_k(\varepsilon_{t-1}, \varepsilon_{t-2}, \dots; \theta_0^{(k)}), \end{cases} \quad (4.4)$$

ou σ_k est une fonction positive, le processus (η_t^*) peut être appelé le vecteur de l'innovation EbE de ε_t .

Remarque 4.1 Dans le modèle (4.3)-(4.4), la volatilité de tout composant de ε_t dépend des valeurs passées de tous les composants, Cette hypothèse représente une extension de la configuration classique des modèles GARCH univariés et pour cette raison, le modèle (4.4) peut être appelé "modèle GARCH augmenté" comme dans la terminologie de Hörmann (2008). Cette extension est d'abord motivée par la généralité: il semble très restrictif de supposer que la variance conditionnelle d'un composant n'est pas influencée par le passé d'autres composants. D'autre part, l'approche de l'estimateur EbE du papier rend cette extension accessible à l'inférence statistique, sans provoquer une explosion du nombre de paramètres. Par exemple, si les volatilités individuelles ont une dynamique de type $GARCH(1,1)$

$$\sigma_{kt}^2 = \omega_k + \sum_{\ell=1}^m \alpha_{k,\ell} \varepsilon_{\ell,t-1}^2 + \beta_k \sigma_{k,t-1}^2, \quad \omega_k, \alpha_{k,\ell}, \beta_k > 0$$

En augmentant de K le nombre de composants implique un nombre supplémentaire de paramètres K par équation. Enfin, cette extension permet d'aborder le problème des données asynchrones en permettant à chaque variance conditionnelle de dépendre des observations les plus récentes.

Remarque 4.2 La positivité de la fonction σ_k implique généralement des restrictions sur les valeurs du paramètre qui ne peuvent pas être explicitées sous la formulation générale,

pour des raisons particulières les modèles avec de telles contraintes peuvent être expliqués, comme dans le cas GARCH(1.1). Notez que les innovations EbE (η_{kt}^*) ne sont pas en général des suite iid, et donc (4.4) n'est pas un processus générateur de données (DGP).

Nous étudions l'estimation du paramètre $\theta_0^{(k)}$ dans le modèle GARCH augmenté (4.4), sous (4.2) pour $k = 1, \dots, m$.

Pour estimer $\theta_0^{(k)}$, nous utiliserons la methode QMV, qui est la méthode d'estimation la plus utilisée pour les modèles GARCH univariés, mais d'autres méthodes pourraient également être considérées.

En vue des remarques (4.1) et (4.2) le modèle GARCH augmenté (4.4) n'est pas, en général, un GARCH univarié et nous ne pouvons pas compter directement sur les résultats existants pour son estimation.

Compte tenu des observations $(\varepsilon_1, \dots, \varepsilon_n)$ et des valeurs initiales arbitraires $\tilde{\varepsilon}_i$ pour $i \leq 0$, nous définissons:

$$\tilde{\sigma}_{kt}(\theta^{(k)}) = \sigma_k(\varepsilon_{t-1}, \dots, \varepsilon_1, \tilde{\varepsilon}_0, \tilde{\varepsilon}_{-1}, \dots; \theta^{(k)})$$

pour $k = 1, \dots, m$ et $\theta^{(k)} \in \Theta_k$ en supposant Que Θ_k est un ensemble de paramètres compact et $\theta_0^{(k)} \in \Theta_k$.

Cette variable aléatoire sera approximée par:

$$\sigma_{kt}(\theta^{(k)}) = \sigma_k(\varepsilon_{t-1}, \varepsilon_{t-2}, \dots; \theta^{(k)}).$$

Soit $\hat{\theta}_n^{(k)}$ l'estimateur EbE de $\theta_0^{(k)}$:

$$\hat{\theta}_n^{(k)} = \arg \min \tilde{Q}_n^{(k)}(\theta^{(k)}), \quad \theta_0^{(k)} \in \Theta_k$$

$$\tilde{Q}_n^{(k)}(\theta^{(k)}) = \frac{1}{2} \sum_{t=1}^n \log \tilde{\sigma}_{kt}^2(\theta^{(k)}) + \frac{\varepsilon_{kt}^2}{\tilde{\sigma}_{kt}^{(2)}(\theta^{(k)})}.$$

4.3 consistance et normalité asymptotique de l'estimateur EbE

Sous les hypothèses suivantes sur le processus ε_t .

H1: ε_t est un processus strictement stationnaire et ergodique satisfaisant (4.1), avec $E|\varepsilon_{kt}|^s < \infty$ pour $s > 0$, de plus $E(\log \sigma_{kt}^2) < \infty$.

H2: On a $\sigma_{kt}(\cdot) > \underline{\omega}$ pour certains $\underline{\omega} > 0$.

H3: nous avons $a_t \leq C\rho^t$, p.s

cette hypothese nous permet de montrer que les valeurs initiales n'ont pas d'effet sur les propriétés asymptotiques de l'estimateur de $\theta_0^{(k)}$. avec $\Delta_{kt}(\theta^{(k)}) = \tilde{\sigma}_{kt}(\theta^{(k)}) - \sigma_{kt}(\theta^{(k)})$, $a_t = \sup_k \sup_{\theta^{(k)} \in \Theta^{(k)} | \Delta_{kt}(\theta^{(k)})|}$, C et ρ sont des constantes générique avec $C > 0$ et $0 < \rho < 1$.

La "constante" C est autorisée à dépendre de variables antérieures à $t = 0$.

H4: $\theta_0^{(k)}$ est dans l'intérieur de $\Theta^{(k)}$.

H5: $E|\eta_{kt}^*|^{4(1+\delta)} < \infty$, pour $\delta > 0$

H6: pour toute suite réel $(e_i)_{i \geq 1}$, la fonction $\theta^{(k)} \rightarrow \sigma_k(e_1, e_2, \dots; \theta^{(k)})$ est continue et deux fois dérivable. cette hypothèse est considéré pour pouvoir dériver la distribution asymptotique de $\hat{\theta}_n$.

H7: pour $k = 1, \dots, m$ et pour n'importe quel $x \in \mathbb{R}^{dk}$, on a:

$$x' \frac{\partial \sigma_{kt}^2(\theta_0^{(k)})}{\partial \theta^{(k)}} = 0, \quad p.s \quad \Rightarrow \quad x = 0$$

cette hypothèse va etre utiliser pour démontrer l'inversibilité de la matrice de covariance asymptotique.

H8: on a $\sigma_{kt}(\theta_0^{(k)}) = \sigma_{kt}(\theta^{(k)})$ p.s si $\theta_0^{(k)} = \theta^{(k)}$.

Théorème 4.2.1 si H1,H2-H3 sont vérifiées, l'estimateur EbE de $\theta_0^{(k)}$ dans le modèle (4.3) est fortement consistant

$$\hat{\theta}_n^{(k)} \rightarrow \theta_0^{(k)}, \quad p.s. \quad \text{quand } n \rightarrow \infty. \quad (4.5)$$

et si de plus, H4-H7 sont verifiées, alors

$$\sqrt{n}(\hat{\theta}_n^{(k)} - \theta_0^{(k)}) \rightarrow \mathcal{N}\{0, J_{kk}^{-1} I_{kk} J_{kk}^{-1}\}. \quad (4.6)$$

Où

$$I_{kk} = E(\{\eta_{kt}^{*4} - 1\} d_{kt} d'_{kt}), \quad J_{kk} = E(d_{kt} d'_{kt}), \quad d_{kt} = \frac{1}{\sigma_{kt}^2} \frac{\partial \sigma_{kt}^2(\theta_0^{(k)})}{\partial \theta^{(k)}}.$$

Preuve de la consistance:

Démonstration. soit

$$Q_n^{(k)}(\theta^{(k)}) = \frac{1}{n} \sum_{t=1}^n \log \sigma_{kt}^2(\theta^{(k)}) + \frac{\varepsilon_{kt}^2}{\sigma_{kt}^2(\theta^{(k)})} = \frac{1}{n} \sum_{t=1}^n \ell_{kt}(\theta^{(k)}).$$

la forte consistance de $\hat{\theta}_n^{(k)}$ est une conséquence des résultats intermediaire suivant:

- i) $\limsup |Q_n^{(k)}(\theta^{(k)}) + \tilde{Q}_n^{(k)}(\theta^{(k)})| = 0$, p.s
- ii) $E|\ell_{k,1}(\theta_0^{(k)})| < \infty$, et si $\theta^{(k)} \neq \theta_0^{(k)}$, $E\ell_{k,1}(\theta_0^{(k)}) < E\ell_{k,1}(\theta^{(k)})$,
- iii) tous $\theta^{(k)} \neq \theta_0^{(k)}$ a un voisinage $V(\theta^{(k)})$ tel que

$$\liminf \inf \tilde{Q}_n^{(k)}(\theta^*) > \limsup \tilde{Q}_n^{(k)}(\theta_0^{(k)}), p.s.$$

C'est facile de voir que i) est verifié par H2-H3 et l'existence de $E|\varepsilon_{kt}|^s$.
maintenant, par (4.2), on a:

$$E\ell_{kt}(\theta^{(k)}) = E\left\{\frac{\sigma_{kt}^2(\theta_0^{(k)})\eta_{kt}^{*2}}{\sigma_{kt}^2(\theta^{(k)})} + \log \sigma_{kt}^2(\theta^{(k)})\right\} = E\frac{\sigma_{kt}^2(\theta_0^{(k)})}{\sigma_{kt}^2(\theta^{(k)})} + \log \sigma_{kt}^2(\theta^{(k)}).$$

comme $\log \sigma_{kt}^2 < \infty$, on a $E\ell_{kt}(\theta_0^{(k)}) < \infty$, tandis que $E\ell_{kt}(\theta^{(k)}) > -\infty$, pour tous $\theta^{(k)} \in \Theta^{(k)}$, par H2 et en utilisant l'egalité $\log x < x - 1$ et H8, ii) est vérifié.

le thérème d'ergodicité nous assure le dernier point, qui peut etre appliquer pour tous $\theta^{(k)} \in \Theta^{(k)}$ sur la suite $\inf_{\theta_* \in V(\theta^{(k)}) \cap \theta^{(k)}} \ell_{kt}(\theta_*)$, qui est strictement stationaire et ergodique sous H1 et admet une prévision en $[-\infty, +\infty]$. .

□

4.4 comparaison entre l'estimateur EbE et l'estimateur QMV

Une question d'intérêt est de savoir si l'approche EbE implique nécessairement une perte d'efficacité (le prix payé pour sa simplicité) par rapport à une méthode QMV dans laquelle les paramètres de volatilité sont estimés conjointement. Pour pouvoir écrire la quasi-vraisemblance globale, il est nécessaire de spécifier la matrice de corrélation conditionnelle. Parce que nous souhaitons comparer les estimateurs des paramètres de volatilité, nous considérons l'estimateur QMV global de θ_0 en supposant que la matrice R_t est constante et connue.

Un estimateur QMV théorique de θ_0 est défini comme toute solution mesurable θ_n^{QMV} de:

$$\theta_n^{QMV} = \arg \min_{\theta \in \Theta} \frac{1}{n} \sum_{t=1}^n \tilde{\ell}_t(\theta), \quad \tilde{\ell}_t(\theta) = \varepsilon_t' \tilde{H}_t^{-1} \varepsilon_t + \log |\tilde{H}_t|.$$

Où $\tilde{H}_t = \tilde{D}_t R \tilde{D}_t$ et $\tilde{D}_t = \text{diag}(\tilde{\sigma}_{1t}(\theta^{(1)}), \dots, \tilde{\sigma}_{mt}(\theta^{(m)}))$.

soit $R^{-1} = (r_{k\ell}^*)$ et la matrice $(d \times d) M = (M_{k\ell})$ où

$$M_{k\ell} = \tau_{k\ell} J_{k\ell} - \sum_{i,j=1}^m \xi_{ki} \xi_{j\ell} (\kappa_{ij} - 1) J_{ki} J_{ii}^{-1} J_{ij} J_{jj}^{-1} J_{j\ell}$$

et

$$\kappa_{kl} = E(\eta_{kt}^{*2} \eta_{\ell t}^{*2}), \quad \xi_{kl} = \frac{1}{2} \begin{cases} r_{kk}^* + 1 & si \quad k = \ell \\ r_{k\ell} r_{k\ell}^* & si \quad k \neq \ell \end{cases}, \quad \tau_{k\ell} = \sum_{i,j=1}^m r_{ki}^* r_{kj}^* E(\eta_{kt}^* \eta_{it}^* \eta_{jt}^* \eta_{\ell t}^*) - 1.$$

Proposition 4.1 Sous les hypothèses du théorème 4.2.1, l'estimateur QMV des paramètres de volatilité, en supposant que la matrice $R_t = R$ est connu, est asymptotiquement plus efficace (respectivement moins efficace) que l'estimateur EbE ssi la matrice M est définie négative (respectivement définie positive).

Quand R est la matrice identité, $M = 0$ les deux méthodes sont équivalentes, elles produisent le même estimateur.

Dans la mise en œuvre pratique du QMV, la matrice R doit être estimée, ce qui peut réduire la précision des estimateurs des paramètres de volatilité. Il est intéressant de noter que l'estimateur QMV n'est pas toujours asymptotiquement plus efficace que l'estimateur EbE, même dans la situation favorable où R est connu (ce qui n'a aucune conséquence pour l'EbE).

D'autres calculs peuvent montrer que l'estimateur EbE peut être Asymptotiquement supérieur à l'estimateur QMV lorsque la distribution de η_t^* est suffisamment éloignée de la loi gaussienne.

Chapitre 5

Etude par simulation sous le langage R

Dans ce chapitre nous allons nous consacrer aux expériences de Monte-Carlo visant à étudier la performance de l'approche EbE dans un échantillon fini. nous donnerons des exemples de données réelles.

5.1 Etude de Monte-Carlo

Nous allons d'abord illustrer les gains en terme de calcul apportés par l'approche EbE en deux étapes par rapport au QMV complet habituel dans lequel tous les paramètres sont estimés en une seule étape. Nous allons étudier également, pour les modèles CCC et DCC, si les gains de complexité numérique ont un prix en termes d'exactitude des échantillons finis N .

Comparaison en temps et en précision de l'EbE et du QMV complet:

Comparons le coût de calcul de l'EbEE avec celui du QMV dans le cas d'un modèle diagonal $CCC - GARCH(1,1)$ de dimension m , c'est-à-dire sous la spécification (4.1) avec $p = q = 1$ et les matrices diagonales A_1 et B_1 . L'estimateur EbE de tous les paramètres du modèle nécessite m estimations des modèles de type GARCH univariés avec 3 paramètres, plus le calcul de la corrélation empirique des résidus EbE. Le QMV complet nécessite l'optimisation d'une fonction de $3m + m(m - 1)/2$ paramètres. Étant donné que la complexité du temps d'une optimisation augmente généralement rapidement avec la dimension de la fonction objectif, le QMV devrait être beaucoup plus coûteux que l'EbEE en termes de

temps de calcul. Les deux estimateurs ont été ajustés sur des simulations de longueur $n = 2000$ du modèle $CCC - GARCH(1,1)$ avec les spécification (4.1) avec $A_1 = 0.05I_m$ et $B_1 = 0.9I_m$ (ces valeurs sont proches de celles généralement adaptées aux séries réelles). La matrice de corrélation utilisée pour les simulations est $R = I_m$, mais les termes sous-diagonaux $m(m-1)/2$ de R ont été estimés, ainsi que les $3m$ autres paramètres du modèle. La distribution de η_t est gaussienne, qui a peu d'impact sur les temps de calcul, mais devrait donner un avantage au QMV en termes de précision. Le tableau 1 compare les temps de calcul effectifs requis par les deux estimateurs en fonction de la dimension m . Comme prévu, la comparaison des processeurs est clairement en faveur de l'EbEE. Notez que ces temps de calcul ont été obtenus en utilisant un seul processeur. Étant donné que l'EbEE est clairement facilement parallélisable (en utilisant un processeur pour chacune des optimisations m), l'avantage de l'EbEE devrait être encore plus prononcé avec une implémentation multiprocessus.

Le code R permettant d'exécuter le programme qui compare les deux estimateurs est :

```
library(compiler)
library(numDeriv)
# vech0
vech0<- function (A){
  d<-nrow(A)
  res<-rep(0,(d*(d-1)/2))
  ind<-0
  for(j in 1:(d-1)){
    for(i in (j+1):d){
      ind<-ind+1
      res[ind]<-A[i,j]}}
  res
# inverse de vech0
inv.vech0<- function (rho){
  m<-length(rho)
  d<-(1+sqrt(1+8*m))/2
```

```

A<-diag(d)
ind<-0
for(j in 1:(d-1)){
  for(i in (j+1):d){
    ind<-ind+1
    A[i,j]<-rho[ind]
    A[j,i]<-rho[ind]
  }}
A
# racine carree d'un matrice symmetrique semi-definie positive
Sqrt <- function(Sigma) {
  n<- nrow(Sigma)
  sp<-eigen(Sigma)
  sp$vector%*%sqrt(diag(sp$values))
# matrice de correlation de m valeurs consécutives d'un AR(1)
R.AR1 <- function(m,rho) {
  R<-diag(rep(1,m))
  for (j in 1:(m-1)){
    for (i in (j+1):m){
      R[i,i-j]<-rho^j
      R[i-j,i]<-R[i,i-j]
    }
  }
  R
# simule un GARCH(1,1)-CCC diagonal avec bruit Student de variance R.mat
Mgarch.sim<- function (n, omega, alpha, beta, R.mat, nu=Inf, valinit=500) {
  m<-length(omega)
  cst<-1
  if(nu>2&nu!=Inf)cst<-sqrt((nu-2)/nu)
  #eta <- matrix(cst*rt(m*(n+valinit),nu),ncol=m,nrow=(n+valinit))%*%t(Sqrt(R.mat))
  eta <- matrix(rnorm(m*(n+valinit)),ncol=m,nrow=(n+valinit))%*%t(Sqrt(R.mat))
  eps<-matrix(0,nrow=n+valinit,ncol=m)
  ht<-matrix(0,nrow=n+valinit,ncol=m)
  for (t in 2:(n+valinit)) {
    ht[t,]<-omega+alpha*(eps[t-1,]^2)+beta*ht[t-1,]
    eps[t,]<-sqrt(ht[t,])*eta[t,]
  }
}

```

```

}
return(eps[(valinit+1):(n+valinit),])}
Mgarch.sim<-cmpfun(Mgarch.sim)
# estime une equation d'un MGARCH(1,1)-CCC diagonal
objf <- function(x, eps, eps2, n, r, tol=sqrt(.Machine$double.eps)){
  omega <- x[1]
  alpha<- x[2]
  beta <- x[3]
  sigma2<-rep(0,n)
sigma2[1]<-var(eps[1:r])
for (t in 2:n) sigma2[t]<-omega+alpha*eps2[t-1]+beta*sigma2[t-1]
sigma2.mod<-pmax(sigma2[(r+1):n],tol)
qml <- mean(log(sigma2.mod)+eps[(r+1):n]**2/sigma2.mod)
  qml }
objf<-cmpfun(objf)
#estim.1equa.Mgarch11<- function(omega,alpha,beta,eps,r)
{
  n <- length(eps)
eps2<-eps^2
valinit<-c(omega,alpha,beta)
res <- nlminb(valinit,objf, lower=c(0.00000000000001,rep(0,2)),
  upper=c(Inf,rep(0.9999,2)), eps=eps, eps2=eps2, n=n,r=r)
omega<-res$par[1]
alpha<-res$par[2]
beta<-res$par[3]
sigma2<-rep(0,n)
sigma2[1]<-var(eps[1:r])
for (t in 2:n) sigma2[t]<-omega+alpha*eps2[t-1]+beta*sigma2[t-1]
eta <- eps[(r+1):n]/sqrt(sigma2[(r+1):n])
list(coef=res$par, minimum=res$objective,eta=eta)
estim.1equa.Mgarch11<-cmpfun(estim.1equa.Mgarch11)
# estime EbE un MGARCH(1,1)-CCC diagonal
estim.EbEE.Mgarch11<- function(Omega,Alpha,Beta,eps,r=10)
{

```

```

m<-length(Omega)
n<-length(eps[,1])
Res<-matrix(nrow=(n-r),ncol=m)
Omega.est<-Omega
Alpha.est<-Alpha
Beta.est<-Beta
for(j in 1:m){
  res<-estim.1equa.Mgarch11(Omega[j],Alpha[j],Beta[j],eps[,j],r)
  Omega.est[j]<-res$coef[1]
  Alpha.est[j]<-res$coef[2]
  Beta.est[j]<-res$coef[3]
  Res[,j]<-res$eta
}
R<-cor(Res)
c(Omega.est,Alpha.est,Beta.est,vech0(R))
estim.EbEE.Mgarch11<-cmpfun(estim.EbEE.Mgarch11)
##### estime un CCC-GARCH(1,1) diagonal par QMLE
objfCCC <- function(x, eps, n, m, r, tol=sqrt(.Machine$double.eps)){
  Omega <- x[1:m]
  Alpha<- x[(m+1):(2*m)]
  Beta <- x[(2*m+1):(3*m)]
  R<-inv.vech0(x[(3*m+1):(3*m+m*(m-1)/2)])
  #{if(kappa(R)<1/tol&det(R)>tol){
  Rinv<-solve(R)
  log.detR<-log(det(R))
  ht<-matrix(0,nrow=m,ncol=n)
  Dtin<-matrix(0,nrow=m,ncol=m)
  l<-as.numeric(n)
  ht[,1]<-pmax(diag(var(eps[1:r,])),tol)
  Dtin<-diag(1/sqrt(ht[,1]))
  l[1]<-eps[1,]%*%Dtin%*%Rinv%*%Dtin%*%eps[1,]+log.detR+sum(log(ht[,1]))
  for (t in 2:n){
    ht[,t]<-pmax(Omega+Alpha*eps[t-1,]^2+Beta*ht[,t-1],tol)
    Dtin<-diag(1/sqrt(ht[,t]))
  }
}

```

```

l[t]<-eps[t,]*%*%Dtinv*%*%Rinv*%*%Dtinv*%*%eps[t,]+log.detR+sum(log(ht[,t]))
  qml <- mean(l[(r+1):n])
return(qml)
#}
#else
#{
#return(Inf)}
#objfCCC<-cmpfun(objfCCC)
CCC.GARCH.QMLE<-function(eps, Omegainit,
  Alphainit,Betainit, Rinit, tol=sqrt(.Machine$double.eps),r=10){
m<-ncol(eps)
n<-nrow(eps)
valinit<-c(Omegainit,Alphainit,Betainit,vech0(Rinit))
born.inf<-c(rep(0.00000000000001,m),rep(0,2*m),rep(-1+tol,m*(m-1)/2))
born.sup<-c(rep(Inf,m),rep(0.9999,2*m),rep(1-tol,m*(m-1)/2))
res <- nlminb(valinit,objfCCC, lower=born.inf,
  upper=born.sup, eps=eps, n=n,m=m,r=r)
Omega<-res$par[1:m]
Alpha<-res$par[(m+1):(2*m)]
Beta<-res$par[(2*m+1):(3*m)]
R<-inv.vech0(res$par[(3*m+1):(3*m+m*(m-1)/2)])
#list(Omega=Omega,A=Alpha,B=Beta,R=R)
res$par
CCC.GARCH.QMLE<-cmpfun(CCC.GARCH.QMLE)
##### estime la matrice J d'information
objfCC <- function(x, eps, n, mm, r, tol=sqrt(.Machine$double.eps)){
  Omega <- x[1:mm]
  Alpha<- x[(mm+1):(2*mm)]
  Beta <- x[(2*mm+1):(3*mm)]
  R<-inv.vech0(x[(3*mm+1):(3*mm+mm*(mm-1)/2)])
Rinv<-solve(R)
log.detR<-log(det(R))
ht<-matrix(0,nrow=mm,ncol=n)
Dtinv<-matrix(0,nrow=mm,ncol=mm)

```

```

l<-as.numeric(n)
ht[,1]<-pmax(diag(var(eps[1:r,])),tol)
Dtin<-diag(1/sqrt(ht[,1]))
l[1]<-eps[1,]%*%Dtin%*%Rinv%*%Dtin%*%eps[1,]+log.detR+sum(log(ht[,1]))
for (t in 2:n){
  ht[,t]<-pmax(Omega+Alpha*eps[t-1,]^2+Beta*ht[,t-1],tol)
  Dtin<-diag(1/sqrt(ht[,t]))
  l[t]<-eps[t,]%*%Dtin%*%Rinv%*%Dtin%*%eps[t,]+log.detR+sum(log(ht[,t]))
  qml <- mean(l[(r+1):n])
}
qml}
objfCC<-cmpfun(objfCC)
J.n <- function(x, eps, n, m, r=10){
Jn<-hessian(objfCC,x,eps=eps,n=n,mm=m,r=r)
Jn/2}
J.n<-cmpfun(J.n)
# estimateur plus rapide de matrice J d'information
VaR.R <- function(x, eps, n, mm, r=10, tol=sqrt(.Machine$double.eps)){
  Omega <- x[1:mm]
  Alpha<- x[(mm+1):(2*mm)]
  Beta <- x[(2*mm+1):(3*mm)]
  R<-inv.vech0(x[(3*mm+1):(3*mm+mm*(mm-1)/2)])
Rinv<-solve(R)
log.detR<-log(det(R))
ht<-matrix(0,nrow=mm,ncol=n)
s1<-3*mm
s0<-length(x)
s2<-s0-s1
der.ht<-array(0,dim=c(s1,mm,n))
Dtin<-matrix(0,nrow=mm,ncol=mm)
l<-as.numeric(n)
der.l<-matrix(0,nrow=s0,ncol=n)
ht[,1]<-pmax(diag(var(eps[1:r,])),tol)
Dtin<-diag(1/sqrt(ht[,1]))
l[1]<-eps[1,]%*%Dtin%*%Rinv%*%Dtin%*%eps[1,]+log.detR+sum(log(ht[,1]))

```

```

for (t in 2:n){
  ht[,t]<-pmax(Omega+Alpha*eps[t-1,]^2+Beta*ht[,t-1],tol)
  Big<-rbind(diag(rep(1,mm)),diag(eps[t-1,]^2),diag(ht[,t-1]))
  der.ht[, ,t]<-Big+der.ht[, ,t-1]%%diag(Beta)
  Dtin<-diag(1/sqrt(ht[,t]))
  l[t]<-eps[t,]%%Dtin%%Rinv%%Dtin%%eps[t,]+log.detR+sum(log(ht[,t]))
  for (i in 1:s1){
    der.Di<-diag(der.ht[i, ,t]/(2*sqrt(ht[,t])))
    c1<--eps[t,]%%Dtin%%der.Di%%Dtin%%Rinv%%Dtin%%eps[t,]
    c1<-c1-eps[t,]%%Dtin%%Rinv%%Dtin%%der.Di%%Dtin%%eps[t,]
    c1<-c1+2*sum(diag(Dtin%%der.Di))
    der.l[i,t]<-c1}
  for (i in 1:s2){
    der.R<-inv.vech0(diag(rep(1,s2))[ ,i])-diag(rep(1,mm))
    c1<--eps[t,]%%Dtin%%Rinv%%der.R%%Rinv%%Dtin%%eps[t,]
    c1<-c1+sum(diag(Rinv%%der.R))
    der.l[s1+i,t]<-c1
  }
  J<-var(t(der.l))
  J/4
  VaR.R<-cmpfun(VaR.R)
  #####
  set.seed(11)
  n<-2000
  nu<-5
  nu<-Inf
  m.min<-2
  m.max<-12
  for(m in m.min:m.max){
    Omega0<-rep(0.01,m)
    Alpha0<-rep(0.05,m)
    Beta0<-rep(0.90,m)
    R0<-diag(rep(1,m))
    eps<-Mgarch.sim(n, Omega0, Alpha0, Beta0, R0, nu=nu)
    vartheta0<-c(Omega0,Alpha0,Beta0,vech0(R0))
  }
}

```

```

print(m)
J<-VaR.R(vartheta0, eps, n, m)
Omegainit<-rep(0.5,m)
Alphainit<-rep(0.1,m)
Betainit<-rep(0.7,m)
Rinit<-diag(rep(1,m))
EbEE<-estim.EbEE.Mgarch11(Omegainit,Alphainit,Betainit,eps)
diff<-EbEE-vartheta0
error.EbEE<-n*as.numeric(diff%*%J%*%diff)
QMLE<-CCC.GARCH.QMLE(eps, Omegainit,Alphainit,Betainit, Rinit)
diff<-QMLE-vartheta0
error.QMLE<-n*as.numeric(diff%*%J%*%diff)
print(error.QMLE/error.EbEE)
#####
set.seed(11)
for(m in m.min:m.max){
  Omega0<-rep(0.01,m)
  Alpha0<-rep(0.05,m)
  Beta0<-rep(0.90,m)
  rho0<-0.9
  R0<-R.AR1(m,rho0)
  R0<-diag(rep(1,m))
  eps<-Mgarch.sim(n, Omega0, Alpha0, Beta0, R0, nu=nu)
  vartheta0<-c(Omega0,Alpha0,Beta0,vech0(R0))
  print(m)
  #ti0<-system.time(VaR.R(vartheta0, eps, n, m))
  #print(ti0)
  Omegainit<-rep(0.5,m)
  Alphainit<-rep(0.1,m)
  Betainit<-rep(0.7,m)
  Rinit<-diag(rep(1,m))
  #Omegainit<-diag(var(eps))*(1-Alphainit-Betainit)
  ti<-system.time(estim.EbEE.Mgarch11(Omegainit,Alphainit,Betainit,eps))
  print(ti)

```

```

ti2<-system.time(CCC.GARCH.QMLE(eps, Omegainit,Alphainit,Betainit, Rinit))
print(ti2)
#####
set.seed(11)
for(m in c(50,100,200,400,800)){
  Omega0<-rep(0.01,m)
  Alpha0<-rep(0.05,m)
  Beta0<-rep(0.90,m)
  R0<-diag(rep(1,m))
  eps<-Mgarch.sim(n, Omega0, Alpha0, Beta0, R0, nu=nu)
  vartheta0<-c(Omega0,Alpha0,Beta0,vech0(R0))
  print(m)
  Omegainit<-rep(0.5,m)
  Alphainit<-rep(0.1,m)
  Betainit<-rep(0.7,m)
  Rinit<-diag(rep(1,m))
  ti<-system.time(estim.EbEE.Mgarch11(Omegainit,Alphainit,Betainit,eps))
  print(ti)
}

```

TAB. 5.1 – *Table 1: Le temps de calcul des deux estimateurs (temps du processeur en secondes) de l'estimateur EbE par rapport a l'estimateur QMV (ND signifie "Non disponible" en raison de l'impossibilité de calculer le QMV) pour le modèles CCC-GARCH (1,1) m-dimensionnel*

Dim m	2	3	4	5	6	7	8	9
Nbre de param	7	12	18	25	33	42	52	63
processeur de l'Ebe	0.57	0.88	1.18	1.31	1.52	1.85	2.04	2.37
processeur du QMV	32.49	100.78	123.33	215.38	317.85	617.33	876.52	1113.68
le ratio	57.00	114.52	104.52	164.41	209.11	333.69	429.67	469.91
Dim m	10	11	12	50	100	200	400	800
Nbre de param	75	88	102	1375	5250	20500	81000	322000
processeur de l'Ebe	2.82	2.98	3.49	13.67	27.89	56.58	110.00	226.32
processeur du QMV	1292.34	1520.60	1986.38	ND	ND	ND	ND	ND
le ratio	458.28	510.27	569.16	ND	ND	ND	ND	ND

Interprétation des résultats:

On remarque que quand la dimension m augmente donc logiquement le nombre de paramètre a estimé augmente aussi, le temps qu'a mis l'estimateur Ebe a estimer tous les paramètres est petit par rapport à l'estimateur QMV.

Tant que la dimension du modèle M-GARCH augmente, on constate que la différence entre les temps des processeurs des estimateur est importante d'ailleurs le ratio qui est le rapport du temps du QMV sur le temps de l'EbeE croit très rapidement au fil de l'augmentation de la dimension m .

On peut dire donc que la méthode EbE est plus pratique en terme de temps que la méthode QMV.

5.2 Estimation des volatilités individuelles d'un DCC par EbEE

Afin d'illustrer la capacité de l'EbeE à estimer les volatilités individuelles d'un DCC, nous avons fait l'expérience Monte Carlo suivante. Nous avons d'abord simulé une suite iid (η_t) , avec $m = 4$ composants indépendants, normalisés de telle sorte que $Var(\eta_t) = I_m$. Nous avons ensuite simulé la matrice de corrélations R_t et l'innovations $\eta_t^* = R_t^{1/2}\eta_t$, où S est la matrice de corrélation de Toeplitz avec l'élément $0.3i$ sur i -ième subdiagonal $\alpha = 0.04$ et $\beta = 0.95$. Pour ces récursions, nous avons pris les valeurs initiales $Q_0 = S$ et $\eta_0^* = 0$. Nous avons ensuite généré la suite $\varepsilon_t = D_t\eta_t^*$, où les éléments de D_t sont obtenue à partir de (2.16) avec $p = q = 1$, $\omega = (0.01, \dots, 0.01)'$, $A = A_1$ la matrice $m \times m$ avec les éléments 0.02 et $B = B_1 = diag(0.91, \dots, 0.91)$. Nous avons enlevé les 500 premières valeurs simulées pour atténuer l'effet des valeurs initiales. Le tableau 3 présente les estimations EbE des coefficients de volatilité ω , A et B , ainsi que la variance des innovations EbE $R^* = Var(\eta_t^*)$, plus de 100 répliques indépendantes de longueur $n = 2000$ du modèle DCC. On constate que le biais d'estimation est très faible. Nous avons également vérifié que les erreurs (RMSE) diminuent lorsque la taille de l'échantillon n augmente et qu'elles ne sont pas trop sensibles aux paramètres nuisants α, β et S impliqués dans la séquence (R_t) des matrices DCC.

Table 2: EbEE moyenne de plus de 100 répétitions du modèle DCC

avec $m = 4$.									
ω	A				diag(B)	R*			
0.013	0.017	0.021	0.020	0.021	0.905	1.000	0.276	0.087	0.023
0.008	0.013	0.011	0.011	0.012	0.024	-	0.096	0.103	0.096
0.014	0.020	0.018	0.022	0.021	0.902	0.278	1.000	0.268	0.086
0.010	0.011	0.012	0.013	0.013	0.025	0.096	-	0.087	0.111
0.013	0.022	0.022	0.016	0.021	0.906	0.087	0.268	1.000	0.299
0.010	0.012	0.011	0.011	0.014	0.027	0.103	0.087	-	0.098
0.018	0.021	0.022	0.021	0.018	0.899	0.023	0.086	0.299	1.000
0.039	0.011	0.011	0.011	0.011	0.064	0.096	0.111	0.098	-

Le code R permettant d'exécuter le programme qui donne les résultats du tableau ci dessus est :

```
library(compiler)
library(numDeriv)
# vech0
vech0<- function (A){
d<-nrow(A)
res<-rep(0,(d*(d-1)/2))
ind<-0
for(j in 1:(d-1)){
for(i in (j+1):d){
ind<-ind+1
res[ind]<-A[i,j]}}
res
}
# inverse de vech0
inv.vech0<- function (rho){
m<-length(rho)
d<-(1+sqrt(1+8*m))/2
A<-diag(d)
```

```

ind<-0
for(j in 1:(d-1)){
  for(i in (j+1):d){
    ind<-ind+1
    A[i,j]<-rho[ind]
    A[j,i]<-rho[ind]
  }
}
A
}

# racine carree d'un matrice symmetrique semi-definie positive
Sqrt <- function(Sigma) {
  n<- nrow(Sigma)
  sp<-eigen(Sigma)
  sp$vector%*%sqrt(diag(sp$values))
}

# matrice de correlation de m valeurs consécutives d'un AR(1)
R.AR1 <- function(m,rho) {
  R<-diag(rep(1,m))
  for (j in 1:(m-1)){
    for (i in (j+1):m){
      R[i,i-j]<-rho^j
      R[i-j,i]<-R[i,i-j]
    }
  }
}

# simule un GARCH(1,1)-DCC (de Engle ou de Aielli) avec beta diagonal et bruit Student
GarchDCC.sim<- function (n, omega, Alpha, beta, aalpha,bbeta, S, nu=Inf, valinit=50)
m<-length(omega)
Qt<-array(dim=c(n+valinit,m,m))
Qt[1,,]<-S
Qt.star<-diag(sqrt(diag(Qt[1,,])))
Qt.star.inv<-diag(sqrt(1/pmax(diag(Qt[1,,]),tol)))
Rt<-Qt
Rt[1,,]<-Qt.star.inv%*%Qt[1,,]%*%Qt.star.inv
cst<-1
if(nu>2&nu!=Inf)cst<-sqrt((nu-2)/nu)

```

```

eta <- matrix(cst*rt(m*(n+valinit)),nu),nrow=(n+valinit),ncol=m)
#plot.ts(eta) var(eta) var(eta.star)
eps<-matrix(0,nrow=(n+valinit),ncol=m)
eta.star<-matrix(0,nrow=(n+valinit),ncol=m)
ht<-matrix(0,nrow=(n+valinit),ncol=m)
eta.star[1,]<-eta[1,]%*%t(Sqrt(Rt[1,,]))
for (t in 2:(n+valinit)) {
ht[t,]<-omega+as.vector(Alpha%*(eps[t-1,]^2))+beta*ht[t-1,]
#
{if(Aielli){
Qt[t,,]<-(1-aalpha-bbeta)*S+aalpha*Qt.star%*eta.star[t-1,]%*%t(eta.star[t-1,])%*Qt.
}
else{
Qt[t,,]<-(1-aalpha-bbeta)*S+aalpha*eta.star[t-1,]%*%t(eta.star[t-1,])+bbeta*Qt[t-1,,]
}}
Qt.star<-diag(sqrt(diag(Qt[t,,])))
Qt.star.inv<-diag(sqrt(1/pmax(diag(Qt[t,,]),tol)))
Rt[t,,]<-Qt.star.inv%*Qt[t,,]%*Qt.star.inv
eta.star[t,]<-eta[t,]%*%t(Sqrt(Rt[t,,]))
eps[t,]<-sqrt(ht[t,])*eta.star[t,]}
#return(eps[(valinit+1):(n+valinit),,])}
list(sim=eps[(valinit+1):(n+valinit),,], cor=Rt[(valinit+1):(n+valinit),,])}
#omega<-c(0.01,0.01);Alpha<-matrix(c(0.03,0.01,0.01,0.03),nrow=2)
#beta<-c(0.8,0.8);S<-matrix(c(1,0.3,0.3,1),nrow=2)
#aalpha<-0.01;bbeta<-0.99-aalpha
# n<-800; nu=7
#res<-GarchDCC.sim(n, omega, Alpha, beta, aalpha,bbeta, S, nu=nu)
#plot.ts(res$sim)
#plot.ts(res$cor[,1,2],ylim=c(-1,1))
GarchDCC.sim<-cmpfun(GarchDCC.sim)
# estime une equation d'un MGARCH(1,1)-CCC diagonal
objf <- function(x, eps, eps2,m, n, r, tol=sqrt(.Machine$double.eps)){
  omega <- x[1]
  Alpha<- x[2:(m+1)]

```

```

        beta <- x[m+2]
        sigma2<-rep(0,n)
sigma2[1]<-var(eps[1:r])
for (t in 2:n) sigma2[t]<-omega+as.numeric(Alpha%%eps2[t-1,])+beta*sigma2[t-1]
sigma2.mod<-pmax(sigma2[(r+1):n],tol)
qml <- mean(log(sigma2.mod)+eps[(r+1):n]**2/sigma2.mod)
      qml }

objf<-cmpfun(objf)
#
estim.1equa.Mgarch11<- function(omega,Alpha,beta,eps,eps2,m,r)
{
  n <- length(eps)
valinit<-c(omega,Alpha,beta)
res <- nlminb(valinit,objf, lower=c(0.00000000000001,rep(0,2)),
  upper=c(Inf,rep(0.9999,2)), eps=eps, eps2=eps2,m=m, n=n,r=r)
omega<-res$par[1]
alpha<-res$par[2:(m+1)]
beta<-res$par[m+2]
sigma2<-rep(0,n)
sigma2[1]<-var(eps[1:r])
for (t in 2:n) sigma2[t]<-omega+as.numeric(alpha%%eps2[t-1,])+beta*sigma2[t-1]
eta <- eps[(r+1):n]/sqrt(sigma2[(r+1):n])
list(coef=res$par, minimum=res$objective,eta=eta)}
estim.1equa.Mgarch11<-cmpfun(estim.1equa.Mgarch11)
# estime EbE les volatilités d'un MGARCH(1,1)-DCC
#omega<-omegainit;Alpha<-Alphainit;beta<-betainit
estim.EbEE.Mgarch11<- function(omega,Alpha,beta,eps,r=10)
{ m<-length(omega)
n<-length(eps[,1])
eps2<-eps^2
Res<-matrix(nrow=(n-r),ncol=m)
omega.est<-omega
Alpha.est<-Alpha

```

```

beta.est<-beta
for(j in 1:m){
res<-estim.1equa.Mgarch11(omega[j],Alpha[j,],beta[j],eps[,j],eps2,m,r)
omega.est[j]<-res$coef[1]
Alpha.est[j,]<-res$coef[2:(m+1)]
beta.est[j]<-res$coef[m+2]
Res[,j]<-res$eta }
R<-cor(Res)
c(omega.est,Alpha.est,beta.est,vech0(R))
}estim.EbEE.Mgarch11<-cmpfun(estim.EbEE.Mgarch11)
#####
set.seed(11)
niter<-100
n<-2000
nu<-Inf
nu<-9
m<-4
rho0<-0.3
omega0<-rep(0.01,m)
Alpha0<-matrix(rep(0.02,m^2),ncol=m)
beta0<-rep(0.91,m)
S0<-R.AR1(m,rho0)
aalpha0<-0.04
bbeta0<-0.99-aalpha0
omegainit<-rep(0.02,m)
Alphainit<-matrix(rep(0.03,m^2),ncol=m)
betainit<-rep(0.7,m)
vartheta0<-c(omega0,Alpha0,beta0,vech0(S0))
res<-vartheta0
#R0<-diag(rep(1,m))
for(iter in 1:niter){
eps<-GarchDCC.sim(n, omega0, Alpha0, beta0, aalpha0,bbeta0, S0, nu=nu, Aielli=FALSE)
#plot.ts(eps) acf(eps^2)
EbEE<-estim.EbEE.Mgarch11(omegainit,Alphainit,betainit,eps)

```

```

res<-rbind(res,EbEE)}
omega<-colMeans(res[2:(niter+1),,])[1:m]
A<-matrix(colMeans(res[2:(niter+1),,])[(m+1):(m+m^2)],nrow=m)
b<-colMeans(res[2:(niter+1),,])[(m+m^2+1):(m+m^2+m)]
R<-inv.vech0(colMeans(res[2:(niter+1),,])[(m+m^2+m+1):(m+m^2+m+m*(m-1)/2)])
sd<-sqrt(diag(var(res[2:(niter+1),,])))
omega.sd<-sd[1:m]
A.sd<-matrix(sd[(m+1):(m+m^2)],nrow=m)
b.sd<-sd[(m+m^2+1):(m+m^2+m)]
R.sd<-inv.vech0(sd[(m+m^2+m+1):(m+m^2+m+m*(m-1)/2)])
for(i in 1:m)
print(round(c(omega[i],A[,i],b[i],R[,i]),digits=3))
print(round(c(omega.sd[i],A.sd[,i],b.sd[i],R.sd[,i]),digits=3))}
bias<-colMeans(res[2:(niter+1),,])-vartheta0
npar<-m*(m+2)+m*(m-1)/2
bias<-sapply(1:npar, function(i) mean(as.numeric(res[2:(niter+1),i])-vartheta0[i]))

```

Conclusion

Le but de ce travail a été de mettre en évidence les modèles non linéaires et l'hétéroscédasticité conditionnelle qui possède, maintenant, des outils puissants d'analyse. Cette modélisation est fondée sur des bases théoriques solides nécessaire pour les séries chronologiques présentant une dynamique non linéaire. Le concept de variance conditionnelle a commencé à jouer un grand rôle au début des années quatre vingt avec l'article fondateur de Engle (1982). Il caractérise les modèles venus élargir la classe des modèles classiques fondés essentiellement sur une structure de dépendance linéaire entre une variable à un instant t et ses valeurs passées et celles d'un bruit blanc et de ses valeurs passées.

Par ailleurs, il est important de noter que les séries financières sont aussi caractérisées par une volatilité non stationnaire et par des phénomènes d'asymétrie qui ne peuvent pas être pris en compte par les modélisations classiques. C'est dans ce contexte que nous nous sommes intéressés aux modèles M-GARCH. Après présentation des différentes approches qui s'offrent à nous pour estimer les paramètres du modèle M-GARCH, entre autres la méthode QMV et la méthode EbE. La consistance et la normalité asymptotique du l'EbE ont été établie. Nous nous sommes aussi basés sur des résultats de simulation afin d'illustrer les résultats théoriques.

Ce travail ouvre de très nombreuses voies de recherche en particulier dans l'estimation des modèles M-GARCH.

Annexe

5.3 Martingale

Définition 5.1 (Martingale). Soit $(X_t)_{t \in \mathbb{N}}$ une suite de variables aléatoires réelles (v.a.r) sur un espace probabilisé $(\Omega, \mathcal{F}, P)$, et $(\mathcal{F}_t)_{t \in \mathbb{N}}$ est une suite de tribus. La suite $\{(X_t, \mathcal{F}_t) : t = 1, 2, \dots\}$ est une martingale si et seulement si :

1. $\mathcal{F}_t \subset \mathcal{F}_{t+1}$;
2. X_t est $\mathcal{F}_t$ -mesurable ;
3. $\mathbb{E}(|X_t|) < \infty$;
4. $\mathbb{E}(X_{t+1} \mid \mathcal{F}_t) = X_t$.

Quand on dit que $(X_t)_{t \in \mathbb{N}}$ est une martingale, on prend implicitement $\mathcal{F}_t = \sigma(X_s, s \leq t)$, c'est-à-dire la tribu engendrée par les valeurs passées et présentes.

Définition 5.2 (Différence de martingale). Soient $(X_t)_{t \in \mathbb{N}}$ une suite de variables aléatoires réelles (v.a.r), et $(\mathcal{F}_t)_{t \in \mathbb{N}}$ est une suite de tribus. La suite $\{(X_t, \mathcal{F}_t) : t = 1, 2, \dots\}$ est une différence de martingale (ou une suite d'accroissements de martingale) si et seulement si :

1. $\mathcal{F}_t \subset \mathcal{F}_{t+1}$;
2. X_t est $\mathcal{F}_t$ -mesurable ;
3. $\mathbb{E}(|X_t|) < \infty$;
4. $\mathbb{E}(X_{t+1} \mid \mathcal{F}_t) = 0$.

5.4 Ergodicité

On dit qu'une suite stationnaire est ergodique si elle satisfait la loi forte des grands nombres. Certaines transformations de suites ergodiques restent ergodiques.

Théorème 5.4.1. *Si $(Z_t)_{t \in \mathbb{Z}}$, est une suite fortement stationnaire et ergodique.*

et si $f : \mathbb{R}^\infty \longrightarrow \mathbb{R}$ est une fonction mesurable.

et soit $Y_t = f(\cdots, Z_{t-1}, Z_t, Z_{t+1}, \cdots)$, alors $(Y_t)_{t \in \mathbb{Z}}$ reste une suite fortement stationnaire et ergodique.

Théorème 5.4.2 (d'ergodicité). Si $(Z_t)_{t \in \mathbb{Z}}$ est strictement stationnaire et ergodique, si f est mesurable et si $\mathbb{E}\{|f(\cdots, Z_{t-1}, Z_t, Z_{t+1}, \cdots)|\} < \infty$, alors:

$$\frac{1}{n} \sum_{t=1}^n f(\cdots, Z_{t-1}, Z_t, Z_{t+1}, \cdots) \longrightarrow \mathbb{E}\{f(\cdots, Z_{t-1}, Z_t, Z_{t+1}, \cdots)\} \quad p.s.$$

5.5 Théorème de Chung-Fuchs

Théorème 5.5.1 (sur les marches aléatoires). Si $X_1, \cdots, X_n$ est une suite iid telle que $\mathbb{E}(X_1) = 0$ et $E|X_1| > 0$, alors p.s.

$$\limsup_{n \rightarrow \infty} \sum_{i=1}^n X_i = +\infty,$$

et

$$\liminf_{n \rightarrow \infty} \sum_{i=1}^n X_i = -\infty.$$

5.6 Critère de Cauchy

Définition 5.6.1 (Critère de Cauchy pour la convergence d'une suite de terme $a_n \geq 0$). Soit $\lambda = \limsup a_n^{\frac{1}{n}}$.

Si $\lambda < 1 \Rightarrow \sum_{n=1}^{\infty} a_n < \infty$,

si $\lambda > 1 \Rightarrow \sum_{n=1}^{\infty} a_n = +\infty$.

5.7 La Régression Linéaire Multiple

En statistiques et en économétrie, un modèle de régression linéaire est un modèle de régression d'une variable expliquée sur une ou plusieurs variables explicatives dans lequel on fait l'hypothèse que la fonction qui relie les variables explicatives à la variable expliquée est linéaire dans ses paramètres, formellement, on modélise la relation entre une variable aléatoire y et un vecteur de variables aléatoires x .

Définition 5.7.1. De manière générale, le modèle linéaire peut s'écrire de la manière

suivante :

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \mu,$$

où la variable y est appelée la variable expliquée ou variable endogène, et les variables $(x_1, x_2, \cdots, x_k)$ sont appelées variables explicatives, variables exogènes ou encore prédicteurs, et μ est appelé terme d'erreur ou perturbation .

5.7.1 Notations

On rencontre principalement trois types de notations :

Notation simple

On considère le modèle pour l'individu i . Pour chaque individu, la variable expliquée s'écrit comme une fonction linéaire des variables explicatives :

$$y_i = \beta_0 + \beta_1 x_{1,i} + \cdots + \beta_k x_{k,i} + \mu_i.$$

Notation vectorielle

La notation vectorielle est similaire à la notation simple mais on utilise la notation vectorielle pour synthétiser la notation. Cette notation est pratique lorsqu'il y a un grand nombre de variables explicatives. On définit $\beta = (\beta_0, \beta_1, \cdots, \beta_k)$ le vecteur des paramètres du modèle, et $x_i = (1, x_{1,i}, \cdots, x_{k,i})$ le vecteur des variables explicatives pour l'individu i , le modèle se réécrit alors de la manière suivante :

$$y_i = x_i' \beta + \mu_i.$$

Notation matricielle

Enfin, on rencontre aussi souvent une notation matricielle, ici, on écrit le modèle pour chacun des n individus présents dans l'échantillon, le modèle s'écrit alors :

$$y = X\beta + \mu,$$

avec :

$$y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, X = \begin{pmatrix} x_1' \\ x_2' \\ \vdots \\ x_n' \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & \cdots & x_{1k} \\ 1 & x_{21} & \cdots & x_{2k} \\ \vdots & \ddots & \vdots & \\ 1 & x_{n1} & \cdots & x_{nk} \end{pmatrix}, \beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{pmatrix}, \mu = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{pmatrix}.$$

Bibliographie

- [1] Bose, A. and Mukherjee, K. (2003) Estimating the ARCH parameters by solving linear equations. *Journal of Time Series Analysis* 24, 127–136.
- [2] Christophe Hurlin (2004) *Econométrie pour la Finance, Modèles ARCH - GARCH*.
- [3] Francq , C. and Zakoian, J.-M. (2008) Can one really estimate nonstationary GARCH models?
- [4] Francq , C. and Zakoian, J.-M. (2010) *GARCH models : Structure, Statistical Inference and Financial Applications*, 1st edition. WILEY.
- [5] Jensen, S.T. and Rahbek, A. (2004a) Asymptotic normality of the QMLE estimator of ARCH in the nonstationary case. *Econometrica* 72, 641–646.
- [6] Zakoian, J.-M. (1992) Modèles ARCH : une revue de la littérature. *Journal de la société statistique de Paris*, tome 133, n° 1-2(1992), p.40-57.
