



وزارة التعليم العالي والبحث العلمي



**BADJI MOKHTAR-ANNABA  
UNIVERSITY  
UNIVERSITE BADJI MOKHTAR-  
ANNABA**

**جامعة باجي مختار  
- عنابة -**

**Faculté des Sciences**

**Année : 2018**

**Département de Mathématiques**



**THÈSE**

Présenté en vue de l'obtention du diplôme de  
Doctorat en Mathématiques

**TITRE**

**Sur les modèles de séries chronologiques temporelles  
multivariées : exemples et application**

**Option**

**Modélisation Mathématique**

**Par**

**GUEROUAH Amin**

**DIRECTEUR DE THÈSE : ZEGHDOUDI Halim M.C.A U.B.M. ANNABA**

Devant le jury

**PRESIDENT : REMITA Mohamed Riad Prof U.B.M. ANNABA**

**EXAMINATEUR : BENCHAAABANE Abbes M.C.A U.Guelma**

**EXAMINATRICE : MERZOUGUI Mouna M.C.A U.B.M. ANNABA**

## Remerciements

Mes sincères remerciements à Dieu «الله عز و جل» le tout puissant, le miséricordieux qui m'a donné la force, la volonté et le courage afin d'élaborer ce modeste travail.

Je remercie infiniment mon directeur de thèse Docteur Zeghdoudi Halim, il restera toujours un mentor pour moi. Merci de m'avoir fait découvrir un très beau sujet et de m'y avoir guidé avec beaucoup de tact. Je l'exprime également mes sincères gratitudees pour son appui, sa sagesse, ses conseils et surtout pour le temps qu'il a consacré pour me guider à améliorer significativement la qualité de ce travail de recherche.

Je remercie également le président de jury Professeur Remita Mohamed Riad, et les membres de jury : Docteur Benchaabane Abbas et Docteur Merzougui Mouna, d'avoir accepté d'évaluer ma thèse.

Un très grand merci à Khochmane Houssein Eddine pour ses efforts et pour sa grande patience et sa disponibilité de m'avoir fait profiter de ses nombreuses compétences.

Enfin je remercie tous ceux qui m'ont aidé de près ou de loin.

## Résumé

Cette thèse étudie les propriétés empiriques des prix du pétrole et les volatilités des rendements des marchés boursiers en utilisant plusieurs modèles GARCH univariés et multivariés et des données mensuelles des États-Unis.

L'étude porte sur la période d'août 1987 à octobre 2016, soit un total de 351 observations. L'objectif de cette thèse est d'examiner la relation entre les marchés boursiers et pétroliers. En outre, nous évaluons la performance de chaque modèle avec plusieurs tests de performance de prévision en utilisant les modèles GARCH univarié (1,1) et les modèles BEKK GARCH (1,1), DCC GARCH (1,1) bivariés.

## **Abstract**

This paper investigates the empirical properties of oil price and Stock market return volatilities using a range of univariate and multivariate GARCH models and monthly data from the U.S.

The study relates the period August 1987 to October 2016, a total of 351 observations given. The aim of this paper is to examine the relationship between stock and oil markets. In addition, we evaluate the performance of each model with a range of diagnostic and forecast performance tests using univariate GARCH(1,1) and bivariate BEKK GARCH(1,1) , DCC GARCH(1,1) models.

## ملخص

في هذه الأطروحة نقوم بدراسة الخصائص التجريبية لأسعار النفط وتقلبات عوائد سوق الأسهم باستخدام مجموعة من نماذج GARCH أحادية ومتعددة المتغيرات بالإضافة الى بيانات شهرية من الولايات المتحدة.

وتتعلق الدراسة بالفترة الممتدة من أوت 1987 إلى أكتوبر 2016 ، بمجموع معطيات قُدرت 351 سعر. والهدف من هذا العمل هو دراسة العلاقة بين أسواق الأسهم وأسواق النفط بالإضافة إلى ذلك، نقوم بتقييم أداء كل نموذج باستخدام مجموعة من اختبارات الأداء والتنبؤ باستخدام نماذج  $GARCH(1,1)$  أحادي المتغير و نماذج  $BEKK-GARCH(1,1)$ ،  $DCC-GARCH(1,1)$  ثنائي المتغير.

# Table des matières

<b>1 Les Modèles ARCH, GARCH Univariés</b>	<b>8</b>
1.1 Notions de base . . . . .	8
1.1.1 Notions de stationnarité . . . . .	8
1.1.2 Ergodicité . . . . .	9
1.1.3 Fonctions d'autocovariance et d'autocorrélation . . . . .	9
1.1.4 Le bruit blanc . . . . .	10
1.1.5 La volatilité . . . . .	10
1.1.6 Séries financières . . . . .	11
1.1.7 Hétéroscédasticité . . . . .	11
1.2 Les modèles ARCH . . . . .	11
1.2.1 Les modèles ARCH(1) . . . . .	12
1.2.2 Les modèles ARCH(q) . . . . .	13
1.3 Les modèles ARCH généralisés (GARCH) . . . . .	14
1.3.1 GARCH(p,q) (Sens faible) . . . . .	15
1.3.2 Processus GARCH(p,q) fort . . . . .	16
1.4 Etude de la stationnarité . . . . .	17
1.4.1 Etude de la stationnarité du modèle GARCH(1,1) . . . . .	17
1.4.2 Etude de la stationnarité du variance conditionnelle . . . . .	23
1.4.3 Etude de la stationnarité du modèle GARCH(p,q) . . . . .	24
1.5 Tests d'hétéroscédasticité : . . . . .	36

1.5.1	Test ARCH :	36
<b>2</b>	<b>Estimation des modèles GARCH</b>	<b>38</b>
2.1	Estimation des paramètres du modèle ARCH par la méthode de MV . . .	38
2.2	Estimation des paramètres du modèle GARCH par la méthode de MV . . .	39
2.3	Estimation des paramètres du modèle GARCH par la méthode de QMV . . .	40
2.4	Estimation par régression artificielle du processus GARCH(1,1) . . . . .	46
<b>3</b>	<b>Les Modèles GARCH multivariés</b>	<b>56</b>
3.1	La notion de causalité . . . . .	56
3.1.1	Causalité au sens de Granger . . . . .	57
3.2	La notion de cointégration . . . . .	58
3.2.1	Principe de la cointégration . . . . .	58
3.2.2	Test de cointégration . . . . .	59
3.3	GARCH(p,q) multivarié . . . . .	60
3.3.1	Modèles GARCH(p,q) Vectoriels . . . . .	61
3.3.2	Modèle BEKK-GARCH . . . . .	65
3.3.3	Modèle CCC-GARCH . . . . .	66
3.3.4	Modèle DCC-GARCH . . . . .	68
3.4	La stationnarité du covariance. . . . .	72
<b>4</b>	<b>Application</b>	<b>74</b>
4.1	Résultats . . . . .	74
4.2	Les données et leurs propriétés statistiques . . . . .	74
4.3	l'étude de la stationnarité des séries. . . . .	76
4.4	Estimation de modèle GARCH(1,1) Univarié . . . . .	81
4.5	Estimation de modèle GARCH(1,1) Bivarié . . . . .	83
4.5.1	BEKK-GARCH(1,1) Bivarié . . . . .	83
4.5.2	DCC-GARCH(1,1)Bivarié . . . . .	86

# Introduction

Jusqu'au début des années 1980, l'économétrie s'est développée à un rythme relativement lent. Elle avait beaucoup de mal à se libérer du paradigme statistique classique. Mais avec la poussée fulgurante de l'informatique, l'économétrie a connu un essor fort appréciable ces vingt dernières années. Que l'on pense simplement à la multiplication effrénée des modèles économétriques non linéaires, des modèles de volatilité et des nouvelles techniques d'estimation comme le GMM ou la méthode des moments simulés, pour ne nommer que quelques nouveaux champs de l'économétrie contemporaine.

Mais ce qui est encore plus saisissant, c'est l'avancée au pas de charge de l'économétrie dans le domaine de la théorie financière. En effet, la théorie des produits dérivés, qui prend sa source au début des années 1970, fait de plus en plus appel aux modèles économétriques de volatilité, tels les modèles GARCH, et à la méthode du GMM pour estimer les paramètres des équations différentielles stochastiques qui servent à la détermination des prix des options, entres autres. L'économétrie a également permis au modèle du CAPM, bien connu en théorie financière, de s'affranchir de son cadre statique. On peut maintenant parler de bêtas variables dans le temps et la transposition de l'approche GARCH au CAPM a permis de le situer dans un cadre multivarié.

D'autre part, la modélisation de la volatilité est un enjeu important de la recherche sur les marchés financiers et de l'énergie. Cependant, il n'y a pas de réponse pour le choix des meilleurs modèles et les mesures de la volatilité des prix sont dues à la complexité de la remise des prix de l'énergie où ce problème a motivé plusieurs chercheurs après la crise financière de 2008. Les relations entre les marchés boursiers et pétroliers ont été attiré une attention particulière de la part des praticiens et des universitaires. Une relation forte entre eux aurait des implications significatives pour les décisions politiques et économiques car les chocs négatifs affectant un marché peuvent être transmis rapidement à l'autre par des effets contagieux.

Les preuves empiriques sur la relation entre les prix boursiers et pétroliers ont été

documentées par de nombreuses études. Plusieurs travaux traitent l'impact du prix du pétrole sur les indices boursiers. Cependant, un nombre réduit de recherches aborde l'effet du marché de pétrole sur rendement des actifs et l'effet de sa volatilité sur les prix des indices.

Dans [33], Jones, Charles M. and Kaul, Gautam en 1996, ont été utilisé des données trimestrielles pour vérifier si la réaction des marchés boursiers internationaux aux variations du prix du pétrole peut être justifiée par le changement des rendements attendus dans les pays tels que les USA, le Japon, et le Canada. Ils trouvent une relation stable et négative entre l'évolution du prix du pétrole et les rendements des indices boursiers. Cette relation est plus pertinente dans le cas du Japon que celui du Canada. Egalement, ces auteurs montrent que les prix actuels et les prix du pétrole décalés affectent négativement le rendement boursier. Jones et Kaul (1996) utilisent des modèles de régression simples pour étudier l'impact des prix du pétrole sur les rendements boursiers aux États-Unis, au Canada, au Japon et au Royaume-Uni et constatent aussi que les prix du pétrole ont un effet négatif sur tous les pays. Dans [18], El-Sharif I, Brown D, Burton B, Nixon B, Russell A.(2005) utilisent des données quotidiennes et mensuelles pour examiner la relation entre les prix du pétrole et les rendements des stocks pour les États-Unis et le Royaume-Uni. Plus récemment, Arouri, M. E. H., & Nguyen, D. K [2]; Arouri, M. E. H., & Rault, C [3]; aussi Ji et Fan [30] ont étudié une approche  $VAR - GARCH$  récemment développée pour analyser les débordements de la volatilité entre les marchés pétrolier et boursier en Europe.

Dans [29], Huang, R.D, Masulis, R. W, Stoll, H.R (1996), ont examiné la relation existante entre les rendements du cours de pétrole et le rendement journalier de l'indice S&P 500 en utilisant le modèle VAR. Ils trouvent qu'il y a une relation entre le rendement du pétrole et la valeur de marché de certaines compagnies pétrolières; en revanche, il n'y a pas une interaction significative entre le prix du pétrole et l'indice composite S&P\_500.

Sadorsky, P (1999) [43], est le premier à s'intéresser à la question de l'asymétrie dans la relation entre le prix du pétrole et les rendements boursiers. Il montre que volatilité

des cours de pétrole, jouent un rôle plus important que celui joué par le taux d'intérêt dans l'explication du comportement des indices boursiers des États-Unis. En effet, la fonction de réponse impulsionnelle au choc du VAR estimé, montre que les chocs positifs matérialisés par une hausse du prix de pétrole, avaient un impact plus important sur l'indice, et qu'un choc négatif illustré par une baisse du prix de pétrole, n'entraîne pas le même effet sur la variation de l'indice S&P\_500.

PARK, J. W. and Ratti, R. A (2008) [40], analysent les impacts respectifs des prix du pétrole et de leur volatilité sur les rendements des actifs boursiers aux États-Unis et dans 13 pays industrialisés d'Europe. Utilisant la même méthodologie que [43], ils montrent que les chocs sur les prix du pétrole ont un effet statistiquement significatif sur le rendement des actifs boursiers, de manière contemporaine au cours du même mois ou avec un décalage d'un mois. De plus, les effets observés varient selon les pays. Pour la plupart des pays de l'UE et non pour les USA, l'accroissement de la volatilité réduit le rendement des actifs de manière contemporaine ou avec un décalage d'un mois.

Les séries financières sont caractérisées par une volatilité, qui varie au cours du temps, par un kurtosis élevé, par une asymétrie et une persistance des chocs sur la volatilité. La variance conditionnelle, qui constitue une mesure de la volatilité d'une série, n'est plus constante et la volatilité dépend de ces valeurs passées.

Pour prendre en compte ce phénomène, Engle (1982) [20] et Bollerselv (1986) [8] ont développé les modèles ARCH (AutoRégressifs Conditionnellement Hétéroscédastiques) et GARCH (AutoRégressifs Conditionnellement Hétéroscédastiques Généralisées). Ces modèles constituent une extension du modèle ARMA, introduit par Box et Jenkins en 1971.

Selon Bera, A. K, et Higgins, M. L (1993) [5], cette modélisation ARCH/GARCH et ses extensions correspondent à une représentation spécifique du non linéarité, qui permet une modélisation simple de l'incertitude.

Dans l'économétrie de séries temporelles univariées, les modèles GARCH semblent être les plus adéquats pour prévoir les séries financières. Le choix de ces modèles est dicté par

les propriétés qui caractérisent ces séries, à savoir : la stationnarité, l'autocorrélation des carrés des variations de prix, les queues de distribution épaisses, les clusters de volatilité, les queues épaisses conditionnelles, l'effet de levier, la saisonnalité, et l'asymétrie.

De plus, les modèles GARCH univariés ont été utilisés pour modéliser la volatilité des prix du pétrole depuis le début des années 1990 et sont devenus une pratique standard. Malgré l'explosion de nouveaux types de modèles GARCH, y compris le modèle GARCH multivarié de Bollerslev (1986) [8], GARCH de Baillie R.T., Bollerslev, T., and Mikkelsen, H.O. (1996) [4], GARCH non paramétriques de Bühlmann, P. and McNeil, A.J. (2002) [12], et GARCH de la composante multiplicative de Engle, R. & Sokalska, M.E (2012) [23], aussi Joshua C.C. Chan , Angelia L. Grant (2015) [34], les modèles simples de type  $GARCH(1,1)$  restent très utiles car ils convergent beaucoup plus rapidement vers un maximum local lorsqu'on utilise une estimation par la méthode de quasi-maximum de vraisemblance, tout en délivrant des performances de prévision qui ne sont pas forcément inférieures aux modèles multivariés de Yudong Wang and Chongfeng Wu (2012) [47].

L'objectif de la thèse est d'analyser et d'étudier le comportement de la dynamique du taux de change des prix du pétrole et les indices boursiers, en appliquant la méthode de VaR. Comme dans la plupart des travaux existants dans la littérature économique, on considère un modèle sous forme multiplication où le bruit blanc au sens faible, dépendant de la volatilité. La théorie classique des séries temporelles n'étant pas suffisante pour décrire les mouvements à court terme du taux de change, nous proposons dans ce travail une méthode qui permet de modéliser la volatilité à l'aide d'un processus GARCH univarié appliquées aux données sur des prix du pétrole et les indices boursiers. De plus, on cherche d'examiner la performance des modèles GARCH multivariés. La contribution de notre recherche puise du fait que nous appliquons les modèles GARCH multivariés dans la gestion active du portefeuille sur une période 351mois et ce avec des données du marchés boursiers et pétroliers utilisant les logiciels Eviews et RATES.

La thèse comporte quatre chapitres, une introduction et une conclusion. Ce travail est composé de deux parties, une partie théorique contient trois chapitres et une partie

pratique qui contient un seul chapitre.

Ainsi, le premier chapitre contient quelques définitions et notions de base sur les séries chronologiques ainsi un panorama des modèles clés de la classe ARCH /GARCH univarié.

Le chapitre 2 se focalise sur les différentes méthodes d'estimations des paramètres des modèles utilisés à savoir :

- la méthode de maximum de vraisemblance ;
- la méthode de quasi maximum de vraisemblance ;
- la méthode de la régression artificielle.

Ainsi que les propriétés asymptotiques de l'estimateur obtenu par ces méthodes.

Le chapitre 3 présente une synthèse sur le modèle GARCH multivarié et leur stationnarité.

Enfin, le dernier chapitre, qui est l'essentiel de cette thèse, contient une application portée sur les corrélations dynamiques entre les prix du pétrole et les marchés boursiers (d'Amérique (SP 500), français (CAC 40) et Japon (NIKKEI 225)) en utilisant les modèles multivariés BEKK et DCC GARCH sur la période du 1er Août 1987 au 30 octobre 2016, soit 351 observations.

# Chapitre 1

## Les Modèles ARCH, GARCH

### Univariés

Dans cette partie, on distingue les différents types des modèles de l'hétéroscédasticité conditionnelle.

#### 1.1 Notions de base

##### 1.1.1 Notions de stationnarité

La stationnarité implique que l'état d'une série temporelle ne dépend pas du temps. En particulier, la stabilité d'une série  $(X_t, t \in \mathbb{Z})$  est vérifiée si elle ne comporte pas de tendance saisonnière, ni de tendance à la hausse ou à la baisse. Formellement, on envisage deux types de stationnarité (stationnarité stricte et stationnarité au second ordre).

**Définition 1 (Stationnarité stricte)** *Le processus  $(X_t)$  est dit strictement stationnaire si  $\forall t_1 < t_2 < \dots < t_n, t_i \in \mathbb{Z}, i = 1, \dots, n$  et quelque soit  $h \in \mathbb{Z}$ , les vecteurs  $(X_{t_1}, \dots, X_{t_n})^T$  et  $(X_{t_1+h}, \dots, X_{t_n+h})^T$  ont la même loi de probabilité.*

**Définition 2 (Stationnarité au second ordre)** *Le processus  $(X_t, t \in \mathbb{Z})$  est stationnaire au second ordre (stationnaire au sens faible) si les conditions suivantes sont*

vérifiées :

$$(i). \forall t \in \mathbb{Z}, E(X_t^2) < \infty,$$

$$(ii). \forall t \in \mathbb{Z}, E(X_t) = m, \text{ indépendant de } t,$$

$$(iii). \forall t, h \in \mathbb{Z}, \text{cov}(X_t, X_{t+h}) = E[(X_{t+h} - m)(X_t - m)] = \gamma(h), \text{ indépendant de } t.$$

La fonction  $\gamma(\cdot)$  (resp.  $\rho(\cdot) := \gamma(\cdot)/\gamma(0)$ ) est appelée fonction d'autocovariance (resp. d'autocorrélation) de  $(X_t)$ .

-La condition (i) signifie l'existence des moments d'ordre deux.

-La condition (ii) implique que les variables aléatoires  $(X_t)$  ayant la même espérance quelle que soit la date  $t$ , i.e  $E(X_t)$  ne dépend pas du temps.

-La condition (iii) signifie que les moments sont indépendants de la date considérée et ne dépendent que uniquement de l'ordre des retards, c'est-à-dire la fonction d'autocovariance de  $(X_t)$  doit être indépendante du temps. Donc on peut conclure qu'un processus faiblement stationnaire si ses moments ne dépend pas du temps. Par conséquent, au cours du temps si la variance de  $(X_t)$  est constante alors le processus  $(X_t)$  est stationnaire sans passer aux test d'hypothèse sur la stationnarité.

### 1.1.2 Ergodicité

Soit  $\{X_t\}_{t=0}^n$  un processus stationnaire au second ordre, le processus  $(X_t)$  est ergodique si

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{h=1}^n \gamma(h) = 0,$$

avec  $\gamma(h)$  représente la fonction d'autocovariance du processus  $X_t$ .

### 1.1.3 Fonctions d'autocovariance et d'autocorrélation

**Définition 3** Soit  $(X_t)$  une série stationnaire,

- La fonction d'autocovariance de  $(X_t)$  est définie par :

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

- La fonction d'autocorrélation de  $(X_t)$  est définie par :

$$\rho_X(h) = \frac{\gamma_X(h)}{\gamma_X(0)} = \text{cor}(X_t, X_{t+h}), \forall t, h \in \mathbb{Z}.$$

### 1.1.4 Le bruit blanc

#### Bruit blanc faible

**Définition 4** Soit  $(\varepsilon_t)_{t \in \mathbb{Z}}$  un processus stochastique, le processus  $(\varepsilon_t)$  est appelé bruit blanc faible si les trois propriétés suivantes soient vérifiées, pour une constante positive  $\sigma^2$  :

- (i).  $E(\varepsilon_t) = 0, \forall t \in \mathbb{Z},$
  - (ii).  $E(\varepsilon_t^2) = \sigma^2, \forall t \in \mathbb{Z},$
  - (iii).  $\text{cov}(\varepsilon_t, \varepsilon_{t+h}) = E(\varepsilon_t \varepsilon_{t+h}) = 0, \forall t, h \in \mathbb{Z}, h \neq 0.$
- on note  $(\varepsilon_t) \sim WN(0, \sigma^2).$

### 1.1.5 La volatilité

**Définition 5** l'instabilité du cours d'un actif financier est mesurée par la volatilité, cette dernière mesure l'amplitude d'un produit dérivé et des variations d'une action. Il s'agit d'un paramètre de quantification du risque de rendement et de prix. Les séries monétaires et financières sont caractérisées par le clustering de volatilité, à savoir les périodes de forte volatilité alternent avec les périodes de faible volatilité. Ce phénomène, que nous appelons aussi l'hétéroscédasticité conditionnelle, est particulièrement fréquent dans les données boursières, les taux de changes ou d'autres prix déterminés sur les marchés financiers.

### 1.1.6 Séries financières

La modélisation des séries financières est très difficile. Cette difficulté due pour les raisons suivantes :

- L'existence de régularités statistique qui est la même dans les différentes séries financières.
- La disponibilité d'échantillons de très grandes tailles.
- L'importance de la fréquence d'observation.
- La variété des séries financières.

Les propriétés concernent les régularités empiriques des séries financières qui sont mises en 1963 par *Mandelbrot*, vérifiées et complétées par nombreux auteurs, ses propriétés qu'on va présenter ci-dessous valent surtout pour les séries quotidiennes de prix d'action.

Soit  $S_t$  le cours d'un actif à la date  $t$  et soit  $\epsilon_t = \log(S_t/S_{t-1})$  présente le rendement. Ainsi que  $(\epsilon_t)$  décrit les variations relatives des prix  $r_t = \frac{S_t - S_{t-1}}{S_{t-1}}$ . Si on pose  $\epsilon_t = \log(1+r_t)$ , dans ce cas, ces deux séries présentent sur celle des prix l'avantage d'être sans unité, cela ne permet de comparer facilement entre différents titres.

### 1.1.7 Hétéroscédasticité

**Définition 6** *En statistique, lorsque les variances des variables étudiées sont différentes, dans ce cas on parle sur l'hétéroscédasticité. Par contre, si la variance de l'erreur des variables est constante, on parle ici sur l'homoscédasticité i.e  $Var(\epsilon_i) = \sigma^2 \forall i$ , nous avons désormais  $Var(\epsilon_i) = \sigma_i^2$ , où  $\sigma_i^2 \neq \sigma_j^2$  si  $i \neq j$ .*

## 1.2 Les modèles ARCH

Une classe de modèles autorégressifs conditionnellement hétéroscédastiques (*ARCH*) a été créée par Engle nous permet de connaître le comportement temporel de la volatilité.

Le modèle est formé de deux équations. La première met en relation le rendement et certaines variables qui l'expliquent et la seconde modélise la variance conditionnelle des résidus.

Le principe proposé par Engle consiste à introduire une dynamique dans la détermination de la volatilité en supposant que la variance est conditionnelle aux informations dont nous disposons. Il avance une spécification  $ARCH(q)$  ont le carré des innovations, c'est-à-dire la variance du terme d'erreur au temps  $t$ , dépend de l'importance des termes d'erreur au carré des  $q$  périodes passées. Le modèle  $ARCH(q)$  permet de générer des épisodes de volatilité importante suivis d'épisodes de volatilité plus faibles.

### 1.2.1 Les modèles ARCH(1)

**Définition 7** *Le processus  $(X_t)_{t \geq 0}$  satisfait une représentation d'un processus autorégressif conditionnellement hétéroscédastique d'ordre 1, noté ARCH(1), si et seulement si*

$$X_t = \varepsilon_t \sqrt{h_t}, \quad (1.1)$$

où

$$h_t = \alpha_0 + \alpha_1 X_{t-1}^2,$$

et  $(\varepsilon_t)$  désigne un bruit blanc faible.

De façon générale,  $\varepsilon_t$  désigne un ensemble de variables aléatoires indépendantes, identiquement distribuées, centrées, réduites. La composante  $h_t$  désigne une variable qui est déterministe et positive conditionnellement à l'ensemble d'informations des valeurs passées de  $X_t$ , *i.e.* à  $\underline{X}_{t-1} = \{X_{t-1}, X_{t-2}, \dots, X_{t-j}, \dots\}$ . Dans ce système, le processus  $(X_t)$  est caractérisé par des autocorrélations nulles et une variance conditionnelle variée dans le temps en fonction de l'ampleur de l'innovation passée.

On peut établir des résultats intéressants en considérant le processus autorégressif sur

$X_t^2$ . Pour simplifier, on se limite au cas du  $ARCH(1)$ . Dans ces conditions :

$$h_t = \alpha_0 + \alpha_1 X_{t-1}^2 \Leftrightarrow X_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + (X_t^2 - h_t),$$

soit encore :

$$X_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \nu_t,$$

où  $\nu_t = X_t^2 - h_t$  vérifiant  $E(\nu_t | X_{t-1}) = 0$  est un processus d'innovation pour  $X_t^2$ . Ainsi, cette écriture précédente correspond celle d'un  $AR(1)$  sur le carré  $X_t^2$ . On sait que ce processus  $X_t^2$  est stationnaire au second ordre si et seulement si  $\alpha_1 < 1$ . On peut déduire de ses différentes écritures, un certain nombre de propriétés qui pourront être étendues au cas des processus  $ARCH(q)$ .

### 1.2.2 Les modèles $ARCH(q)$

**Définition 8** *On considère un modèle linéaire autorégressif exprimé sous la forme suivante :*

$$r_t = \mu_t + X_t,$$

où

$$\mu_t = E(r_t | \psi_{t-1}),$$

*On dit que le processus  $(X_t)$  suit un processus  $ARCH(q)$  s'il est défini par une équation de la forme*

$$X_t = \varepsilon_t \sqrt{h_t},$$

*l'équation de la variance conditionnelle d'un processus  $ARCH$  s'écrit :*

$$h_t = \alpha_0 + \sum_{i=1}^q \alpha_i X_{t-i}^2, \quad (1.2)$$

où :

$r_t$  : rendements d'un actif au temps  $t$ ,

$\mu_t$  : l'espérance conditionnelle de  $r_t$ ,

$X_t$  : erreurs conditionnelles d'actif au temps  $t$ ,

$\psi_t$  : ensemble d'informations au temps  $t$ ,

$h_t$  : la variance conditionnelle au temps  $t$  (la volatilité au carré),

$\varepsilon_t$  : bruit blanc faible,

$\alpha_0$  : paramètre qui représente la variance moyenne à long terme,

$\alpha_i$  : paramètre qui mesure la sensibilité à la volatilité conditionnelle,

$q$  : l'ordre du processus ARCH

Afin de s'assurer d'obtenir une variance conditionnelle positive, il faut respecter les conditions suivantes :  $\alpha_0 > 0$ ,  $\alpha_i \geq 0$  pour tout  $i = 1, 2, 3, \dots, q$ .

### 1.3 Les modèles ARCH généralisés (GARCH)

Les modèles ARCH (autorégressifs conditionnellement hétéroscédastiques) ont été introduits par Engle (1982) [20] et leur extension GARCH (ARCH généralisés) est due à Bollerslev (1986) [8]. Leur caractérisation repose essentiellement sur le concept de variance conditionnelle. Dans ces modèles, celle-ci s'écrit comme une fonction affine des valeurs passées du carré de la série. Cette spécification particulière se révèle très fructueuse car elle permet une étude complète des propriétés des solutions tout en étant assez générale. Les modèles GARCH sont en effet susceptibles de capter les propriétés caractéristiques de certaines séries.

### 1.3.1 GARCH(p,q) (Sens faible)

**Définition 9** On dit que  $\{X_t\}_{t \in \mathbb{N}}$  est un processus GARCH(p, q) si ses deux premiers moments conditionnels existent et vérifient :

(i).  $E(X_t | X_s, s < t) = 0, t \in \mathbb{Z}$ ;

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

$$h_t = V(X_t | X_s, s < t) = \alpha_0 + \sum_{i=1}^q \alpha_i X_{t-i}^2 + \sum_{j=1}^p \beta_j h_{t-j}, t \in \mathbb{Z}, \quad (1.3)$$

où  $\varepsilon_t$  est un bruit blanc faible.

L'équation (1.3) peut-être écrit de manière symbolique sous la forme plus compacte

$$h_t = \alpha_0 + \alpha(L)X_t^2 + \beta(L)h_t, t \in \mathbb{Z}, \quad (1.4)$$

où  $L$  est l'opérateur retard ( $L^i X_t^2 = X_{t-i}^2$  et  $L^i h_t = h_{t-i}$  pour tout entier  $i$ ),  $\alpha$  et  $\beta$  sont les polynômes de degrés  $q$  et  $p$  :

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

**Définition 10** L'innovation du processus  $X_t^2$  est par définition la variable  $\nu_t = X_t^2 - h_t$ . En remplaçant, dans l'équation (1.3), les variables  $h_{t-j}$  par  $X_{t-j}^2 - \nu_{t-j}$ ; on obtient la représentation

$$X_t^2 = \alpha_0 + \sum_{i=1}^r (\alpha_i + \beta_i) X_{t-i}^2 + \nu_t - \sum_{j=1}^p \beta_j \nu_{t-j}, t \in \mathbb{Z},$$

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

- On retrouve ainsi dans cette équation la structure linéaire des modèles ARMA,

- Sous l'hypothèse suivante :

$X_t^2$  est stationnaire, si  $(X_t)$  est un GARCH(p, q), alors  $(X_t^2)$  est un processus ARMA(r, p).

### 1.3.2 Processus GARCH(p,q) fort

**Définition 11** Dans le cas d'un GARCH fort, on dit que  $\{X_t\}_{t \in \mathbb{N}}$  suit un processus GARCH(p, q) si :

$$\begin{cases} X_t = \varepsilon_t \sqrt{h_t} \\ h_t = \alpha_0 + \sum_{i=1}^q \alpha_i X_{t-i}^2 + \sum_{j=1}^p \beta_j h_{t-j}, t \in \mathbb{Z}' \end{cases} \quad (1.5)$$

où  $\alpha_0 > 0$ ,  $\alpha_i \geq 0$ ,  $i = 1, \dots, q$  et  $\beta_j \geq 0$ ,  $j = 1, \dots, p$ ,  $p \geq 0$  et  $\varepsilon_t$  est un bruit blanc fort gaussien  $N(0, 1)$ , ou on peut écrire  $h_t$  sous la forme suivante :

$$h_t = \alpha_0 + \sum_{i=1}^q \alpha_i h_{t-i} \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j h_{t-j}, \quad (1.6)$$

ou bien

$$h_t = \alpha_0 + \sum_{i=1}^r a_i(\varepsilon_{t-i}) h_{t-i},$$

où  $a_i(z) = \alpha_i z^2 + \beta_i$ ,  $i = 1, \dots, r$ . Cette représentation montre que dans le cas d'un GARCH fort, le processus de volatilité vérifie une équation autorégressive, mais avec coefficients aléatoires.

#### Propriétés

-Si le processus GARCH(p, q) est stationnaire au second ordre, alors nécessairement

$$\sum_{i=1}^p \alpha_i + \sum_{j=1}^q \beta_j < 1,$$

et dans ce cas, la variance de  $X_t$  est donnée par

$$V(X_t) = \frac{\alpha_0}{1 - \sum_{i=1}^q \alpha_i + \sum_{j=1}^p \beta_j}.$$

- Les conditions  $\alpha_0 > 0$ ,  $\alpha_i \geq 0$  et  $\beta_j \geq 0$  suffisantes pour garantir la positivité de  $h_t$ .
- Bollerslev (1986) a montré que, dans le cas d'un GARCH(p, q), la variance du processus reste finie si la somme des paramètres  $\alpha_1 + \dots + \alpha_q + \beta_1 + \dots + \beta_p$  est plus petite

que 1.

- Un  $GARCH(0, q)$  est un  $ARCH(q)$ .

## 1.4 Etude de la stationnarité

### 1.4.1 Etude de la stationnarité du modèle $GARCH(1,1)$

Nous allons chercher sous quelles conditions il existe des processus stationnaires (au sens strict et au second-ordre) vérifiant les définitions 9 et/ou 11. On s'intéresse plus particulièrement aux solutions non anticipatives du modèle (1.5), c'est-à-dire aux processus  $(X_t)$  tel que  $X_t$  soit une fonction mesurable des variables  $\varepsilon_{t-s}$ ,  $s \geq 0$ .

Nous examinons d'abord le cas du modèle  $GARCH(1, 1)$  qui peut se traiter avec des techniques élémentaires. On notera, pour  $x > 0$ ,  $\log^+ x = \max(\log x, 0)$ .

#### Modèle $GARCH(1,1)$

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

$$\begin{cases} X_t = \varepsilon_t \sqrt{h_t} \\ h_t = \alpha_0 + \alpha X_{t-1}^2 + \beta h_{t-1} \end{cases} \quad (1.7)$$

avec  $\alpha_0 > 0$ ,  $\alpha \geq 0$ , et  $\beta \geq 0$ , on pose  $a(z) = \alpha z^2 + \beta$ .

#### Stationnarité stricte

##### **Théorème 12 Stationnarité stricte du modèle $GARCH(1,1)$ fort**

*Si*

$$-\infty \leq \gamma := E (\log \{ \alpha \varepsilon_t^2 + \beta \}) < 0, \quad (1.8)$$

*la série*

$$s_t = \left\{ 1 + \sum_{i=1}^{\infty} a(\varepsilon_{t-1}) \dots a(\varepsilon_{t-i}) \right\} \alpha_0, \quad (1.9)$$

converge presque sûrement (*p.s.*) et le processus  $(X_t)$  défini par  $X_t = \varepsilon_t \sqrt{s_t}$  est l'unique solution strictement stationnaire du modèle (1.7).

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

**Remarque 1**

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

2) Dans le cas où  $\alpha_0 = 0$  et  $\gamma < 0$ , il est clair d'après (1.9) que la seule solution strictement stationnaire du modèle est  $X_t = 0$ . Il n'est donc naturel d'imposer  $\alpha_0 > 0$  dans la pratique.

3) On voit que la condition (1.8) dépend de la loi du processus  $(\varepsilon_t)$  et qu'elle n'est pas symétrique en  $\alpha$  et  $\beta$ .

4) La condition (1.8) implique  $\beta < 1$ . Inversement, si  $\alpha + \beta < 1$ , (1.8) est vérifiée, car par application de l'inégalité de Jensen

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

5) Si (1.8) est satisfaite, elle l'est également pour tout couple  $(\alpha_1, \beta_1)$  tel que  $\alpha_1 \leq \alpha$  et  $\beta_1 \leq \beta$ . En particulier la stationnarité stricte du modèle *GARCH* implique celle du modèle *ARCH* obtenu en supprimant  $\beta$ .

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

$$0 \leq \alpha < \exp\{-E(\log \varepsilon_t^2)\}. \quad (1.10)$$

**Preuve du théorème 12**

En utilisant de manière itérative la seconde équation du modèle (1.7), on obtient, pour  $N \geq 1$ ,

$$\begin{aligned}
h_t &= \alpha_0 + a(\varepsilon_{t-1})h_{t-1} \\
&= \alpha_0 \left[ 1 + \sum_{n=1}^N a(\varepsilon_{t-1}) \dots a(\varepsilon_{t-n}) \right] + a(\varepsilon_{t-1}) \dots a(\varepsilon_{t-N-1})h_{t-N-1} \\
&: = s_t(N) + a(\varepsilon_{t-1}) \dots a(\varepsilon_{t-N-1})h_{t-N-1}.
\end{aligned} \tag{1.11}$$

Le processus limite  $s_t = \lim_{N \rightarrow \infty} s_t(N)$  existe dans  $\overline{\mathbb{R}}^+ = [0, +\infty]$ , puisque les termes de la somme sont positifs. De plus, en faisant tendre  $N$  vers l'infini dans la relation  $s_t(N) = \alpha_0 + a(\varepsilon_{t-1})s_{t-1}(N-1)$ , on obtient :

$$s_t = \alpha_0 + a(\varepsilon_{t-1})s_{t-1}.$$

Nous allons montrer que  $s_t$  est presque sûrement finie si et seulement si  $\gamma < 0$ .

Supposons  $\gamma < 0$ . On utilise la règle de Cauchy pour les séries à termes positifs.

On a

$$[a(\varepsilon_{t-1}) \dots a(\varepsilon_{t-n})]^{1/n} = \exp \left[ \frac{1}{n} \sum_{i=1}^n \log \{a(\varepsilon_{t-i})\} \right] \rightarrow e^\gamma \text{ p.s.}, \tag{1.12}$$

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

La série définie dans (1.9) converge alors 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  $(X_t)$  défini par :

$$X_t = \varepsilon_t \sqrt{s_t} = \left\{ \alpha_0 + \sum_{i=1}^{\infty} a(\varepsilon_{t-1}) \dots a(\varepsilon_{t-i}) \alpha_0 \right\}^{1/2} \varepsilon_t, \tag{1.13}$$

est strictement stationnaire (voir [14]). Il est non anticipatif comme fonction mesurable des variables  $\varepsilon_{t-i}, i \geq 0$ . De plus  $(X_t)$  vérifie le modèle (1.7).

Nous montrons maintenant l'unicité. Soit  $\tilde{X}_t = \varepsilon_t \sqrt{\tilde{h}_t}$  une solution strictement sta-

tionnaire. D'après (1.11) on a

$$h_t = s_t(N) + a(\varepsilon_{t-1}) \dots a(\varepsilon_{t-N-1}) h_{t-N-1}.$$

Par suite

$$h_t - s_t = \{s_t(N) - s_t\} + a(\varepsilon_{t-1}) \dots a(\varepsilon_{t-N-1}) h_{t-N-1}.$$

Le terme entre accolades à droite de l'égalité tend vers 0 p.s. quand  $N \rightarrow \infty$ . Par ailleurs, puisque la série définissant  $s_t$  converge p.s., on a  $a(\varepsilon_{t-1}) \dots a(\varepsilon_{t-n}) \rightarrow 0$  avec probabilité 1 quand  $n \rightarrow \infty$ . De plus la loi de  $h_{t-N-1}$  est indépendante de  $N$  par stationnarité. Par suite  $a(\varepsilon_{t-1}) \dots a(\varepsilon_{t-N-1}) h_{t-N-1} \rightarrow 0$  en probabilité lorsque  $N \rightarrow \infty$ . On a montré que  $h_t - s_t \rightarrow 0$  en probabilité quand  $N \rightarrow \infty$ . Ce terme étant indépendant de  $N$ , on a nécessairement  $s_t = h_t$  pour tout  $t$ , p.s.

Si  $\gamma > 0$ , d'après (1.12) et la règle de Cauchy,  $\sum_{n=1}^N a(\varepsilon_{t-1}) \dots a(\varepsilon_{t-n}) \rightarrow +\infty$ , p.s. lorsque  $N \rightarrow \infty$ . Donc si  $\alpha_0 > 0$ ,  $s_t = +\infty$ , p.s. D'après (1.11), il est clair que  $h_t = +\infty$ , p.s. Par suite, il n'existe pas de solution finie p.s. de (1.7).

Dans le cas  $\gamma = 0$ , nous procéderons par l'absurde. Supposons qu'il existe une solution strictement stationnaire  $(X_t, h_t)$  de (1.7). Nous avons pour  $n > 0$ ,

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

d'où on déduit que le terme général  $a(\varepsilon_{t-1}) \dots a(\varepsilon_{t-n}) \alpha_0$  converge vers zéro, p.s., quand  $n \rightarrow \infty$ , ou, de manière équivalente, que

$$\sum_{i=1}^n \log a(\varepsilon_i) + \log \alpha_0 \rightarrow -\infty \text{ p.s. quand } n \rightarrow \infty. \quad (1.14)$$

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

## Stationnarité au second ordre

### **Théorème 13 Stationnarité au second ordre du GARCH(1,1)**

*Supposons  $\alpha_0 > 0$ .*

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

*Si  $\alpha + \beta < 1$ , le processus  $(X_t)$  défini par (1.13), est stationnaire au second ordre.*

*Plus précisément,  $(X_t)$  est un bruit blanc. De plus, il n'existe pas d'autre solution stationnaire au second ordre et non anticipative.*

### **Preuve**

Si  $X_t$  est un processus GARCH(1,1), au sens de la définition 9, stationnaire au second-ordre et non anticipatif, on a

$$E(X_t^2) = E \{E(X_t^2 | X_s, s < t)\} = E(h_t) = \alpha_0 + (\alpha + \beta)E(X_{t-1}^2)$$

soit

$$(1 - \alpha - \beta)E(X_t^2) = \alpha_0$$

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

Inversement, supposons  $\alpha + \beta < 1$ . D'après la remarque 4 précédente, la condition de stationnarité stricte est vérifiée. Il suffit donc de montrer que la solution strictement stationnaire définie dans (1.13) admet une variance finie. La variable  $s_t$  étant une limite croissante de variables aléatoires positives, d'après le théorème de Beppo Levi, on peut intervertir espérance et somme infinie et écrire

$$\begin{aligned}
E(X_t^2) &= E(s_t) = \left[ 1 + \sum_{n=1}^{+\infty} E \{a(\varepsilon_{t-1}) \dots a(\varepsilon_{t-n})\} \right] \alpha_0 \\
&= \left[ 1 + \sum_{n=1}^{+\infty} \{Ea(\varepsilon_t)\}^n \right] \alpha_0 \\
&= \left[ 1 + \sum_{n=1}^{+\infty} (\alpha + \beta)^n \right] \alpha_0 \\
&= \frac{\alpha_0}{1 - (\alpha + \beta)}
\end{aligned}$$

Cela suffit à prouver la stationnarité au second-ordre de la solution. De plus cette solution est un bruit blanc car  $E(X_t) = E\{E(X_t|X_s, s < t)\} = 0$  et pour tout  $h > 0$ ,

$$\text{cov}(X_t, X_{t-h}) = E\{X_{t-h}E(X_t|X_s, s < t)\} = 0.$$

Soit  $\tilde{X}_t = \varepsilon_t \sqrt{\tilde{s}_t}$  une autre solution stationnaire au second ordre et non anticipative.

On a

$$|s_t - \tilde{s}_t| = a(\varepsilon_{t-1}) \dots a(\varepsilon_{t-n}) |s_{t-n-1} - \tilde{s}_{t-n-1}|$$

et par suite,

$$\begin{aligned}
E|s_t - \tilde{s}_t| &= E\{a(\varepsilon_{t-1}) \dots a(\varepsilon_{t-n})\} E|s_{t-n-1} - \tilde{s}_{t-n-1}| \\
&= (\alpha + \beta)^n E|s_{t-n-1} - \tilde{s}_{t-n-1}|.
\end{aligned}$$

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

## 1.4.2 Etude de la stationnarité du variance conditionnelle

### **Théorème 14**

Soit  $(Y_t)$  est un processus stochastique défini par

$$Y_t = A_t + B_t Y_{t-1}, \quad t \in \mathbb{N},$$

On suppose que  $Y_0$  est indépendant de  $((A_t, B_t)_{t \geq 0})$  qui sont iid. Supposons que

$$E \ln^+ |A| < \infty \text{ et } -\infty \leq E \ln |B| < 0.$$

Alors

a–  $Y_t \xrightarrow{\mathcal{D}} Y$  pour chaque variable  $Y$ , telle qu'elle vérifie l'identité en loi

$$Y = A + BY,$$

où  $Y$  et  $(A, B)$  sont indépendantes.

b– L'équation précédente a une unique solution au sens de distribution, qui est donnée par

$$Y_t \xrightarrow{\mathcal{D}} \sum_{m=1}^{\infty} A_m \prod_{j=1}^{m-1} B_j. \quad (1.15)$$

Le côté droit de l'équation (1,15) converge absolument avec une probabilité 1.

c– Si on choisit  $Y_0 \xrightarrow{\mathcal{D}} Y$  comme dans (1,15), Alors le processus  $(Y_t)_{t \geq 0}$  est strictement stationnaire.

### **Théorème 15**

Soit  $h_t$  la variance conditionnelle du processus GARCH(1,1) définie dans (1,7), et on suppose que

$$E [\ln(\alpha_1 \varepsilon_0^2 + \beta_1)] < 0$$

et que  $h_0$  est indépendant de  $(\varepsilon_t)$ .

Alors le processus  $(h_t)$  est strictement stationnaire si

$$h_0 \stackrel{\mathcal{D}}{=} \alpha_0 \sum_{m=1}^{\infty} \prod_{j=1}^{m-1} (\beta_1 + \alpha_1 \varepsilon_{j-1}^2). \quad (1.16)$$

et la série (1, 16) est convergente absolument avec une probabilité égale à 1.

### Preuve

De l'équation (1, 7) on a

$$h_t = \alpha_0 + \alpha_1 X_{t-1}^2 + \beta_1 h_{t-1},$$

ou

$$h_t = \alpha_0 + (\alpha_1 \varepsilon_{t-1}^2 + \beta_1) h_{t-1}$$

qui représente une équation de différence stochastique

$$Y_t = A_t + B_t Y_{t-1},$$

où  $Y_t = h_t$ ,  $A_t = \alpha_0$  et  $B_t = \alpha_1 \varepsilon_{t-1}^2 + \beta_1$ . A partir des hypothèses du théorème nous avons  $E \ln^+ |A| < \infty$  et  $E \ln |B| = E [\ln(\beta_1 + \alpha_1 \varepsilon_{t-1}^2)] < 0$ . Alors d'après le théorème (14), on peut conclure que  $(h_t)$  est strictement stationnaire avec une distribution marginale unique donnée par (1, 16), ceci montre le résultat du théorème.

### 1.4.3 Etude de la stationnarité du modèle GARCH(p,q)

Dans le cas général du  $GARCH(p, q)$  fort, l'écriture vectorielle suivante sera très utile.

On a

$$z_t = \underline{b}_t + A_t z_{t-1}, \quad (1.17)$$

où

$$\underline{b}_t = \underline{b}_t(\varepsilon_t) = \begin{pmatrix} \alpha_0 \varepsilon_t^2 \\ 0 \\ \vdots \\ \alpha_0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in \mathbb{R}^{p+q}, \underline{z}_t = \begin{pmatrix} X_t^2 \\ \vdots \\ X_{t-q+1}^2 \\ h_t \\ \vdots \\ h_{t-p+1} \end{pmatrix} \in \mathbb{R}^{p+q}$$

et

$$A_t = \begin{pmatrix} \alpha_1 \varepsilon_t^2 & \cdots & \alpha_q \varepsilon_t^2 & \beta_1 \varepsilon_t^2 & \cdots & \beta_p \varepsilon_t^2 \\ 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 1 & 0 & 0 & \cdots & 0 \\ \alpha_1 & \cdots & \alpha_q & \beta_1 & \cdots & \beta_p \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 1 & 0 \end{pmatrix} \quad (1.18)$$

est une matrice de dimension  $(p+q) \times (p+q)$ . Dans le cas  $ARCH(q)$ ,  $\underline{z}_t$  ne contient que  $X_t^2$  et ses  $q-1$  premières valeurs passées, et  $A_t$  se limite au bloc supérieur à gauche de la matrice ci-dessus. L'équation (1.17) constitue un modèle vectoriel autorégressif d'ordre un, avec coefficients positifs et iid. La loi de  $\underline{z}_t$  conditionnelle à son passé infini coïncide avec sa loi conditionnelle à  $z_{t-1}$  seulement, ce qui fait que  $(\underline{z}_t)$  un processus de Markov. On parle ainsi de représentation markovienne.

En itérant (1.17) on obtient

$$\underline{z}_t = \underline{b}_t + \sum_{k=1}^{\infty} A_t A_{t-1} \dots A_{t-k+1} \underline{b}_{t-k} \quad (1.19)$$

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 (1.19) a un sens, cela n'assure pas pour autant que les composantes de ce vecteur sont positives. Une condition suffisante pour que, presque sûrement,

$$\underline{b}_t + \sum_{k=1}^{\infty} A_t A_{t-1} \dots A_{t-k+1} \underline{b}_{t-k} > 0 \quad (1.20)$$

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

$$\alpha_0 > 0, \alpha_i \geq 0 (i = 1, \dots, q), \beta_j \geq 0 (j = 1, \dots, p). \quad (1.21)$$

Cette condition, très simple à utiliser, n'est cependant pas toujours nécessaire comme nous le verrons plus loin.

### Stationnarité stricte

L'outil principal pour l'étude de la stationnarité stricte est le concept d'exposant de *Lyapounov*. Soit  $A$  une matrice de dimension  $(p+q) \times (p+q)$ . 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  $(p+q) \times (p+q)$ . On a le résultat d'algèbre suivant

$$\lim_{t \rightarrow \infty} \frac{1}{t} \log \|A^t\| = \log \rho(A) \quad (1.22)$$

### Théorème 16

*Soit  $\{A_t, t \in \mathbb{Z}\}$  une suite de matrices aléatoires, strictement stationnaire et ergodique, telle que  $E \log^+ \|A^t\|$  est finie. On a*

$$\lim_{t \rightarrow \infty} \frac{1}{t} E(\log \|A_t A_{t-1} \dots A_1\|) = \gamma = \inf_{t \in \mathbb{N}^*} \frac{1}{t} E(\log \|A_t A_{t-1} \dots A_1\|) \quad (1.23)$$

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

$$\gamma = \lim_{t \rightarrow \infty} p.s. \frac{1}{t} \log \|A_t A_{t-1} \dots A_1\| \quad (1.24)$$

### Remarque 2

- 1) On a toujours  $\gamma \leq E(\log \|A_1\|)$ , avec égalité en dimension 1.
- 2) Si  $A_t = A$  pour tout  $t \in Z$ , on a  $\gamma = \log \rho(A)$  d'après (1.22).
- 3) Toutes les normes étant équivalentes sur un espace de dimension fini, il est facile de voir que  $\gamma$  est indépendant du choix de la norme.

4) L'équivalence entre les définitions de  $\gamma$  se montre en utilisant le théorème ergodique sous-additif, voir Kingmam (1973) théorème 6 [36]. La caractérisation (1.24) est particulièrement intéressante car elle permet de calculer des approximations de ce coefficient par simulation. Des intervalles de confiance asymptotiques peuvent également être obtenus, voir Goldsheid (1991) [26].

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

**Lemme 17** Soit  $\{A_t, t \in Z\}$  une suite de matrices aléatoires iid telle que  $E(\log^+ \|A_t\|)$  est finie et de plus grand exposant de Lyapounov  $\gamma$ . Alors

$$\lim_{t \rightarrow \infty} p.s. \|A_0 \dots A_{-t}\| = 0 \Rightarrow \gamma < 0 \quad (1.25)$$

Comme pour les modèles ARMA, nous nous intéressons plus particulièrement aux solutions  $(X_t)$  non anticipatives du modèle (1.5), c'est à dire telles que  $X_t$  appartient à la tribu engendrée par  $\{\varepsilon_t, \varepsilon_{t-1}, \dots\}$ .

### Théorème 18 Stationnarité stricte du modèle GARCH(p, q)

Une condition nécessaire et suffisante d'existence d'un processus GARCH(p, q) stric-

tement stationnaire, solution du modèle (1.5) est que

$$\gamma < 0$$

où  $\gamma$  est le plus grand exposant de Lyapounov de la suite  $\{A_t, t \in \mathbb{Z}\}$  définie par (1.18).

Lorsqu'elle existe, la solution strictement stationnaire est unique, non anticipative et ergodique.

### Preuve

Nous utiliserons la norme définie par  $\|A\| = \sum |a_{ij}|$ . 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  $\varepsilon_t$  étant de variance finie, tous les termes de la matrice  $A_t$  sont intégrables. On a donc

$$E(\log^+ \|A_t\|) \leq E \|A_t\| < \infty.$$

Supposons d'abord  $\gamma < 0$ . Alors, l'égalité (1.24) implique que la série

$$\tilde{z}_t = \underline{b}_t + \sum_{n=0}^{\infty} A_t A_{t-1} \dots A_{t-n} \underline{b}_{t-n-1}$$

converge presque sûrement pour tout  $t$ . On a en effet, en utilisant la multiplicativité de la norme,

$$\|\tilde{z}_t\| \leq \|\underline{b}_t\| + \sum_{n=0}^{\infty} \|A_t A_{t-1} \dots A_{t-n}\| \|\underline{b}_{t-n-1}\|, \quad (1.26)$$

et

$$\|A_t \dots A_{t-n}\|^{1/n} \|b_{t-n-1}\|^{1/n} = \exp \left[ \frac{1}{n} \log \|A_t \dots A_{t-n}\| + \frac{1}{n} \log \|b_{t-n-1}\| \right] \xrightarrow{p.s.} e^\gamma < 1.$$

Par suite, par la règle de Cauchy,  $\tilde{z}_t$  est bien défini dans  $(R^{*+})^{p+q}$ . Soit  $\tilde{z}_{q+1,t}$  la  $q+1$ ème composante de  $\tilde{z}_t$ . En posant  $X_t = \sqrt{\tilde{z}_{q+1,t}} \varepsilon_t$ , on définit une solution strictement stationnaire du modèle (1.5). D'après (1.19),  $X_t$  s'exprime comme fonction mesurable de  $\varepsilon_t, \varepsilon_{t-1}, \dots$ . La solution est donc non anticipative et ergodique puisque  $(\varepsilon_t)$  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 du modèle (1.5), ou de manière équivalente une autre solution strictement stationnaire positive  $(z_t^*)$  de (1.17). Alors, pour tout  $n \geq 0$ ,

$$\|\tilde{z}_t - z_t^*\| = \|A_t(\tilde{z}_{t-1} - z_{t-1}^*)\| = \dots \leq \|A_t \dots A_{t-n}\| \|\tilde{z}_{t-n-1} - z_{t-n-1}^*\|.$$

On a  $P(\|\tilde{z}_t - z_t^*\| \neq 0) > 0$ . Or on sait que  $\|A_t \dots A_{t-n}\| \rightarrow 0$  *p.s.* quand  $n \rightarrow \infty$  car la série intervenant dans (1.26) converge. Par suite  $P(\|\tilde{z}_{t-n-1} - z_{t-n-1}^*\| \rightarrow \infty) > 0$ , ce qui implique que  $\|\tilde{z}_{t-n-1}\| \rightarrow \infty$  ou  $\|z_{t-n-1}^*\| \rightarrow \infty$  avec une probabilité positive.

Ceci est impossible car les suites  $(\tilde{z}_t)_t$  et  $(z_t^*)_t$  sont stationnaires. On en conclut que  $\tilde{z}_t = z_t^*$  pour tout  $t$ , *p.s.*

Nous montrons finalement la partie nécessaire du théorème. D'après le lemme 1, il suffit d'établir (1.25). Nous allons montrer que, pour  $1 \leq i \leq p+q$

$$\lim_{t \rightarrow \infty} A_0 \dots A_{-t} e_i = 0, \quad p.s. \tag{1.27}$$

où  $e_i$  est le  $i$  - ème élément de la base canonique de  $\mathbb{R}^{p+q}$ . Soit  $(X_t)$  une solution strictement stationnaire de (1.5) et soit  $(z_t)$  défini par (1.17). On a pour  $t > 0$

$$\begin{aligned}
z_0 &= \underline{b}_0 + A_0 z_{-1} \\
&= \underline{b}_0 + \sum_{k=0}^{t-1} A_0 \dots A_{-k} \underline{b}_{-k-1} + A_0 \dots A_{-t} z_{-t-1} \\
&\geq \sum_{k=0}^{t-1} A_0 \dots A_{-k} \underline{b}_{-k-1}
\end{aligned}$$

car les coefficients des matrices  $A_t, \underline{b}_0$  et  $z_t$  sont positifs. Par suite la série  $\sum_{k=0}^{t-1} A_0 \dots A_{-k} \underline{b}_{-k-1}$  converge et donc  $A_0 \dots A_{-k} \underline{b}_{-k-1}$  tend presque sûrement vers 0 quand  $k \rightarrow \infty$ . Or  $\underline{b}_{-k-1} = \alpha_0 \varepsilon_{-k-1}^2 e_1 + \alpha_0 e_{q+1}$ . Donc  $A_0 \dots A_{-k} \underline{b}_{-k-1}$  se décompose en deux termes positifs et on a

$$\lim_{k \rightarrow \infty} A_0 \dots A_{-k} \alpha_0 \varepsilon_{-k-1}^2 e_1 = 0, \quad \lim_{k \rightarrow \infty} A_0 \dots A_{-k} \alpha_0 e_{q+1} = 0, \quad p.s. \quad (1.28)$$

Puisque  $\alpha_0 \neq 0$ , (1.27) est vraie pour  $i = q + 1$ . En utilisant la relation

$$A_{-k} e_{q+i} = \beta_i \varepsilon_{-k}^2 e_1 + \beta_i e_{q+1} + e_{q+i+1}, \quad i = 1, \dots, p, \quad (1.29)$$

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

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

donc (1.27) est vraie pour  $i = q + 2$ , et par récurrence, pour  $i = q + j, j = 1, \dots, p$  en utilisant (1.29). Par ailleurs, on remarque que  $A_{-k} e_q = \alpha_q \varepsilon_{-k}^2 e_1 + \alpha_q e^{q+1}$  ce qui permet de voir, d'après (1.28), que (1.27) est vérifiée pour  $i = q$ . On conclut pour les autres valeurs de  $i$  en utilisant

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

et une récurrence ascendante. Le théorème 18 est donc démontré.

**Remarque 3**

1) On aurait pu mener une démonstration analogue en utilisant la représentation vectorielle markovienne suivante fondée sur (1.6)

$$\underline{h}_t = \underline{\alpha}_0 + B_t \underline{h}_{t-1}, \tag{1.30}$$

avec  $\underline{\alpha}_0 = (\alpha_0, 0, \dots, 0) \in \mathbb{R}^r$ ,

$$\underline{h}_t = (\sigma_t^2, \dots, \sigma_{t-r+1}^2) \in \mathbb{R}^r, \quad B_t = \begin{pmatrix} a_1(\varepsilon_{t-1}) & \cdots & a_r(\varepsilon_{t-r}) \\ I_{r-1} & & 0 \end{pmatrix},$$

où  $I_{r-1}$  est la matrice identité de taille  $r - 1$ . Notons que les matrices  $B_t$  ne sont pas indépendantes, contrairement aux  $A_t$ . Il est amusant de noter

$$E \prod_{t=0}^n B_t = \prod_{t=0}^n E B_t. \tag{1.31}$$

La propriété d'indépendance des  $A_t$  sera cruciale pour trouver les conditions d'existence des moments, c'est pourquoi il est préférable de travailler avec la représentation (1.17).

2) Pour vérifier que  $\gamma < 0$ , il suffit de vérifier que

$$E(\log \|A_t A_{t-1} \dots A_1\|) < 0$$

pour un  $t > 0$ .

3) Si un modèle *GARCH* admet une solution strictement stationnaire, tout modèle *GARCH* obtenu en remplaçant les  $\alpha_i$  et  $\beta_j$  par des coefficients plus petits en admet également une. En effet le coefficient  $\gamma$  du modèle ainsi défini sera nécessairement inférieur à celui du modèle initial car, avec la norme utilisée,  $0 \leq A \leq B$  implique  $\|A\| \leq \|B\|$ . En particulier la stationnarité stricte du modèle *GARCH* implique celle du modèle *ARCH* obtenu en supprimant les coefficients  $\beta_j$ .

Le résultat suivant (établi par Bougerol et Picard (1992) [11]) fournit une condition nécessaire simple de stationnarité stricte.

**Corollaire 19** *Soit  $\gamma$  le plus grand exposant de Lyapounov de la suite  $\{A_t, t \in \mathbb{Z}\}$  défini par (1.18). Alors*

$$\gamma < 0 \Rightarrow \sum_{j=1}^p \beta_j < 1.$$

**Preuve**

Comme tous les termes des matrices  $A_t$  sont positifs, il est clair que  $\gamma$  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 (16) on voit que

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

où  $B$  est la matrice bloc de  $A_t$  définie par :

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

Il est facile de montrer (par récurrence sur  $p$  et en développant par rapport à la dernière colonne) que, pour  $\lambda \neq 0$ ,

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

où  $\mathbf{B}(z) = 1 - \beta_1 z - \cdots - \beta_p z^p$ . On en déduit que si  $\gamma < 0$  alors  $\mathbf{B}(z) = 0$  a toutes ses racines en dehors du cercle unité.

Montrons maintenant que

$$\{\mathbf{B}(z) = 0 \Rightarrow |z| > 1\} \Leftrightarrow \sum_{j=1}^p \beta_j < 1. \quad (1.32)$$

On a  $\mathbf{B}(0) = 1$  et  $\mathbf{B}(1) = 1 - \sum_{j=1}^p \beta_j$ . Donc si  $\sum_{j=1}^p \beta_j \geq 1$  alors  $\mathbf{B}(1) \leq 0$  et, par continuité, il existe une racine dans  $]0, 1]$ .

Inversement si  $\sum_{j=1}^p \beta_j < 1$  et si  $\mathbf{B}(z_0) = 0$  pour un  $z_0$  de module inférieur ou égal à 1 alors  $1 = \sum_{j=1}^p \beta_j z_0^j = \left| \sum_{j=1}^p \beta_j z_0^j \right| \leq \sum_{j=1}^p \beta_j |z_0^j| \leq \sum_{j=1}^p \beta_j$ , ce qui est impossible. D'où (1.32) et finalement le corollaire.

### Stationnarité au second ordre

Le théorème suivant donne des conditions nécessaires et suffisantes de stationnarité au second-ordre.

#### **Théorème 20 Stationnarité au second ordre du modèle GARCH(p, q)**

*S'il existe un processus GARCH(p, q), au sens de la définition 9, stationnaire au second-ordre et non anticipatif, et si  $\alpha_0 > 0$ , alors*

$$\sum_{i=1}^q \alpha_i + \sum_{j=1}^p \beta_j < 1. \quad (1.33)$$

*Inversement, si (1.33) est vérifiée, l'unique solution strictement stationnaire du modèle (1.5) est un bruit blanc (donc est stationnaire au second ordre). Il n'existe pas d'autre solution stationnaire au second ordre.*

#### **Preuve**

Montrons d'abord que la condition (1.33) est nécessaire. Soit  $(X_t)$  un processus GARCH(p, q), stationnaire au second ordre et non anticipatif. Alors la quantité

$$E(X_t^2) = E \{ EX_t^2 / X_s, s < t \} = E(\sigma_t^2)$$

est un réel positif, indépendant de  $t$ . En prenant l'espérance des deux membres de l'égalité (1.3), on tire donc

$$E(X_t^2) = \alpha_0 + \sum_{i=1}^q \alpha_i E(X_t^2) + \sum_{j=1}^p \beta_j E(X_t^2)$$

soit

$$\left(1 - \sum_{i=1}^q \alpha_i - \sum_{j=1}^p \beta_j\right) E(X_t^2) = \alpha_0. \quad (1.34)$$

Puisque  $\alpha_0$  est strictement positif, on doit avoir (1.33).

Supposons maintenant que (1.33) soit vraie et cherchons une solution *GARCH* au sens fort qui soit stationnaire. Pour  $t, k \in \mathbb{Z}$ , définissons les vecteurs à valeurs dans  $\mathbb{R}^d$  suivants :

$$Z_k(t) = \begin{cases} 0 & \text{si } k < 0 \\ \underline{b}_t + A_t Z_{k-1}(t-1) & \text{si } k \geq 0. \end{cases}$$

On a

$$Z_k(t) - Z_{k-1}(t) = \begin{cases} 0 & \text{si } k < 0 \\ \underline{b}_t & \text{si } k = 0 \\ A_t \{Z_{k-1}(t-1) - Z_{k-2}(t-1)\} & \text{si } k \geq 1 \end{cases}$$

En itérant ces relations on obtient, pour  $k > 0$

$$Z_k(t) - Z_{k-1}(t) = A_t A_{t-1} \dots A_{t-k+1} \underline{b}_{t-k}.$$

Par ailleurs, pour la norme victorienne  $\|V\| = \sum_i |v_i|$ , on a pour toute matrice aléatoire  $V$  à coefficients positifs,  $E \|V\| = E \sum_i |v_i| = \sum_i E |v_i| = \|E(V)\|$ . Donc, pour  $k > 0$

$$E \|Z_k(t) - Z_{k-1}(t)\| = \|E(A_t A_{t-1} \dots A_{t-k+1} \underline{b}_{t-k})\|,$$

car la matrice  $A_t A_{t-1} \dots A_{t-k+1} \underline{b}_{t-k}$  est positive. Tous les termes du produit  $A_t A_{t-1} \dots A_{t-k+1} \underline{b}_{t-k}$  sont indépendants (car le processus  $(\varepsilon_t)$  est *iid* et chaque terme du produit est fonction d'une variable  $\varepsilon_{t-i}$ , les dates  $t-i$  étant distinctes).

Par ailleurs,  $A := E(A_t)$  et  $\underline{b} = E(\underline{b}_t)$  ne dépendent évidemment pas de  $t$ . Finalement, pour  $k > 0$

$$E \|Z_k(t) - Z_{k-1}(t)\| = \|A^k \underline{b}\| = (1, \dots, 1) A^k \underline{b},$$

car tous les termes du vecteur  $A^k \underline{b}$  sont positifs.

La condition (1.33) implique que les valeurs propres de  $A$  sont de module strictement inférieur à 1 : en effet on vérifie que

$$\det(\lambda Id - A) = \lambda^{p+q} \left( 1 - \sum_{i=1}^q \alpha_i \lambda^{-i} - \sum_{j=1}^p \beta_j \lambda^{-j} \right), \quad (1.35)$$

par exemple en retranchant la  $q + 1$  - ème ligne de  $\lambda Id - A$  à la première, puis en développant le déterminant par rapport à la première ligne. Donc si  $|\lambda| \geq 1$ , en utilisant l'inégalité  $|a - b| \geq |a| - |b|$ , on obtient

$$|\det(\lambda Id - A)| \geq \left| 1 - \sum_{i=1}^q \alpha_i \lambda^{-i} - \sum_{j=1}^p \beta_j \lambda^{-j} \right| \geq 1 - \sum_{i=1}^q \alpha_i - \sum_{j=1}^p \beta_j > 0.$$

Par suite, en utilisant la décomposition de Jordan, ou (1.22), il est facile de voir que  $A^k \rightarrow 0$  à vitesse exponentielle quand  $k \rightarrow \infty$ . Donc pour  $t$  fixé,  $Z_k(t)$  converge à la fois aux sens  $L^1$ , en utilisant le critère de *Cauchy*, et presque sur quand  $k \rightarrow \infty$ . Soit  $\underline{z}_t$  la limite de  $(Z_k(t))_k$ . Ak fixé, le processus  $(Z_k(t))_{t \in \mathbb{Z}}$  est strictement stationnaire.

Le processus limite  $\underline{z}_t$  l'est donc également. Par ailleurs il est clair que  $\underline{z}_t$  vérifie l'équation (1.17).

#### Remarque 4

1) Sous les conditions du théorème, l'unique solution stationnaire du modèle (1.5) est, en utilisant (1.34), un bruit blanc de variance :

$$Var(X_t) = \frac{\alpha_0}{\left( 1 - \sum_{i=1}^q \alpha_i - \sum_{j=1}^p \beta_j \right)}.$$

2) Les conditions des théorèmes 18 et 20 étant nécessaires et suffisantes, on a forcément

$$\left[ \sum_{i=1}^q \alpha_i + \sum_{j=1}^p \beta_j < 1 \right] \Rightarrow \gamma < 0.$$

puisque la solution stationnaire au second ordre du théorème 20 l'est également strictement. On vérifie directement cette implication en remarquant que si (1.33) est vraie, d'après la preuve précédente, le rayon spectral  $\rho(A)$  est strictement inférieur à 1. Par ailleurs, d'après un résultat de Kesten et Spitzer (1984, (1.4)) [35], on a toujours :  $\gamma \leq \log \rho(A)$ .

Lorsque

$$\sum_{i=1}^q \alpha_i + \sum_{j=1}^p \beta_j = 1$$

## 1.5 Tests d'hétéroscédasticité :

Il existe plusieurs tests possibles : test de Goldfeld et Quandt, test de White, test de Breusch et Pagan et test ARCH de Engle. Nous étudierons ici le test ARCH car il est très fréquemment employé en économétrie des séries temporelles financières.

### 1.5.1 Test ARCH :

Le test ARCH consiste à effectuer une régression autorégressive des résidus carrés sur  $q$  retards :

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

où  $\varepsilon_t$  désigne le résidu à l'instant  $t$  issu de l'estimation,

Pour déterminer le nombre de retards  $q$ , on étudie le corrélogramme des résidus au carré.

Les hypothèses du test *ARCH* sont les suivantes :

$H_0$  : homoscedasticité et  $\alpha_1 = \dots = \alpha_q = 0$

$H_1$  : hétéroscédasticité et il y a au moins un coefficient  $\alpha_i$  significativement différent de 0.

Pour mener le test, on utilise la statistique de test  $n \times R^2$  où  $n$  correspond au nombre d'observations de la série  $e_t$  et  $R^2$  représente le coefficient de détermination associé à la régression  $\hat{u}_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i \hat{u}_{t-i}^2$ .

Sous l'hypothèse  $H_0$ , la statistique de test  $n \times R^2$  suit la loi du Khi-deux de degrés de liberté  $q$ . La règle de décision est alors :

-Si  $n \times R^2 \leq \chi^2(q)$  où  $\chi^2(q)$  désigne la valeur critique figurant dans la table du Khi-deux , on accepte ici l'hypothèse  $H_0$  d'homoscédasticité.

-Si  $n \times R^2 > \chi^2(q)$ , on rejette ici l'hypothèse  $H_0$  d'homoscédasticité et on admet qu'il y a de l'hétéroscédasticité.

.

# Chapitre 2

## Estimation des modèles GARCH

Dans cette section, nous allons traiter l'estimation des paramètres d'un modèle *GARCH*. Les modèles introduits reposent sur des formulations des moyennes et variances conditionnelles.

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

- La méthode de MV,
- La méthode de QMV,
- La méthode de régression artificielle.

### 2.1 Estimation des paramètres du modèle ARCH par la méthode de MV

L'estimation des paramètres de modèles *ARCH* se base très souvent sur la maximisation de la fonction de vraisemblance. Nous supposons que le processus  $(X_t)$  est conditionnellement gaussien. La vraisemblance associée à  $(X_t)$  conditionnellement au passé  $X_{t-1}$

est donné par :

$$L(X_t|X_{\underline{t-1}}; \theta) = \frac{1}{\sigma_t \sqrt{2\pi}} \exp\left(-\frac{X_t^2}{2\sigma_t^2}\right)$$

et dépend du vecteur de paramètres  $\theta = (\alpha_0, \alpha_1, \dots, \alpha_q)'$  à travers  $\sigma_t$ . La fonction de vraisemblance de  $(X_1; \dots; X_T)'$  conditionnelle à  $X_0 = 0$ , est par conséquent

$$L_T(X_1, X_2, \dots, X_T|X_{\underline{t-1}}; \theta) = \prod_{t=1}^T \frac{1}{\sigma_t \sqrt{2\pi}} \exp\left(-\frac{X_t^2}{2\sigma_t^2}\right)$$

l'estimateur est alors défini comme le vecteur

$$\hat{\theta}_T = (\hat{\alpha}_{0;T}, \dots, \hat{\alpha}_{q;T})'$$

qui maximise le logarithme de cette fonction de vraisemblance :

$$\hat{\theta}_T = \arg \max_{\theta} \ln L_T(X_1, X_2, \dots, X_T; \theta)$$

Sous diverses conditions de régularité, l'estimateur est convergent (Weiss 1986 ) [45].

## 2.2 Estimation des paramètres du modèle GARCH par la méthode de MV

La vraisemblance associée à  $(X_t)$  conditionnellement au passé  $X_{\underline{t-1}}$  s'écrit :

$$L(X_t|X_{\underline{t-1}}; \theta) = \frac{1}{\sigma_t \sqrt{2\pi}} \exp\left(-\frac{X_t^2}{2\sigma_t^2}\right)$$

et dépend du vecteur de paramètres  $\theta = (\alpha_0, \dots, \alpha_q, \beta_1, \dots, \beta_p)'$ , La fonction de vraisemblance de  $(X_1, \dots, X_T)'$  conditionnelle à  $X_0 = 0$ , est par conséquent

$$L(X_t|X_{\underline{t-1}}; \theta) = \frac{1}{\sigma_t \sqrt{2\pi}} \exp\left(-\frac{X_t^2}{2\sigma_t^2}\right)$$

et la fonction de vraisemblance totale est donnée par :

$$L_T(X_1, X_2, \dots, X_T | X_{\underline{t-1}}; \theta) = \prod_{t=1}^T L(X_t | X_{\underline{t-1}}; \theta)$$

Cette fonction de vraisemblance peut être calculée pour différentes valeurs du vecteur  $\theta_0$  et sa maximisation livre l'estimateur de maximum de vraisemblance.

La contribution de l'observation  $t$  à la fonction de log-vraisemblance à la forme :

$$l_t(\theta) = \frac{-1}{2} \ln(2\pi\sigma_t^2) - \frac{X_t^2}{2\sigma_t^2}.$$

Pour un échantillon de taille  $T$  la fonction de log-vraisemblance est s'écrit par :

$$L_T = \frac{1}{T} \sum_{t=1}^T l_t(\theta).$$

l'estimateur est alors défini par le vecteur

$$\hat{\theta}_T = (\hat{\alpha}_{0;T}, \dots, \hat{\alpha}_{q;T}, \hat{\beta}_{0;T}, \dots, \hat{\beta}_{p;T})'$$

qui maximise le logarithme de cette fonction de vraisemblance :

$$\hat{\theta}_T = \arg \max_{\theta} \ln L_T(X_1, X_2, \dots, X_T; \theta)$$

## 2.3 Estimation des paramètres du modèle GARCH par la méthode de QMV

Dans cette approche, l'utilisation de la méthode du QMV ou PMV est particulièrement intéressante pour les modèles GARCH car elle est valide, asymptotiquement, pour tout processus GARCH strictement stationnaire (sous des conditions de régularité mineure), sans hypothèse de moments sur le processus observé. Par conséquent, la fonction de vraisemblance définissant l'estimateur du MV sous l'hypothèse de normalité et la fonction de quasi-vraisemblance de l'estimateur QMV (ou PMV) sont les mêmes.

Les conditions de régularité sont toujours de trois types :

- Des conditions de stationnarité forte du processus,
- Des conditions d'existence des dérivées et des moments apparaissant dans les diverses formules,
- Des conditions d'indentifiabilité des paramètres  $\theta$ , qui doit pouvoir être retrouvé, sans ambiguïté à partir des deux premiers moments conditionnels.

On supposera que les observations  $X_1, \dots, X_n$  constituent une réalisation (de longueur  $n$ ) d'un processus  $GARCH(p, q)$ , solution strictement stationnaire non anticipative du modèle qui définit dans (1.5)

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

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

appartient à un espace de paramètres

$$\Theta \subset ]0, +\infty[ \times ]0, \infty[^{p+q}.$$

La vraie valeur du paramètre est inconnue. Elle est notée

$$\theta_0 = (\alpha_0, \alpha_{01}, \dots, \alpha_{0q}, \beta_{01}, \dots, \beta_{0p})'.$$

Pour écrire la vraisemblance du modèle, il faut spécifier une distribution particulière pour les variables iid  $\varepsilon_t$ . On considère généralement la quasi-vraisemblance gaussienne, i.e. la vraisemblance obtenue à partir d'une loi normale centrée réduite pour les  $\varepsilon_t$ . Nous ne ferons cependant pas l'hypothèse que cette loi constitue la vraie distribution du processus iid.

La spécification d'une distribution gaussienne pour les variables  $\varepsilon_t$  ne permet pas d'en déduire simplement la loi de l'échantillon. On travaille avec la vraisemblance de  $X_1, \dots, X_n$  conditionnellement à certaines valeurs initiales.

Étant donné des valeurs initiales  $X_1, \dots, X_{1-q}, \tilde{\sigma}_0^2, \dots, \tilde{\sigma}_{1-p}^2$  que nous allons préciser, la

vraisemblance conditionnelle gaussienne  $L_n(\theta)$  s'écrit

$$L_n(\theta) = Ln(\theta; X_1, \dots, X_n) = \prod_{t=1}^n \frac{1}{\sqrt{2\pi\tilde{\sigma}_t^2}} \exp\left(\frac{-X_t^2}{2\tilde{\sigma}_t^2}\right),$$

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

$$\tilde{\sigma}_t^2 = \tilde{\sigma}_t^2(\theta) = \alpha_0 + \sum_{i=1}^q \alpha_i X_{t-i}^2 + \sum_{j=1}^p \beta_j \tilde{\sigma}_{t-j}^2 \quad (2.1)$$

Pour une valeur donnée de  $\theta$ , sous l'hypothèse de stationnarité au second ordre, la variance non conditionnelle (correspondant à cette valeur de  $\theta$ ) est un choix raisonnable pour les valeurs initiales inconnues :

$$X_0^2 = \dots = X_{1-q}^2 = \sigma_0^2 = \dots = \sigma_{1-p}^2 = \frac{\alpha_0}{1 - \sum_{i=1}^q \alpha_i - \sum_{j=1}^p \beta_j}. \quad (2.2)$$

De telles valeurs initiales ne conviennent pas notamment pour les modèles *IGARCH*, pour lesquels l'hypothèse de stationnarité au second ordre est relâchée, car la constante (2.2) prendrait des valeurs négatives pour certaines valeurs de  $\theta$ . On peut alors prendre comme valeurs initiales

$$X_0^2 = \dots = X_{1-q}^2 = \tilde{\sigma}_0^2 = \dots = \tilde{\sigma}_{1-p}^2 = \alpha_0 \quad (2.3)$$

ou encore

$$X_0^2 = \dots = X_{1-q}^2 = \tilde{\sigma}_0^2 = \dots = \tilde{\sigma}_{1-p}^2 = X_1^2 \quad (2.4)$$

Un estimateur du QMV de  $\theta$  est défini comme toute solution mesurable  $\hat{\theta}_n$  de

$$\hat{\theta}_n = \arg_{\theta \in \Theta} \max L_n(\theta). \quad (2.5)$$

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

$$\tilde{\mathbf{I}}_n = n^{-1} \sum_{t=1}^n \tilde{\ell}_t, \text{ où } \tilde{\ell}_t = \tilde{\ell}_t(\theta) = \frac{X_t^2}{\tilde{\sigma}_t^2} + \log \tilde{\sigma}_t^2 \quad (2.6)$$

et  $\tilde{\sigma}_t^2$  est définie en (2.1). Un estimateur du quasi-maximum de vraisemblance est donc une solution mesurable de l'équation

$$\hat{\theta}_n = \arg_{\theta \in \Theta} \min \tilde{\mathbf{I}}_n(\theta). \quad (2.7)$$

Nous montrerons que le choix des valeurs initiales n'a pas d'importance pour les propriétés asymptotiques de l'estimateur du QMV. En pratique ce choix peut cependant avoir de l'importance. Notons que d'autres méthodes sont possibles pour générer la suite  $\tilde{\sigma}_t^2$ ; par exemple en prenant  $\tilde{\sigma}_t^2 = c_0(\theta) + \sum_{i=1}^{t-1} c_i(\theta) X_{t-i}^2$  ou les  $c_i(\theta)$  sont calculés récursivement (voir [6]). Remarquons que pour calculer  $\tilde{\mathbf{I}}_n(\theta)$ , cette procédure nécessite un nombre d'opérations de l'ordre de  $n^2$ , tandis que celle que nous proposons ne requiert qu'un ordre de  $n$  opérations.

### Equations de vraisemblance

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

$$\frac{1}{n} \sum_{t=1}^n (X_t^2 - \tilde{\sigma}_t^2) \frac{1}{\tilde{\sigma}_t^4} \frac{\partial \tilde{\sigma}_t^2}{\partial \theta} = 0. \quad (2.8)$$

Ces équations s'interprètent, pour  $n$  grand, comme des relations d'orthogonalité.

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

$$\frac{1}{n} \sum_{t=1}^n (X_t^2 - \sigma_t^2) \frac{1}{\sigma_t^4} \frac{\partial \sigma_t^2}{\partial \theta} \quad (2.9)$$

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

$$E_{\theta_0} \left( v_t \frac{1}{\sigma_t^4(\theta_0)} \frac{\partial \sigma_t^2(\theta_0)}{\partial \theta} \right) = 0$$

car  $\frac{1}{\sigma_t^4(\theta_0)} \frac{\partial \sigma_t^2(\theta_0)}{\partial \theta}$  est une fonction mesurable des  $X_{t-i}, i > 0$ . Ce résultat n'est autre que la version asymptotique de (2.8) en  $\theta_0$ , en utilisant le théorème d'ergodicité.

### Propriétés asymptotiques de l'estimateur du QMV

Nous utiliserons comme norme matricielle de  $A = (a_{ij})$  quelconque la norme  $\|A\| = \sum |a_{ij}|$ . Le rayon spectral d'une matrice carrée  $A$  sera noté par  $\rho(A)$ . Le produit de Kronecker sera noté  $\otimes$  et le symbole  $\xrightarrow{\mathcal{L}}$  désignera la convergence en loi.

### Convergence forte

Rappelons que le modèle (1.5) possède une solution strictement stationnaire si et seulement si la suite de matrices  $\mathbf{A}_0 = (A_{0t})$ , où

$$A_{0t} = \begin{pmatrix} \alpha_{01}\varepsilon_t^2 & \cdots & \alpha_{0q}\varepsilon_t^2 & \beta_{01}\varepsilon_t^2 & \cdots & \beta_{0p}\varepsilon_t^2 \\ 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 1 & 0 & 0 & \cdots & 0 & 0 \\ \alpha_{01} & \cdots & \alpha_{0q} & \beta_{01} & \cdots & \beta_{0p} \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 1 & 0 \end{pmatrix}$$

admet un coefficient de *Lyapounov* strictement négatif,  $\gamma(\mathbf{A}_0) < 0$ , où

$$\gamma(\mathbf{A}_0) := \inf_{t \in \mathbb{N}^*} \frac{1}{t} E(\log \|A_t A_{t-1} \dots A_1\|) = \lim_{t \rightarrow \infty} \frac{1}{t} \log \|A_t A_{t-1} \dots A_1\| \quad p.s. \quad (2.10)$$

Notons

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

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

Pour montrer la convergence forte, les hypothèses suivantes seront faites.

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

**A2** :  $\gamma(\mathbf{A}_0) < 0$  et  $\forall \theta \in \Theta, \sum_{j=1}^p \beta_j < 1$ .

**A3** :  $\varepsilon_t^2$  a une loi non dégénérée et  $E(\varepsilon_t^2) = 1$ .

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

### **Théorème 21 (Convergence forte de l'estimateur du QMV)**

Soit  $(\hat{\theta}_n)$  une suite d'estimateurs du QMV satisfaisant (2.5), avec les conditions initiales (2.3) ou (2.4). Sous les hypothèses A1-A4,

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

### **Remarque 5**

(1). On ne suppose pas que la vraie valeur  $\theta_0$  du paramètre appartient à l'intérieur de  $\Theta$ . Le théorème permet donc de traiter les cas où certains coefficients,  $\alpha_i$  ou  $\beta_j$ , sont nuls.

(2). Il est important de noter qu'on ne suppose pas non plus que la condition de stationnarité stricte est vérifiée sur tout  $\Theta$ . D'après le corollaire 17, la condition  $\sum_{j=1}^p \beta_j < 1$  est plus faible que celle de stationnarité stricte.

(3). L'hypothèse **A4** disparaît dans le cas ARCH. Dans le cas général, elle permet de sur identifier l'un des ordres,  $p$  ou  $q$ , mais pas les deux. On estime ainsi de manière convergente les paramètres d'un processus  $GARCH(p-1, q)$  (ou d'un  $GARCH(p, q-1)$ ) si on utilise un  $GARCH(p, q)$ .

(4). Quand  $p \neq 0$ , l'hypothèse **A4** exclut le cas où tous les  $\alpha_{0i}$  sont nuls. Ceci est

évidemment nécessaire, sinon le modèle a pour solution un bruit blanc fort qui peut s'écrire de multiples manières. Par exemple, un bruit blanc fort de variance 1 peut s'écrire sous la forme d'un  $GARCH(1, 1)$  avec  $\sigma_t^2 = \sigma^2(1 - \beta) + 0 \times X_{t-1}^2 + \beta\sigma_{t-1}^2$ .

(5).L'hypothèse d'absence de racines communes, dans **A4**, n'est restrictive que si  $p > 1$  et  $q > 1$ . En effet si  $q = 1$ , la seule racine de  $\mathcal{A}_{\theta_0}(z)$  est 0 et  $\mathcal{B}_{\theta_0}(z) \neq 0$ .

Si  $p = 1$  et  $\beta_{01} \neq 0$ , la seule racine de  $\mathcal{B}_{\theta_0}(z)$  est  $1/\beta_{01} > 0$  (si  $\beta_{01} = 0$ , le polynôme n'admet pas de racine). En raison de la positivité des coefficients  $\alpha_{0i}$ , cette valeur ne peut annuler  $\mathcal{A}_{\theta_0}(z)$ .

(6).L'hypothèse que  $E(\varepsilon_t) = 0$  n'est pas nécessaire pour les propriétés asymptotiques de l'estimateur du QMV d'un GARCH. La variance conditionnelle de  $t$  est donc, en général, seulement proportionnelle à  $h_t$  dans ce modèle :  $Var(X_t|X_u, u < t) = \{1 - (E\varepsilon_t)^2\}h_t$ . L'hypothèse  $E(\varepsilon_t^2) = 1$  est faite pour des raisons d'identifiabilité et n'est pas restrictive dès que  $E(\varepsilon_t^2) < \infty$ .

## 2.4 Estimation par régression artificielle du processus GARCH(1,1)

Soit un modèle de régression linéaire simple où les erreurs suivent un processus  $GARCH(1, 1)$  introduit par Bollerslev en 1986 [8]

$$Z = Y\delta + X, \tag{2.11}$$

avec

$$X_t = \sigma_t\varepsilon_t, \text{ où } \varepsilon_t \sim IIDN(0, 1)$$

et

$$\sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \beta_1 \sigma_{t-1}^2,$$

où  $Z$  est un vecteur  $(n \times 1)$ ,  $Y$  est une matrice  $(n \times k)$  et  $\delta$  un vecteur  $(k \times 1)$  : On notera  $\theta$  le vecteur  $(k + 3 \times 1)$  des paramètres du modèle (2.11)

$$\theta = (\delta, \alpha_0, \alpha_1, \beta_1)'$$

On considère la régression artificielle qui va nous permettre d'estimer le modèle de régression linéaire (1.11) à erreurs  $GARCH(1, 1)$ . Pour chaque observation correspondent deux nouvelles observations artificielles (empilées). Soit  $R(\theta)$  la matrice  $(2n \times k + 3)$  des régresseurs et  $r(\theta)$  le vecteur  $(2n)$  de la régressande, la régression artificielle est

$$r(\theta) = R(\theta)\Phi + \text{residus}$$

où

$$\begin{pmatrix} r_1(\theta) \\ r_2(\theta) \end{pmatrix} = \begin{pmatrix} R_1(\theta) \\ R_2(\theta) \end{pmatrix} \Phi + \text{residus}$$

Une estimation par la méthode des moindres carrés ordinaire (OLS) de cette régression, nous donne

$$\Phi = (R(\theta)^T R(\theta))^{-1} R(\theta)^T Y(\theta)$$

- Pour l'observation  $t$  les deux éléments qui correspondent la régressande  $r(\theta)$  sont

$$r_1(\theta) = \frac{X_t}{\sigma_t} \tag{2.12}$$

$$r_2(\theta) = \frac{(X_t^2 - \sigma_t^2)}{\sigma_t^2 \sqrt{2}} \tag{2.13}$$

- Les éléments de  $R(\theta)$  de la régression artificielle qui correspondent au paramètre  $\delta$  sont les éléments

$$R_{1t}(\delta) = \frac{Y_t}{\sigma_t} \quad (2.14)$$

$$R_{2t}(\delta) = -\frac{\alpha_1 \sqrt{2}}{\sigma_t^2} \sum_{s=1}^{t-1} \beta^{s-1} Y_{t-s} X_{t-s} \quad (2.15)$$

- Similairement, les éléments des régresseurs qui correspondent au paramètre  $\alpha_0$  sont

$$R_{1t}(\alpha_0) = 0 \quad (2.16)$$

$$R_{2t}(\alpha_0) = \frac{1}{\sigma_t^2 \sqrt{2}} \left( \frac{1 - \beta^{t-1}}{1 - \beta} + \beta^{t-1} \frac{\partial \sigma_x^2}{\partial \alpha_1} \right) \quad (2.17)$$

- et les régresseurs correspondant au paramètre  $\alpha_1$  sont

$$R_{1t}(\alpha_1) = 0 \quad (2.18)$$

$$R_{2t}(\alpha_1) = \frac{1}{\sigma_t^2 \sqrt{2}} \left( \sum_{s=1}^{t-1} \beta^{s-1} X_{t-s}^2 + \beta^{t-1} \frac{\partial \sigma_x^2}{\partial \alpha_0} \right) \quad (2.19)$$

- Finalement, les éléments des régresseurs qui correspondent à  $\beta$  sont

$$R_{1t}(\beta) = 0 \quad (2.20)$$

$$R_{2t}(\beta) = \frac{1}{\sigma_t^2 \sqrt{2}} \left( -\frac{\alpha_0 (t-1) \beta^{t-2}}{1 - \beta} + \frac{\alpha_1 (1 - \beta^{t-1})}{(1 - \beta)^2} + \alpha_1 \sum_{s=1}^{t-1} (s-1) \beta^{t-2} X_{t-s}^2 + \beta^{t-1} \frac{\partial \sigma_x^2}{\partial \beta} \right) \quad (2.21)$$

On montre que le gradient du log-vraisemblance est nul quand il est évalué en des valeurs des paramètres pour lesquels la régressande est orthogonale aux régresseurs. Plus précisément, pour chaque observation et chaque paramètre, le produit des premiers éléments de la régressande avec les premiers éléments correspondants des régresseurs, plus le

produit des seconds éléments de la régressande avec les seconds éléments des régresseurs est égal à la dérivée partielle de  $l_t$  par rapport au paramètre soit

$$R_{1t}(\delta)r_{1t}(\delta) + R_{2t}(\delta)r_{2t}(\delta) = \frac{\partial l_t}{\partial \delta}$$

$$R_{1t}(\alpha_0)^T r_{1t}(\alpha_0) + R_{2t}(\alpha_0)^T r_{2t}(\alpha_0) = \frac{\partial l_t}{\partial \alpha_0}$$

$$R_{1t}(\alpha_1)^T r_{1t}(\alpha_1) + R_{2t}(\alpha_1)^T r_{2t}(\alpha_1) = \frac{\partial l_t}{\partial \alpha_1}$$

$$R_{1t}(\beta)^T r_{1t}(\beta) + R_{2t}(\beta)^T r_{2t}(\beta) = \frac{\partial l_t}{\partial \beta}$$

Finalement, on peut déduire que la contribution au gradient du log-vraisemblance est donnée comme suit

$$\frac{\partial l_t}{\partial \delta_i} = \frac{\partial l_t}{\partial X_t} \frac{\partial X_t}{\partial \delta_i} + \frac{\partial l_t}{\partial \sigma_t^2} \frac{\partial \sigma_t^2}{\partial \delta_i} = \frac{Y_{ti} X_t}{\sigma_t^2} - \frac{\alpha_1 (X_t^2 - \sigma_t^2)}{\sigma_t^4} \sum_{s=1}^{t-1} \beta^{s-1} Y_{ti} X_{t-s}^2$$

$$\frac{\partial l_t}{\partial \alpha_0} = \frac{\partial l_t}{\partial \sigma_t^2} \frac{\partial \sigma_t^2}{\partial \alpha_0} = \frac{(X_t^2 - \sigma_t^2)}{2\sigma_t^4} \left( \frac{1 - \beta^{t-1}}{1 - \beta} + \beta^{t-1} \frac{\partial \sigma_1^2}{\partial \alpha_0} \right)$$

$$\frac{\partial l_t}{\partial \alpha_1} = \frac{\partial l_t}{\partial \sigma_t^2} \frac{\partial \sigma_t^2}{\partial \alpha_1} = \frac{(X_t^2 - \sigma_t^2)}{2\sigma_t^4} \left( \sum_{s=1}^{t-1} \beta^{s-1} X_{t-s}^2 + \beta^{t-1} \frac{\partial \sigma_1^2}{\partial \alpha_1} \right)$$

$$\begin{aligned} \frac{\partial l_t}{\partial \beta} &= \frac{\partial l_t}{\partial \sigma_t^2} \frac{\partial \sigma_t^2}{\partial \beta} = \frac{(X_t^2 - \sigma_t^2)}{2\sigma_t^4} \left( -\frac{\alpha_0(t-1)\beta^{t-2}}{1-\beta} + -\frac{\alpha_0(1-\beta^{t-1})}{(1-\beta)^2} \right. \\ &\quad \left. + \alpha_1 \sum_{s=1}^{t-1} (s-1)\beta^{t-2} X_{t-s}^2 + \beta^{t-1} \frac{\partial \sigma_u^2}{\partial \beta} \right) \end{aligned}$$

Il est bien sûr impossible d'utiliser ces dernières expressions sans connaître la valeur de  $\sigma_t^2$ .

Pour cela, on a écrit  $\sigma_t^2$  sous une forme suivante qui nous permet de calculer simplement ces dérivées partielles :

$$\sigma_t^2 = \frac{\alpha_0(1 - \beta^{t-1})}{1 - \beta} + \alpha_1 \sum_{s=1}^{t-1} \beta^{t-1} X_{t-s}^2 + \beta^{t-1} \sigma_1^2$$

En premier pour les éléments du vecteur  $\delta$  on forme l'expression

$$\begin{aligned} & \frac{Y_{ti} X_t}{\sigma_t \sigma_t} - \frac{X_t^2 - \sigma_t^2}{\sigma_t^2 \sqrt{2}} \frac{\alpha_1 \sqrt{2}}{\sigma_t^2} \sum_{s=1}^{t-1} \beta^{s-1} Y_{ti} X_{t-s}^2 \\ &= \frac{Y_{ti} X_t}{\sigma_t^2} - \frac{\alpha_1 (X_t^2 - \sigma_t^2)}{\sigma_t^4} \sum_{s=1}^{t-1} \beta^{s-1} Y_{ti} X_{t-s}^2 = \frac{\partial l_t}{\partial \delta} \end{aligned}$$

De même, pour  $\alpha_0$ , on a

$$\frac{X_t^2 - \sigma_t^2}{\sigma_t^2 \sqrt{2}} \frac{1}{\sigma_t^2 \sqrt{2}} \left( \frac{1 - \beta^{t-1}}{1 - \beta} + \beta^{t-1} \frac{\partial \sigma_1^2(\alpha_0, \alpha_1, \beta)}{\partial \alpha_0} \right) = \frac{\partial l_t}{\partial \alpha_0}$$

puis pour  $\alpha_1$

$$\frac{X_t^2 - \sigma_t^2}{\sigma_t^2 \sqrt{2}} \frac{1}{\sigma_t^2 \sqrt{2}} \left( \sum_{s=1}^{t-1} \beta^{s-1} X_{t-s}^2 + \beta^{t-1} \frac{\partial \sigma_1^2(\alpha_0, \alpha_1, \beta)}{\partial \alpha_1} \right) = \frac{\partial l_t}{\partial \alpha_1}$$

Et enfin pour  $\beta$

$$\begin{aligned} & \frac{X_t^2 - \sigma_t^2}{\sigma_t^2 \sqrt{2}} \frac{1}{\sigma_t^2 \sqrt{2}} \left( -\frac{\alpha_0(t-1)\beta^{t-2}}{1 - \beta} + -\frac{\alpha_0(1 - \beta^{t-1})}{(1 - \beta)^2} + \alpha_1 \sum_{s=1}^{t-1} (s-1)\beta^{t-2} X_{t-s}^2 \right. \\ & \left. + \beta^{t-1} \frac{\partial \sigma_1^2(\alpha_0, \alpha_1, \beta)}{\partial \beta} \right) = \frac{\partial l_t}{\partial \beta} \end{aligned}$$

On va maintenant montrer que  $R(\theta)^T R(\theta)$  est la matrice d'information de la vraisemblance  $I$

$$R(\theta)^T R(\theta) + I$$

où en terme de contribution, si  $\phi_1$  et  $\phi_2$  sont des paramètres du modèle, c'est-à-dire des éléments de  $\delta$  ou de la variance conditionnelle

$$E \left( \frac{\partial l_t}{\partial \phi_1} \frac{\partial l_t}{\partial \phi_2} / \Omega_t \right) = R_{1t}(\phi_1)^T R_{1t}(\phi_2) + R_{2t}(\phi_1)^T R_{2t}(\phi_2)$$

Pour cela on remarque que  $\sigma_t^2$  fait partie de l'ensemble d'informations  $\Omega_t$ . on voit qu'une contribution à la matrice d'information peut s'exprimer comme

$$\begin{aligned} E \left( \frac{\partial l_t}{\partial \phi_1} \frac{\partial l_t}{\partial \phi_2} / \Omega_t \right) &= E \left( \left( \frac{\partial l_t}{\partial X_t} \frac{\partial X_t}{\partial \phi_1} + \frac{\partial l_t}{\partial \sigma_t^2} \frac{\partial \sigma_t^2}{\partial \phi_1} \right) \left( \frac{\partial l_t}{\partial X_t} \frac{\partial X_t}{\partial \phi_2} + \frac{\partial l_t}{\partial \sigma_t^2} \frac{\partial \sigma_t^2}{\partial \phi_2} \right) / \Omega_t \right) \\ &= \frac{1}{\sigma_t^4} E(X_t^2) \frac{\partial X_t}{\partial \phi_1} \frac{\partial X_t}{\partial \phi_2} + \frac{1}{4\sigma_t^8} E((X_t^2 - \sigma_t^2)^2) \frac{\partial \sigma_t^2}{\partial \phi_1} \frac{\partial \sigma_t^2}{\partial \phi_2} \\ &= \frac{1}{\sigma_t^2} \frac{\partial X_t}{\partial \phi_1} \frac{\partial X_t}{\partial \phi_2} + \frac{1}{2\sigma_t^4} \frac{\partial \sigma_t^2}{\partial \phi_1} \frac{\partial \sigma_t^2}{\partial \phi_2} \end{aligned}$$

Commençons par regarder pour  $\phi_1 = \delta_i, \phi_2 = \delta_j$ ,

$$\begin{aligned} E \left( \frac{\partial l_t}{\partial \delta_i} \frac{\partial l_t}{\partial \delta_j} / \Omega_t \right) &= \frac{1}{\sigma_t^2} \frac{\partial X_t}{\partial \delta_i} \frac{\partial X_t}{\partial \delta_j} + \frac{1}{2\sigma_t^4} \frac{\partial \sigma_t^2}{\partial \delta_i} \frac{\partial \sigma_t^2}{\partial \delta_j} \\ &= \frac{1}{\sigma_t^2} X_{ti} X_{tj} + \frac{2\alpha_1}{\sigma_t^4} \left( \sum_{s=1}^{t-1} \beta^{s-1} X_{t-s}^2 Y_{(t-s)i} \right) \left( \sum_{s=1}^{t-1} \beta^{s-1} X_{t-s}^2 Y_{(t-s)j} \right) \\ &= R_{1t}(\delta_i)^T R_{1t}(\delta_j) + R_{2t}(\delta_i)^T R_{2t}(\delta_j) \end{aligned}$$

Ce qui est bien la contribution de l'observation  $t$  de l'élément qui correspond à  $\delta_i$  et  $\delta_j$  du produit matriciel  $R(\delta, \theta)^T R(\delta, \theta)$ . Pour les paramètres  $\phi_1 = \delta_i, \phi_2 = \alpha_0$

$$\begin{aligned}
E\left(\frac{\partial l_t}{\partial \delta_i} \frac{\partial l_t}{\partial \alpha_0} / \Omega_t\right) &= \frac{1}{\sigma_t^2} \frac{\partial X_t}{\partial \delta_i} \frac{\partial X_t}{\partial \alpha_0} + \frac{1}{2\sigma_t^4} \frac{\partial \sigma_t^2}{\partial \delta_i} \frac{\partial \sigma_t^2}{\partial \alpha_0} \\
&= 0 + \frac{1}{2\sigma_t^4} \left( -2\alpha_1 \sum_{s=1}^{t-1} \beta^{s-1} X_{t-s}^2 Y_{(t-s)i} \right) \left( \frac{1 - \beta^{t-1}}{1 - \beta} + \beta^{t-1} \frac{\partial \sigma_x^2}{\partial \alpha_0} \right) \\
&= \frac{-\alpha_1}{\sigma_t^4} \left( \sum_{s=1}^{t-1} \beta^{s-1} X_{t-s}^2 Y_{(t-s)i} \right) \left( \frac{1 - \beta^{t-1}}{1 - \beta} + \beta^{t-1} \frac{\partial \sigma_x^2}{\partial \alpha_0} \right) \\
&= R_{1t}(\delta_i)^T R_{1t}(\alpha_0) + R_{2t}(\delta_i)^T R_{2t}(\alpha_0)
\end{aligned}$$

Pour les paramètres  $\phi_1 = \delta_i$ ,  $\phi_2 = \alpha_1$

$$\begin{aligned}
E\left(\frac{\partial l_t}{\partial \delta_i} \frac{\partial l_t}{\partial \alpha_0} / \Omega_t\right) &= \frac{1}{\sigma_t^2} \frac{\partial X_t}{\partial \delta_i} \frac{\partial X_t}{\partial \alpha_1} + \frac{1}{2\sigma_t^4} \frac{\partial \sigma_t^2}{\partial \delta_i} \frac{\partial \sigma_t^2}{\partial \alpha_1} \\
&= 0 + \frac{1}{2\sigma_t^4} \left( -2\alpha_1 \sum_{s=1}^{t-1} \beta^{s-1} X_{t-s}^2 Y_{(t-s)i} \right) \left( \sum_{s=1}^{t-1} \beta^{s-1} X_{t-s}^2 + \beta^{t-1} \frac{\partial \sigma_x^2}{\partial \alpha_1} \right) \\
&= \frac{-\alpha_1}{\sigma_t^4} \left( \sum_{s=1}^{t-1} \beta^{s-1} X_{t-s}^2 Y_{(t-s)i} \right) \left( \sum_{s=1}^{t-1} \beta^{s-1} X_{t-s}^2 + \beta^{t-1} \frac{\partial \sigma_x^2}{\partial \alpha_1} \right) \\
&= R_{1t}(\delta_i)^T R_{1t}(\alpha_1) + R_{2t}(\delta_i)^T R_{2t}(\alpha_1)
\end{aligned}$$

Pour les paramètres  $\phi_1 = \delta_i$ ,  $\phi_2 = \beta$

$$\begin{aligned}
E\left(\frac{\partial l_t}{\partial \delta_i} \frac{\partial l_t}{\partial \beta} / \Omega_t\right) &= \frac{1}{\sigma_t^2} \frac{\partial X_t}{\partial \delta_i} \frac{\partial X_t}{\partial \beta} + \frac{1}{2\sigma_t^4} \frac{\partial \sigma_t^2}{\partial \delta_i} \frac{\partial \sigma_t^2}{\partial \beta} \\
&= 0 - \frac{\alpha_1}{\sigma_t^4} \left( \sum_{s=1}^{t-1} \beta^{s-1} X_{t-s}^2 Y_{(t-s)i} \right) \left( -\frac{\alpha_0(t-1)\beta^{t-2}}{1-\beta} + \frac{\alpha_0(1-\beta^{t-1})}{(1-\beta)^2} \right. \\
&\quad \left. + \alpha_1 \sum_{s=1}^{t-1} (s-1)\beta^{t-2} X_{t-s}^2 + \beta^{t-1} \frac{\partial \sigma_x^2}{\partial \beta} \right) \\
&= R_{1t}(\delta_i)^T R_{1t}(\beta) + R_{2t}(\delta_i)^T R_{2t}(\beta)
\end{aligned}$$

Pour les paramètres  $\phi_1 = \alpha_0$ ,  $\phi_2 = \alpha_0$

$$\begin{aligned}
E\left(\frac{\partial l_t}{\partial \alpha_0} \frac{\partial l_t}{\partial \alpha_0} / \Omega_t\right) &= \frac{1}{\sigma_t^2} \frac{\partial X_t}{\partial \alpha_0} \frac{\partial X_t}{\partial \alpha_0} + \frac{1}{2\sigma_t^4} \frac{\partial \sigma_t^2}{\partial \alpha_0} \frac{\partial \sigma_t^2}{\partial \alpha_0} \\
&= \frac{1}{2\sigma_t^4} \left( \sum_{s=1}^{t-1} \beta^{s-1} X_{t-s}^2 + \beta^{t-1} \frac{\partial \sigma_x^2}{\partial \alpha_1} \right)^2 \\
&= R_{1t}(\alpha_0)^T R_{1t}(\alpha_0) + R_{2t}(\alpha_0)^T R_{2t}(\alpha_0)
\end{aligned}$$

Pour les paramètres  $\phi_1 = \alpha_0$ ,  $\phi_2 = \alpha_1$

$$\begin{aligned}
E\left(\frac{\partial l_t}{\partial \alpha_0} \frac{\partial l_t}{\partial \alpha_1} / \Omega_t\right) &= \frac{1}{\sigma_t^2} \frac{\partial X_t}{\partial \alpha_0} \frac{\partial X_t}{\partial \alpha_1} + \frac{1}{2\sigma_t^4} \frac{\partial \sigma_t^2}{\partial \alpha_0} \frac{\partial \sigma_t^2}{\partial \alpha_1} \\
&= 0 + \frac{1}{2\sigma_t^4} \left( \frac{1 - \beta^{t-1}}{1 - \beta} + \beta^{t-1} \frac{\partial \sigma_x^2}{\partial \alpha_0} \right) \left( \sum_{s=1}^{t-1} \beta^{s-1} X_{t-s}^2 + \beta^{t-1} \frac{\partial \sigma_x^2}{\partial \alpha_1} \right) \\
&= R_{1t}(\alpha_0)^T R_{1t}(\alpha_1) + R_{2t}(\alpha_0)^T R_{2t}(\alpha_1)
\end{aligned}$$

Pour les paramètres  $\phi_1 = \alpha_0$ ,  $\phi_2 = \beta$

$$\begin{aligned}
E\left(\frac{\partial l_t}{\partial \alpha_0} \frac{\partial l_t}{\partial \beta} / \Omega_t\right) &= \frac{1}{\sigma_t^2} \frac{\partial X_t}{\partial \alpha_0} \frac{\partial X_t}{\partial \beta} + \frac{1}{2\sigma_t^4} \frac{\partial \sigma_t^2}{\partial \alpha_0} \frac{\partial \sigma_t^2}{\partial \beta} \\
&= \frac{1}{2\sigma_t^4} \left( \frac{1 - \beta^{t-1}}{1 - \beta} + \beta^{t-1} \frac{\partial \sigma_x^2}{\partial \alpha_0} \right) + \left( -\frac{\alpha_0(t-1)\beta^{t-2}}{1 - \beta} + \right. \\
&\quad \left. + \frac{\alpha_0(1 - \beta^{t-1})}{(1 - \beta)^2} + \alpha_1 \sum_{s=1}^{t-1} (s-1)\beta^{t-2} X_{t-s}^2 + \beta^{t-1} \frac{\partial \sigma_x^2}{\partial \beta} \right) \\
&= R_{1t}(\alpha_0)^T R_{1t}(\beta) + R_{2t}(\alpha_0)^T R_{2t}(\beta)
\end{aligned}$$

Pour les paramètres  $\phi_1 = \alpha_0$ ,  $\phi_2 = \alpha_1$

$$\begin{aligned}
E \left( \frac{\partial l_t}{\partial \alpha_1} \frac{\partial l_t}{\partial \alpha_1} / \Omega_t \right) &= \frac{1}{\sigma_t^2} \frac{\partial X_t}{\partial \alpha_1} \frac{\partial X_t}{\partial \alpha_1} + \frac{1}{2\sigma_t^4} \frac{\partial \sigma_t^2}{\partial \alpha_1} \frac{\partial \sigma_t^2}{\partial \alpha_1} \\
&= 0 + \frac{1}{2\sigma_t^4} \left( \sum_{s=1}^{t-1} \beta^{s-1} X_{t-s}^2 + \beta^{t-1} \frac{\partial \sigma_x^2}{\partial \alpha_1} \right)^2 \\
&= R_{1t}(\alpha_1)^T R_{1t}(\alpha_1) + R_{2t}(\alpha_1)^T R_{2t}(\alpha_1)
\end{aligned}$$

Pour les paramètres  $\phi_1 = \alpha_0$ ,  $\phi_2 = \beta$

$$\begin{aligned}
E \left( \frac{\partial l_t}{\partial \alpha_1} \frac{\partial l_t}{\partial \beta} / \Omega_t \right) &= \frac{1}{\sigma_t^2} \frac{\partial X_t}{\partial \alpha_1} \frac{\partial X_t}{\partial \beta} + \frac{1}{2\sigma_t^4} \frac{\partial \sigma_t^2}{\partial \alpha_1} \frac{\partial \sigma_t^2}{\partial \beta} \\
&= \frac{1}{2\sigma_t^4} \left( \sum_{s=1}^{t-1} \beta^{s-1} X_{t-s}^2 + \beta^{t-1} \frac{\partial \sigma_x^2}{\partial \alpha_1} \right) + \left( -\frac{\alpha_0(t-1)\beta^{t-2}}{1-\beta} + \frac{\alpha_0(1-\beta^{t-1})}{(1-\beta)^2} \right. \\
&\quad \left. + \alpha_1 \sum_{s=1}^{t-1} (s-1)\beta^{t-2} X_{t-s}^2 + \beta^{t-1} \frac{\partial \sigma_x^2}{\partial \beta} \right) \\
&= R_{1t}(\alpha_1)^T R_{1t}(\beta) + R_{2t}(\alpha_1)^T R_{2t}(\beta)
\end{aligned}$$

Pour les paramètres  $\phi_1 = \beta$ ,  $\phi_2 = \beta$

$$\begin{aligned}
E \left( \frac{\partial l_t}{\partial \beta} \frac{\partial l_t}{\partial \beta} / \Omega_t \right) &= \frac{1}{\sigma_t^2} \frac{\partial X_t}{\partial \beta} \frac{\partial X_t}{\partial \beta} + \frac{1}{2\sigma_t^4} \frac{\partial \sigma_t^2}{\partial \beta} \frac{\partial \sigma_t^2}{\partial \beta} \\
&= \frac{1}{2\sigma_t^4} \left( -\frac{\alpha_0(t-1)\beta^{t-2}}{1-\beta} + \frac{\alpha_0(1-\beta^{t-1})}{(1-\beta)^2} + \alpha_1 \sum_{s=1}^{t-1} (s-1)\beta^{t-2} X_{t-s}^2 + \beta^{t-1} \frac{\partial \sigma_x^2}{\partial \beta} \right)^2 \\
&= R_{1t}(\beta)^T R_{1t}(\beta) + R_{2t}(\beta)^T R_{2t}(\beta)
\end{aligned}$$

On vient donc de montrer que la régression artificielle que nous avons construite correspond à une maximisation par quasi-newton avec comme approximation numérique de la Hessienne la matrice d'information du maximum de vraisemblance, on parle d'estimateur du score efficace, le processus itératif est le suivant

$$\phi_{i+1} = \phi_i + \kappa (R(\theta)^T R(\theta))^{-1} R(\theta)^T r(\theta)$$

équivalent à

$$\theta_{i+1} = \theta_i + \kappa I^{-1} \frac{\partial L}{\partial \theta}$$

en terme d'algorithme du gradient.

# Chapitre 3

## Les Modèles GARCH multivariés

Les modèles univariés ne permettent que d'analyser une série financière à la fois. Mais dans la pratique, le fait d'analyser qu'une série à la fois n'est pas très utile.

L'intérêt pour un tel modèle est de pouvoir examiner et analyser les diverses relations qu'ont les différentes séries entre elles. Afin de pouvoir estimer plusieurs séries financières pour analyser leurs corrélations et les transferts de volatilité, il faut utiliser les modèles GARCH multivariés (MVGARCH). Ce modèle multivarié est souvent utilisé afin de tester les transferts de volatilité ainsi que les effets de débordement dans les études portant sur la contagion (Silvennoinen et Teräsvirta, 2008) [44]. Ce modèle multivarié permet non seulement d'analyser les volatilités, mais aussi les corrélations de différents marchés. Il existe plusieurs spécifications au modèle GARCH multivarié; les modèles linéaires, non linéaires, asymétriques, modèles à sauts, modèles à corrélations conditionnelles et plusieurs autres.

### 3.1 La notion de causalité

La loi du processus est la loi du couple  $(X_t, Y_t)$ . Cette dernière s'écrit, à la date  $t$ , conditionnellement au passé (noté  $\underline{x_{t-1}}$  et  $\underline{y_{t-1}}$ ),  $\ell(x_t, y_t / \underline{x_{t-1}}, \underline{y_{t-1}})$ . Dans le cas où les processus  $(X_t)$  et  $(Y_t)$  sont indépendants, alors

$$\ell(x_t, y_t / \underline{x}_{t-1}, \underline{y}_{t-1}) = \ell(x_t / \underline{x}_{t-1}) \times \ell(y_t / \underline{y}_{t-1})$$

### 3.1.1 Causalité au sens de Granger

#### Causalité en moyenne

Soient  $(X_t)$  et  $(Y_t)$  deux séries temporelles, et notons le passé de  $(X_t)$  et  $(Y_t)$ ,  $\underline{X}_t = \{X_t, X_{t-1}, \dots\}$  et  $\underline{Y}_t = \{Y_t, Y_{t-1}, \dots\}$ . Granger a introduit en 1969 [28] différentes notions de causalité :

*i)*  $Y$  cause  $X$  à la date  $t$  si et seulement si

$$E(X_t / \underline{X}_{t-1}, \underline{Y}_{t-1}) \neq E(X_t / \underline{X}_{t-1}).$$

*ii)*  $Y$  cause  $X$  instantanément à la date  $t$  si et seulement si

$$E(X_t / \underline{X}_{t-1}, \underline{Y}_t) \neq E(X_t / \underline{X}_{t-1}, \underline{Y}_{t-1}).$$

- On dit que  $Y$  ne cause pas  $X$  si et seulement si

$$E(X_t / \underline{X}_{t-1}, \underline{Y}_{t-1}) = E(X_t / \underline{X}_{t-1}), \text{ presque sûrement pour tout } t.$$

Il y a une équivalence entre

- 1)  $X$  ne cause pas  $Y$  instantanément à la date  $t$
- 2)  $Y$  ne cause pas  $X$  instantanément à la date  $t$ .

#### Causalité en variance

*i)*  $Y$  cause  $X$  ssi

$$Var(X_t / \underline{X}_{t-1}, \underline{Y}_{t-1}) < Var(X_t / \underline{X}_{t-1}).$$

pour au moins une valeur de  $t$ .

ii)  $Y$  cause  $X$  instantanément ssi

$$\text{Var}(X_t/\underline{X}_{t-1}, \underline{Y}_t) < \text{Var}(X_t/\underline{X}_{t-1}, \underline{Y}_{t-1}).$$

pour au moins une valeur de  $t$ .

iii) **Rétroaction.** Il y a rétroaction entre  $X$  et  $Y$  ssi  $X$  cause  $Y$  et  $Y$  cause  $X$ .

- On dit que  $Y$  ne cause pas  $X$  ssi

$$\text{Var}(X_t/\underline{X}_{t-1}, \underline{Y}_{t-1}) = \text{Var}(X_t/\underline{X}_{t-1}), \forall t.$$

En d'autres termes,  $Y$  ne cause pas  $X$  si et seulement si la prédiction de  $X$  sachant l'historique de  $Y$  et  $X$  reste la même que si on connaît seulement l'historique de  $X$ .

## 3.2 La notion de cointégration

L'analyse de la cointégration permet d'identifier la relation entre plusieurs variables. Cette notion a été introduite dès 1974 par Granger et Newbold [27], sous le nom de "spurious régressions", ou régressions fallacieuses, puis formalisée par Engle et Granger en 1987 [21], et enfin par Johansen en 1991 [31] et en 1995 [32].

### 3.2.1 Principe de la cointégration

En scrutant les relations d'équilibre à long terme des variables, l'analyse de la cointégration permet d'identifier l'existence des séries suivant un même sentier d'équilibre à long terme (l'existence d'un éventuel vecteur de cointégration) et d'éliminer son effet le cas échéant.

Une série est intégrée d'ordre  $d$  s'il convient de la différencier  $d$  fois avant de la stationnariser.

**Définition 22** *La série  $(X_t)$  sera dite intégrée d'ordre  $d(d \geq 1)$  si  $\Delta^{d-1}X_t$  n'est pas stationnaire et  $\Delta^d X_t$  est stationnaire. Une série stationnaire sera dite intégrée d'ordre 0.*

### Remarque 6

Soient  $(X_t)$  une série stationnaire et  $(Y_t)$  intégrée d'ordre 1, alors  $(X_t + Y_t)$  est intégrée d'ordre 1.

Toutefois, si  $(X_t)$  et  $(Y_t)$  sont intégrées d'ordre  $d$ , alors  $(X_t + Y_t)$  peut être soit intégrée d'ordre  $d$ , soit stationnaire (dans le cas où les tendances s'annulent).

Deux séries  $(X_t)$  et  $(Y_t)$  sont cointégrées si

(i)  $(X_t)$  et  $(Y_t)$  sont intégrées d'ordre  $d$

(ii) il existe une combinaison linéaire de ces séries qui soit intégrée d'ordre strictement inférieur à  $d$ , noté  $d - b$

Dans le cas de l'intégration, on notera  $X_t \sim I(d)$ , et pour la cointégration  $X_t, Y_t \sim CI(d, b)$ . Le vecteur  $(\alpha, \beta)$  tel que  $\alpha X_t + \beta Y_t \sim I(d - b)$  sera appelé vecteur de cointégration.

On généralisé (à plus de deux séries) comme suit :

Soit  $Z_t$ , le vecteur de séries chronologiques, les composantes de  $Z_t$  sont dites cointégrées d'ordre  $(d, b)$  si :

(i) toutes les composantes de  $Z_t$  sont  $I(d)$

(ii) il y a au moins un vecteur de coefficients  $\beta = [\beta_1, \beta_2, \dots, \beta_n]$  tel que :  $\beta Z_t \sim I(d - b)$

## 3.2.2 Test de cointégration

Pour identifier l'existence d'un éventuel vecteur de cointégration, plusieurs méthodes peuvent être utilisées, notamment les approches proposées par : Engel – Granger(1987) [21], Johansen (1991) [31], et Phillips-Ouliaris (1990) [41]. Dans cette écriture, nous décrivons l'algorithme à deux pas proposé par Engel et Granger et celui de Johansen appliqué en cas de séries co-intégrées.

### Test de Engel et Granger

Ce test est construit par deux étapes :

**-Etape1** : tester l'ordre d'intégration des variables : une des conditions nécessaires pour qu'il y ait cointégration étant que les deux séries doivent être intégrées du même ordre. Ces tests reposent sur l'utilisation des tests de Dikey & Fuller.

On cherche alors  $d$  tel que  $X_t \sim I(d)$  et  $Y_t \sim I(d)$ .

**-Etape2** : Si  $X_t \sim I(1)$  et  $Y_t \sim I(1)$  : estime par les MCO la relation de long terme :

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

En réécrivant l'équation précédente comme suit, on a :

$$\varrho_t = Y_t - \alpha X_t - \beta t \sim I(0) \text{ pour } \alpha \neq 0$$

Ainsi, le test valide la présence de cointégration si le vecteur résiduel issu de cette estimation est stationnaire.

### 3.3 GARCH(p,q) multivarié

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

$X_t = (X_{1t}, X_{2t}, \dots, X_{nt})'$  et la matrice de covariance conditionnelle  $H_t$  du processus  $\{X_t\}_{t \in N}$  à la date  $t$  de terme général  $h_{ij,t}$  :

$$H_t = E_{t-1}(X_t \acute{X}_t).$$

Dans les modèles *GARCH* multivariés, on exprime cette matrice de covariance conditionnelle en fonction du passé du processus  $\{X_t\}_{t \in N}$  et des valeurs propres passées de  $H_t$ . Une première approche consiste à considérer un processus *GARCH* multivarié  $\{X_t\}_{t \in N}$  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 proces-

sus *GARCH* multivariés. Ainsi, différentes formulations tenant compte des corrélations entre les composantes d'un processus *GARCH* multivarié (ou *MVGARCH*) ont été introduites. De la même manière que dans le cas univarié, on dit que  $\{X_t\}_{t \in \mathbb{N}}$  est un processus *MVGARCH*( $p, q$ ) si :

$$X_t = H_t^{1/2} \varepsilon_t \quad (3.1)$$

où  $H_t = [h_{ij,t}]$  est la matrice de covariance conditionnelle de  $X_t$  de taille  $N \times N$  et  $\varepsilon_t$  un processus indépendant et identiquement distribué de moyenne nulle et tel que  $E(\varepsilon_t \varepsilon_t') = I$  où  $I$  est une matrice identité de dimension  $N$ .

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.

Le vecteur  $X_t$  vérifie les conditions suivantes sur les moments :

$$\begin{cases} E(X_t / \mathcal{F}_{t-1}) = 0 \\ E(X_t \dot{X}_t' / \mathcal{F}_{t-1}) = H_t \end{cases}$$

### 3.3.1 Modèles *GARCH*( $p, q$ ) Vectoriels

Le modèle vectoriel a été introduit par Bollerslev, Engle et Wooldridge (1988) [7], 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 dans le cas univarié :

$$\begin{aligned} Vech(H_t) &= W + \sum_{i=1}^q A_i Vech(X_{t-i} \dot{X}_{t-i}') + \sum_{j=1}^p B_j Vech(H_{t-j}), \quad t \in \mathbb{Z} \quad (3.2) \\ &= W + A(L) Vech(X_t \dot{X}_t') + B(L) Vech(H_t), \end{aligned}$$

où  $H_t$  est symétrique et  $Vech(H_t)$  est un vecteur de dimension  $\frac{N(N+1)}{2} \times 1$ ,  $W$  est un vecteur de dimension  $\frac{N(N+1)}{2} \times 1$ ,  $A_i, B_j$  sont des matrices de paramètres de dimension

$\frac{N(N+1)}{2} \times \frac{N(N+1)}{2}$ ,  $A(L) = A_1L + A_2L^2 + \dots + A_qL^q$  et  $B(L) = B_1L + B_2L^2 + \dots + B_pL^p$ .

L'opérateur *Vech* appliqué à 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,

Pour un modèle  $GARCH(p, q)$  multivarié à  $N$  équations, le nombre de paramètres à estimer est donné par la formule suivante :

$$\frac{N(N+1)}{2} + (p+q)\left(\frac{N(N+1)}{2}\right)^2.$$

### Modèles VEC-GARCH(1,1)

Dans ce modèle, la matrice de covariance conditionnelle est donnée par  $H_t = [h_{ij,t}]$ , telle que

$$h_{ij,t} = \omega_{ij} + \alpha_{ij}X_{i;t-1}X_{j;t-1} + \beta_{ij}h_{ij;t-1}, \quad i, j = 1, \dots, k. \quad (3.3)$$

### Modèles VEC-GARCH(1,1) bivarié

Partant d'abord d'un  $GARCH(1, 1)$  bivarié pour rendre compte de la problématique se rapportant au nombre de paramètres à estimer pour les  $VEC - GARCH(p, q)$  multivariés à  $N$  composantes. Soient  $r_{1,t}$  et  $r_{2,t}$  les rendements de deux actifs 1 et 2 obéissant aux processus suivants :

$$r_{1,t} = \mu_{1,t} + X_{1,t},$$

$$r_{2,t} = \mu_{2,t} + X_{2,t},$$

Dans le cas  $GARCH(1, 1)$ , où  $\omega_0 > 0, \alpha_1, \beta_1 \geq 0$ ,

On pose  $X_t = (X_{1,t} X_{2,t})$ , Alors la matrice de covariance conditionnelle  $H_t$  du processus  $\{X_t\}_{t \in N}$  à la date  $t$  de terme général  $h_{ij,t}$  :

$$H_t = E_{t-1}(X_{1,t}X_{2,t}') = E_{t-1} \begin{pmatrix} X_{1,t}^2 & X_{1,t}X_{2,t} \\ X_{2,t}X_{1,t} & X_{2,t}^2 \end{pmatrix} = \begin{pmatrix} h_{11,t} & h_{12,t} \\ h_{21,t} & h_{22,t} \end{pmatrix},$$

en utilisant l'opérateur *vech*(.), on obtient

$$\begin{aligned}
h_t &= \begin{pmatrix} h_{11,t} \\ h_{21,t} \\ h_{22,t} \end{pmatrix} = \begin{pmatrix} \omega_{11} \\ \omega_{21} \\ \omega_{22} \end{pmatrix} + \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} X_{1,t-1}^2 \\ X_{1,t-1}X_{2,t-1} \\ X_{2,t-1}^2 \end{pmatrix} \\
&+ \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix} \begin{pmatrix} h_{11,t-1} \\ h_{21,t-1} \\ h_{22,t-1} \end{pmatrix}.
\end{aligned}$$

Rappelons au travers que non contraint signifie que chaque élément de la matrice variance – covariance conditionnelle est généré par le même type de processus *GARCH*. Dans la littérature économétrique, on utilise souvent le nom *VECH* pour désigner cette représentation.

Elle nous permet d'extraire les variances et les covariances conditionnelles comme suit :

$$h_{11,t} = \omega_{11} + a_{11}X_{1,t-1}^2 + a_{12}X_{1,t-1}X_{2,t-1} + a_{13}X_{2,t-1}^2 + b_{11}h_{11,t-1} + b_{12}h_{21,t-1} + b_{13}h_{22,t-1}$$

$$h_{12,t} = \omega_{21} + a_{21}X_{1,t-1}^2 + a_{22}X_{1,t-1}X_{2,t-1} + a_{23}X_{2,t-1}^2 + b_{21}h_{11,t-1} + b_{22}h_{21,t-1} + b_{23}h_{22,t-1}$$

$$h_{22,t} = \omega_{22} + a_{31}X_{1,t-1}^2 + a_{32}X_{1,t-1}X_{2,t-1} + a_{33}X_{2,t-1}^2 + b_{31}h_{11,t-1} + b_{32}h_{21,t-1} + b_{33}h_{22,t-1}$$

Précisons que  $h_{11,t}$ , et  $h_{22,t}$ , ne sont que les variances conditionnelles de nos deux actifs 1 et 2 et enfin  $h_{12,t}$ , leur covariance conditionnelle.

Il est clair que ce processus *GARCH*(1, 1) bivarié non contraint nécessite l'estimation de 21 paramètres.

### Modèles **VEC-GARCH(p,q) Diagonal**

C'est aux travaux de Bollerslev, Engle et Wooldridge (1988) [7] que remonte l'un des plus anciens modèles contraints connus sous le nom de *GARCH* diagonal, La différence entre le modèle actuel et le précédent est dans le présent nous prendrons les deux matrices  $A_i$  et  $B_j$  diagonales.

Le nombre de paramètres à estimer est

$$(p + q + 1) \frac{N(N + 1)}{2}.$$

### Modèles VEC-GARCH(1,1) Diagonal

Dans ce modèle, les composantes de la matrice de covariance conditionnelle sont données par

$$h_{ij,t} = \omega_i + \alpha_{ii} X_{i;t-1} X_{j;t-1} + \beta_{ii} h_{ij;t-1}, \quad i, j = 1, \dots, k. \quad (3.4)$$

### Modèles VEC-GARCH(1,1) Diagonal bivarié

si on prenant  $p = q = 1$  et  $N = 2$ , alors elle nous permet d'extraire les variances et les covariances conditionnelles comme suit :

$$\begin{aligned} h_{11,t} &= \omega_{11} + a_{11} X_{1,t-1}^2 + b_{11} h_{11,t-1} \\ h_{12,t} &= \omega_{21} + a_{22} X_{1,t-1} X_{2,t-1} + b_{22} h_{21,t-1} \\ h_{22,t} &= \omega_{22} + a_{33} X_{2,t-1}^2 + b_{33} h_{22,t-1} \end{aligned}$$

Cette technique permet de contourner le problème du nombre de paramètres à estimer.

### Remarque 7

La représentation vectorielle (et donc a priori le modèle diagonal) telle qu'elle est définie dans la partie précédente ne tient pas compte de la nécessité que la matrice  $H_t$  est définie positivité. 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.

### 3.3.2 Modèle BEKK-GARCH

Le modèle *BEKK – GARCH* a été introduit par Engle et Kroner (1995) [19] pour imposer la restriction de la 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* est donnée par la relation suivante

$$H_t = \acute{C}C + \sum_{k=1}^K \sum_{i=1}^q \acute{A}_{ik} X_{t-i} \acute{X}_{t-i} A_{ik} + \sum_{k=1}^K \sum_{j=1}^p \acute{B}_{jk} H_{t-j} B_{jk}, \quad (3.5)$$

avec  $X_0$  et  $H_0$  fixés, de plus  $H_0$  est défini positif, 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  $n \times n$ . De plus  $C$  est la matrice triangulaire inférieure.

Le but de décomposer le terme constant en un produit de deux matrices triangulaires est d'avoir une matrice de covariance conditionnelle semi-définie positive  $H_t$  et  $\sqrt{H_t}$  est la racine carrée symétrique de  $H_t$ .

Le nombre de paramètres à estimer est :  $(p + q)kN^2 + \frac{N(N+1)}{2}$ .

Considérons le modèle *GARCH*(1, 1). L'expression du modèle est la suivante

$$H_t = \acute{C} C + \sum_{k=1}^K \acute{A}_{1k} X_{t-1} \acute{X}_{t-1} A_{1k} + \sum_{k=1}^K \acute{B}_{1k} H_{t-1} B_{1k}, \quad (3.6)$$

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

$$\begin{aligned} H_t = & \acute{C}C + \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}' \begin{pmatrix} X_{1,t-1}^2 & X_{1,t-1} X_{2,t-1} \\ X_{2,t-1} X_{1,t-1} & X_{2,t-1}^2 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \\ & + \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}' \begin{pmatrix} h_{11,t-1} & h_{12,t-1} \\ h_{21,t-1} & h_{22,t-1} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \end{aligned}$$

On remarque que ce modèle nécessite moins de paramètres que le modèle *GARCH*(1, 1)

proposé dans l'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 assurons l'identifiabilité et la généralité du modèle.

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

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

On voit que les conditions pour obtenir une matrice  $H_t$  définie positive quelque soit  $t \in N$  sont faibles : la condition portant sur les noyaux des matrices  $C$  et  $B_{jk}$   $j \in \{1, \dots, p\}$  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  $A'_{ik} X_{t-1} \hat{X}_{t-1} A_{ik}$  est semi-définie positive car  $X_{t-1} \hat{X}_{t-1}$  est semi-définie positive. Ainsi, pour assurer la définie positivité de  $H_t$  il suffit que

$$\acute{C}C + \sum_{k=1}^K \sum_{j=1}^p \acute{B}_{jk} H_{t-j} B_{jk}$$

soit définie positive.

Une conséquence de cette proposition est que pour le modèle *MVARCH*, la condition nécessaire pour que  $H_t$  soit définie positive,  $t \in N$ , devient  $\ker\{C\} = \{0\}$ , c'est-à-dire que  $C$  soit de plein rang.

### 3.3.3 Modèle CCC-GARCH

Les modèles *GARCH* multivariés à corrélations conditionnelles décomposent la matrice de corrélations en deux composantes, soit celle des écarts-types et celle des corrélations conditionnelles. Bollerslev (1990) [9] introduit le premier modèle de corrélations, le modèle à corrélations conditionnelles constantes (*CCC-GARCH*). Il propose un modèle

où les variances et covariances conditionnelles varient dans le temps et les corrélations conditionnelles restent constantes :

$$r_t = \mu_t + X_t \quad (3.7)$$

$$X_t = H_t^{1/2} \varepsilon_t \quad (3.8)$$

$$H_t = D_t R D_t \quad (3.9)$$

où

$r_t$  : vecteur  $n \times 1$  des rendements de  $n$  actifs au temps  $t$ ,

$\mu_t$  : vecteur  $n \times 1$  des espérances conditionnelles de  $r_t$ ,

$X_t$  : vecteur  $n \times 1$  des erreurs conditionnelles de  $n$  actifs au temps  $t$ , avec

$$E[X_t] = 0 \text{ et } Cov[X_t] = H_t,$$

$H_t$  : matrice  $n \times n$  des variances conditionnelles de  $\varepsilon_t$  au temps  $t$ ,

$D_t$  : matrice diagonale  $n \times n$  des écarts-types conditionnels de  $\varepsilon_t$  au temps  $t$ ,

$R$  : matrice  $n \times n$  des corrélations constantes,

$\varepsilon_t$  : vecteur  $n \times 1$  des erreurs i.i.d. avec

$$E[\varepsilon_t] = 0 \text{ et } E[\varepsilon_t \varepsilon_t'] = I_n$$

Les composantes de la matrice  $H_t = (h_{ij,t})$ , sont données par

$$\begin{cases} h_{ii,t} = \omega_i + \sum_{j=1}^q \alpha_{ij} X_{i,t-j}^2 + \sum_{j=1}^p \beta_{ij} h_{ii,t-j}, & i = 1, \dots, N \\ h_{ij,t} = \rho_{ij} \sqrt{h_{ii,t} h_{jj,t}}, & i \neq j. \end{cases} \quad (3.10)$$

et la matrice  $D_t$  est donnée par

$$D_t = \begin{pmatrix} \sqrt{h_{11,t}} & 0 & 0 & \dots & 0 \\ 0 & \sqrt{h_{22,t}} & 0 & \dots & 0 \\ 0 & 0 & \sqrt{h_{33,t}} & \dots & 0 \\ \dots & \dots & \dots & \dots & 0 \\ 0 & 0 & \dots & 0 & \sqrt{h_{NN,t}} \end{pmatrix}$$

et  $R$  est une matrice symétrique définie positive contenant les corrélations conditionnelles constantes  $\rho_{ij}$  telle que  $\rho_{ii} = 1, i = 1, \dots, N$ .

$$R = \begin{pmatrix} 1 & \rho_{12} & \rho_{13} & \dots & \rho_{1n} \\ \rho_{21} & 1 & \rho_{23} & \dots & \rho_{2n} \\ \rho_{31} & \rho_{32} & 1 & \dots & \dots \\ \dots & \dots & \dots & \ddots & \rho_{n-1n} \\ \rho_{n1} & \rho_{n2} & \dots & \rho_{nn-1} & 1 \end{pmatrix}$$

Les éléments de la matrice d'écart-types  $D_t$  sont obtenus à partir des résidus des *GARCH* univariés de chacune des variables à estimer et les éléments de la matrice de corrélations  $R$  ne varient pas dans le temps.

Le nombre de paramètres à estimer est :  $(p + q + 1)N + \frac{N(N-1)}{2}$ .

### 3.3.4 Modèle DCC-GARCH

Engle (2002) [22] introduit le modèle à corrélations conditionnelles dynamiques, le *DCC – GARCH*, en permettant à la matrice de corrélations conditionnelles de varier dans le temps. Ce modèle est une généralisation du modèle *CCC – GARCH*.

C'est un modèle d'estimation en deux étapes. La première étape consiste à estimer la variance conditionnelle avec un *GARCH* univarié pour chacune des séries. Dans la deuxième étape, on utilise ensuite les résidus standardisés obtenus dans la première

étape pour estimer les paramètres de la matrice des corrélations dynamiques. Ce modèle comporte des conditions permettant à la matrice de covariances d'être définie positive en tout temps ainsi qu'à la covariance d'être stationnaire. Analogiquement au modèle *CCC – GARCH*, la matrice  $H_t$  est divisée en deux matrices,  $D_t$  et  $R_t$ . Les éléments de la matrice  $D_t$  proviennent des GARCH univarié estimés pour chacune des séries :

$$H_t = D_t R_t D_t,$$

$H_t$  : matrice  $n \times n$  des variances conditionnelles de  $X_t$  au temps  $t$ ,

$D_t$  : matrice diagonale  $n \times n$  des écarts-types conditionnels de  $\varepsilon_t$  temps  $t$ ,

$R_t$  : est la matrice symétrique définie positive contenant les corrélations conditionnelles dynamiques  $\rho_{ij,t}$ .

La matrice  $D_t$  est donnée par

$$D_t = \begin{pmatrix} \sqrt{h_{1,t}} & 0 & 0 & \dots & 0 \\ 0 & \sqrt{h_{2,t}} & 0 & \dots & 0 \\ 0 & 0 & \sqrt{h_{3,t}} & \dots & 0 \\ \dots & \dots & \dots & \dots & 0 \\ 0 & 0 & \dots & 0 & \sqrt{h_{N,t}} \end{pmatrix},$$

Les éléments contenus dans  $D_t$  sont générés selon un processus GARCH (P, Q), qui peut être formulé comme suit

$$h_{it} = \alpha_{0,i} + \sum_{q=1}^{Q_i} \alpha_{iq} X_{i,t-q}^2 + \sum_{p=1}^{P_i} \beta_{ip} h_{i,t-p}, \text{ ou } \sum_{q=1}^{Q_i} \alpha_{iq} + \sum_{p=1}^{P_i} \beta_{ip} < 1, \quad (3.11)$$

La matrice  $R_t$  est celle des corrélations conditionnelles des résidus standardisés  $E_t$ , elle est maintenant dynamique :

$$R_t = \begin{pmatrix} 1 & \rho_{12,t} & \rho_{13,t} & \cdots & \rho_{1n,t} \\ \rho_{21,t} & 1 & \rho_{23} & \cdots & \rho_{2n,t} \\ \rho_{31,t} & \rho_{32,t} & 1 & & \cdots \\ \cdots & \cdots & & \ddots & \rho_{n-1,n,t} \\ \rho_{n1,t} & \rho_{n2,t} & \cdots & \rho_{n,n-1,t} & 1 \end{pmatrix},$$

La matrice  $H_t$  doit être définie positive, car elle est une matrice de variances-covariances. Afin d'assurer que  $H_t$  soit définie positive, il faut aussi que  $R_t$  soit définie positive, car  $H_t = D_t R_t D_t$ . La matrice  $D_t$  est toujours positive, car  $D_t = \text{diag}(\sqrt{h_{1,t}}, \dots, \sqrt{h_{N,t}})$ , donc ses éléments sont toujours positifs. Il faut aussi s'assurer que les éléments de  $R_t$  soient plus petits ou égaux à 1, car ce sont des corrélations. Afin de s'assurer de la positivité de  $R_t$ , ce dernier est décomposé en deux matrices :

$$R_t = Q_t^{*-1} Q_t Q_t^{*-1},$$

où

$$Q_t^* = \begin{bmatrix} \sqrt{q_{11,t}} & 0 & 0 & \cdots & 0 \\ 0 & \sqrt{q_{22,t}} & 0 & \cdots & 0 \\ 0 & 0 & \sqrt{q_{33,t}} & 0 & \\ \vdots & \vdots & & \ddots & 0 \\ 0 & 0 & \cdots & 0 & \sqrt{q_{NN,t}} \end{bmatrix},$$

et

$$Q_t = (1 - \alpha_{DCC} - \beta_{CCC})\bar{Q} + \alpha_{DCC}X_{t-1}X'_{t-1} + \beta_{DCC}Q_{t-1}, \quad (3.12)$$

avec

$$\bar{Q} = \text{cov}(X_t X'_t) = E(X_t X'_t),$$

$\bar{Q}$  soit la covariance non conditionnelle des résidus standardisés obtenue par les GARCH univariés,

alors nous pouvons écrire la matrice  $Q_t$  comme suit

$$Q_t = \begin{bmatrix} q_{11,t} & q_{12,t} & \cdots & q_{1N,t} \\ q_{21,t} & q_{22,t} & \cdots & q_{2N,t} \\ \vdots & \vdots & \ddots & \vdots \\ q_{N1,t} & q_{N2,t} & \cdots & q_{NN,t} \end{bmatrix},$$

La matrice  $Q_t$  doit être définie positive afin que  $R_t$  le soit aussi.

Notons que dans l'équation (3.13),  $\alpha_{DCC}$  et  $\beta_{DCC}$  sont des scalaires. Les conditions suivantes doivent être respectées afin d'assurer que  $H_t$  soit définie positive :

$$\alpha_{DCC} \geq 0, \beta_{DCC} \geq 0 \text{ and } \alpha_{DCC} + \beta_{DCC} < 1$$

La structure générale de corrélation dynamique  $DCC(p, q)$  est la suivante :

$$Q_t = \left( 1 - \sum_{i=1}^p \alpha_{DCC,i} - \sum_{j=1}^q \beta_{DCC,j} \right) \bar{Q} + \sum_{i=1}^p \alpha_{DCC,i} (X_{t-i} X'_{t-i}) + \sum_{j=1}^q \beta_{DCC,j} Q_{t-j}. \quad (3.13)$$

Les corrélations conditionnelles dynamiques  $\rho_{ij} = \frac{q_{ij,t}}{(q_{ii,t} q_{jj,t})^{1/2}}$ .

Le nombre de paramètres à estimer est :  $(p + q) + \frac{N(N+1)}{2}$ .

Cependant, dans cette recherche nous utiliserons un  $DCC - GARCH(1, 1)$ .

### Modèle DCC-GARCH (1,1)

Dans cette recherche, nous utilisons la forme la plus simple du modèle  $DCC - GARCH$  pour analyser les relations de transfert de volatilité entre le prix de pétrole et les différents marchés boursiers, soit le  $DCC - GARCH(1, 1)$  bivarié. La construction

du modèle  $DCC - GARCH(1, 1)$  est :

$$\begin{aligned} h_{11,t} &= \alpha_{0,1} + \alpha_{11}X_{1,t-1}^2 + \beta_{11}h_{11,t-1} \\ h_{22,t} &= \alpha_{0,2} + \alpha_{21}X_{2,t-1}^2 + \beta_{21}h_{22,t-1} \\ R_t &= Q_t^{*-1}Q_tQ_t^{*-1} \\ Q_t &= (1 - \alpha_{DCC} - \beta_{DCC})\bar{Q} + \alpha_{DCC}X_{t-1}X'_{t-1} + \beta_{DCC}Q_{t-1} \end{aligned}$$

avec

$$H_t = \begin{bmatrix} h_{11,t} & h_{12,t} \\ h_{21,t} & h_{22,t} \end{bmatrix}.$$

Les avantages du modèle  $DCC - GARCH$  sont la modélisation directe de la variance et de la covariance et ainsi que sa flexibilité. Cependant il a aussi des faiblesses, la fonction de vraisemblance devient compliquée quand le nombre de variables est plus grand ou égal à 3 et la matrice de corrélations conditionnelles doit être définie positive pour tout  $t$ . Dans notre recherche, le nombre de variables se limite à 2 et le logiciel utilisé s'assure d'obtenir une matrice  $H_t$  définie positive.

Ce modèle permettra d'examiner les transferts de volatilité entre le le prix de pétrole et les différents marchés boursiers.

### 3.4 La stationnarité du covariance.

Finalement, on a donné dans la proposition suivante des conditions nécessaires et suffisantes pour que la covariance du processus GARCH multivarié soit stationnaire.

#### **Proposition**

On suppose que le processus  $\{X_t\}$  commence arbitrairement loin dans le passé, à l'instant  $t - \tau$  et on suppose aussi que  $\lim_{\tau \rightarrow \infty} H_{t-\tau}$  est finie et définie positive,

Si le modèle  $Vech - GARCH$  est défini positif, alors sa covariance est stationnaire si et seulement si toutes les valeurs propres de

$$\sum_{i=1}^q A_i + \sum_{j=1}^p B_j \text{ sont strictement inférieures à 1 en valeur absolue.}$$

De plus, le modèle *BEKK – GARCH* est de covariance stationnaire si et seulement si toutes les valeurs propres de

$$\sum_{i=1}^q \sum_{k=1}^K (A_{ik} \otimes A_{ik}) + \sum_{j=1}^p \sum_{k=1}^K (B_{jk} \otimes B_{jk}) \text{ sont strictements inférieurs à un en valeur absolue,}$$

où  $\otimes$  désigne le produit de Kronecker de deux matrices.

Dans le modèle *CCC – GARCH*, les conditions habituelles pour assurer la positivité des variances et la stationnarité sont

$$\omega_i > 0, \alpha_{ij}, \beta_{ij} > 0 \text{ et } \sum_{j=1}^q \alpha_{ij} + \sum_{j=1}^p \beta_{ij} < 1.$$

Les conditions nécessaires pour que le modèle *DCC – GARCH* soit stationnaire sont les suivantes

$$0 \leq \alpha_{DCC,i}, \beta_{DCC,j} \leq 1, \text{ et } \sum_{i=1}^p \alpha_{DCC,i} + \sum_{j=1}^Q \beta_{DCC,j} < 1.$$

# Chapitre 4

## Application

Nous avons vu dans les trois premiers chapitres les formulations des modèles avec hétéroscédasticité conditionnelle, les méthodes d'estimation utilisées en pratique. Dans ce chapitre, nous nous proposons d'illustrer certaines de ces notions.

### 4.1 Résultats

Les principaux résultats de cet article sont la modélisation et l'application en utilisant le logiciel EViews. En appliquant aux données du taux de change des prix du pétrole et aux trois séries de rendements boursiers qui sont CAC 40, S&P 500 et NIKKEI 225. Les données représentent les prix du pétrole mensuels wti, l'étude se rapporte à la période d'août 1987 à octobre 2016, Les données proviennent de *investing.com*.

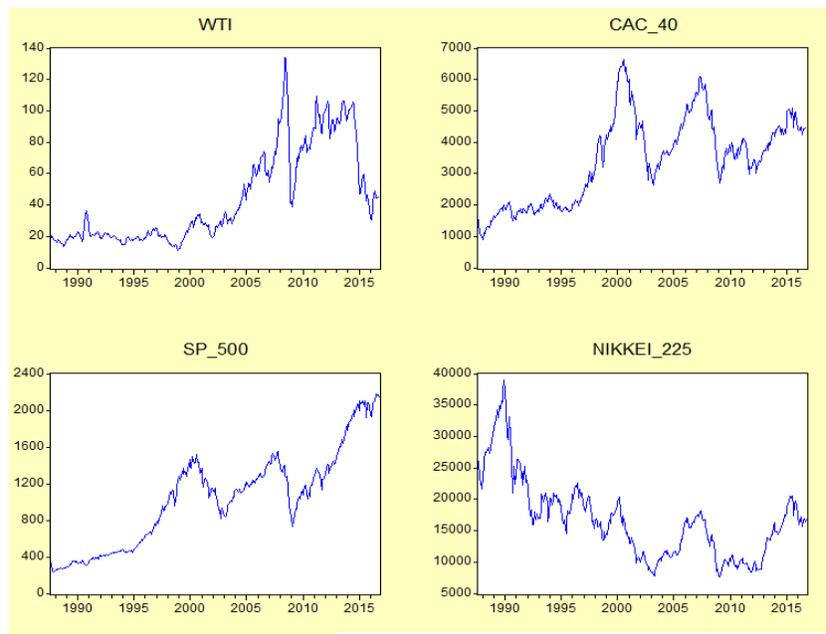
### 4.2 Les données et leurs propriétés statistiques

Dans notre application, nous étudierons d'abord ses différentes séries de données graphiquement (analyse descriptive). Par la suite nous nous interrogeons sur la question de la stationnarité des données, nous étudierons alors la stationnarité à l'aide des tests appropriés. Puis, nous envisagerons un traitement (sous forme de transformation) de ces

données afin de les stationnariser, s'il s'avère qu'elles ne le sont pas.

Enfin nous prouvons, successivement, par le calcul des différents Kurtosis, Skewness et les P-values du test de Jarque-Bera, que le processus est non gaussien et que sa variance est hétéroscédastique.

La figure 1 représente l'évolution temporelle des prix du pétrole et comportement du marché boursier au fil du temps.



Prix du ptrole et comportement du march boursier au fil du temps

La lecture visuelle de ces graphes indique une non-stationnarité en moyenne et en variance (surtout en variance). Ce même graphe montre des regroupements de volatilité, ce qui signifie que ces séries sont volatiles. Nous remarquons, par ailleurs, que cette volatilité évolue au cours du temps.

### 4.3 l'étude de la stationnarité des séries.

Pour cela, on a utilisé le test de racine unitaire (test de non-stationnarité) de Dickey Fuller (ADF) et de Philips Péron (PP).

#### **Test de Dickey et Fuller augmenté(ADF)**

Un test de non-stationnarité largement utilisé et répandu est le test de racine unitaire proposé par Dickey et Fuller en 1979.

$H_0$  : La série comporte une racine unitaire (ou de non-stationnarité),

$H_1$  : La série ne comporte pas de racine unitaire, la série est stationnaire.

Étant donné que le p-valu calculé est inférieur au niveau de signification  $\alpha = 1\%$ , on doit rejeter l'hypothèse nulle  $H_0$  de la non-stationnarité, et retenir l'hypothèse alternative  $H_1$ .

#### **Tests de Phillips et Perron(PP)**

Phillips et Perron (1987, 1988) et Phillips (1987) proposent un autre test pour détecter la non-stationnarité d'une série temporelle. Ce test est une adaptation non paramétrique du test de Dickey et Fuller. L'hypothèse nulle du test est, comme pour le test DF, la présence d'une racine unitaire. La présentation détaillée de ce test se retrouve dans la plupart des manuels consacrés à l'économétrie des séries temporelles (Bresson et Pirotte, 1995).

$H_0$  : La série comporte une racine unitaire,

$H_1$  : La série ne comporte pas de racine unitaire, la série est stationnaire.

Étant donné que le p-valu calculé est inférieur au niveau de signification  $\alpha = 1\%$ , on doit rejeter l'hypothèse nulle  $H_0$  de non-stationnarité, et retenir l'hypothèse alternative  $H_1$ .

Le tableau 1 représente le test de stationnarité (Test de racine unitaire)

Series	Test	ADF	P-valu	PP	P-valu
WTI	wti	-2.158718	0.2221	-1.468379	0.5487
	Logwti	-1.754289	0.4030	-1.462268	0.5517
	DLogwti	-13.503220*** <sup>1</sup>	0.0000	-12.926670***	0.0000
	Dwti	-12.201670***	0.0000	-11.864940***	0.0000
CAC_40	cac_40	-1.576910	0.4933	-1.767885	0.3962
	Logcac_40	-1.630469	0.4659	-1.702235	0.4293
	DLogcac_40	-17.347770***	0.0000	-17.357700***	0.0000
	Dcac_40	-17.000150***	0.0000	-17.129590***	0.0000
SP_500	sp_500	-0.009170	0.9562	-0.154496	0.9410
	Logsp_500	-0.894766	0.7894	-0.930747	0.7778
	DLogsp_500	-17.205800***	0.0000	-17.198700***	0.0000
	Dsp_500	-18.117990***	0.0000	-18.172110***	0.0000
NIKKEI_225	nikkei_225	-1.952964	0.3078	-2.010034	0.2825
	Lognikkei_225	-1.879445	0.3419	-1.988244	0.2921
	DLognikkei_225	-17.648070***	0.0000	-17.676180***	0.0000
	Dnikkei_225	-18.003820***	0.0000	-17.996750***	0.0000

**Table 1.** Test de racine unitaire  $\alpha = 1\%$ (-3.448889)

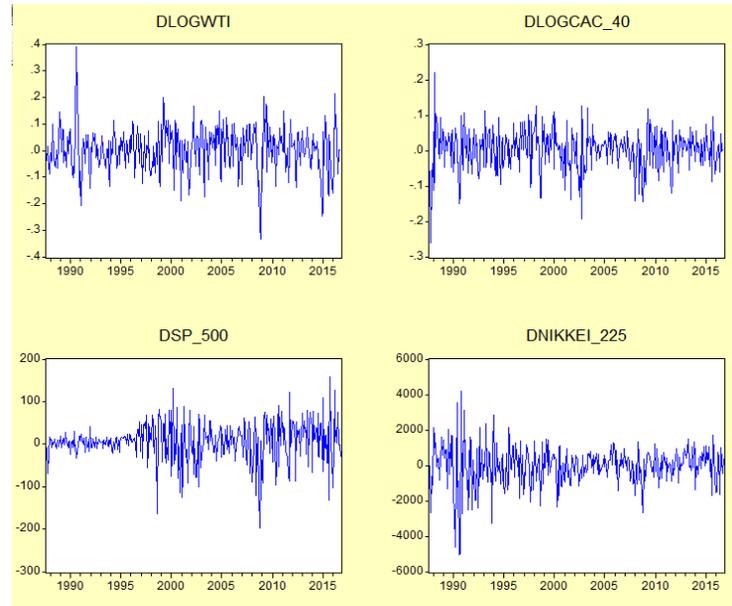
Le test de racine unitaire confirme la stationnarité des séries.

À partir du tableau 1, on va utiliser dans ce qui suit les séries suivantes : Dlogwti, Dlogcac\_40, Dsp\_500, DNikkei\_225 car elles sont stationnaires.

---

<sup>1</sup>\*\*\* dénote les statistiques significatives au seuil de 1%.

Les graphes associés aux séries stationnaires sont présentés dans la figure 2



Evolution temporelle des titres

Tester la présence d'effet  $ARCH(q)$  consiste à faire une régression simple par les moindres carrés ordinaires de  $\hat{u}_t^2$ , sur  $\hat{u}_t^2, \hat{u}_{t-1}^2, \dots$  et à considérer la statistique  $LM = nR^2$ , où  $R^2$  et  $n$  sont respectivement coefficient de corrélation multiple et le nombre d'observations. Sous l'hypothèse nulle, cette statistique suit asymptotiquement une loi de Khi-deux avec autant de degrés de liberté que le nombre de retards associé à  $u_t$ . La régression a été faite pour plusieurs degrés de ARCH. Le tableau 2 présente la statistique de test, ainsi que les P-values correspondantes.

	DLogwti	DLogcac_40	Dsp_500	Dnikkei_225
nR <sup>2</sup>	20.73225	51.5284	561.78815	51.0783
P-valus	0.0000	0.0000	0.0000	0.0000

**Table 2.** Statistiques du test d'homoscédasticité

Il est clair que toutes les P-values ne dépassent en aucun cas les 1%, ce qui nous conduit à rejeter l'hypothèse nulle d'homoscédasticité, et à accepter l'hypothèse de l'hétéroscédasticité.

### Statistiques descriptives

Le tableau 3 présente des statistiques sommaires des quatre séries.

	Dlogwti	Dlogcac_40	Dsp_500	DNikkei_225
Mean	0.002273	0.003051	5.162657	-26.25429
Median	0.009274	0.009999	7.555000	35.02500
Std.Dev	0.084150	0.057653	45.45415	1105.5990
Skewness	-0.274326	-0.520575	-0.567477	-0.663243
Kurtosis	5.208957	4.460995	5.054818	6.346034
Jarque-Bera	75.54910	46.93647	80.35997	188.9345
P-valus	0.000000	0.000000	0.000000	0.000000

**Table 3.** Statistiques descriptives pour chaque série stationnaire

Le tableau 3, montre que le rendement moyen le plus faible est attribué à l'indice boursier Nikkei\_225 avec une moyenne de  $-26.25$ ; alors que le plus remarquable est enregistré en Dsp\_500 avec une moyenne de 5.16. En terme de risque, l'indice boursier Cac\_40 possède le risque le moins élevé avec un écart type de 0.05 alors que celui le plus élevé est observé à l'indice boursier Nikkei\_225 avec un écart-type de 1105.59. À propos du marché du pétrole, il a connu un rendement moyen positif, ce qui n'est pas surprenant compte tenu de la tendance à la hausse du prix du pétrole au cours de la dernière décennie.

D'un autre côté, on peut voir que les coefficients de l'asymétrie (Skewness) de DLogwti, DLogcac\_40, Dsp\_500 et Dnikkei\_225 est  $-0.274326$ ,  $-0.520575$ ,  $-0.567477$  et  $-0.663243$  respectivement, donc ils sont négatifs, ce qui illustre bien le fait qu'un choc négatif a plus d'impact qu'un choc positif, de plus la moyenne est inférieur à la médiane, et il y a une

distribution asymétrique à gauche.

Le coefficient d'aplatissement (Kurtosis) de DLogwti, DLogcac\_40, Dsp\_500 et Dnikkei\_225 est 5.208957, 4.460995, 5.054818 et 6.346034 respectivement supérieur à 3, dans ce cas la distribution est pointue (distribution leptokurtique) et cela indique aussi des queues plus grosses que la distribution normale avec un pic supérieur.

Les valeurs de Jarque-Bera de DLogwti, DLogcac\_40, Dsp\_500 et Dnikkei\_225 est 75.54910, 46.93647, 80.35997 et 188.9345 respectivement et on voit que les différentes valeurs de la P-value du test de Jarque-Bera sont nulles, ce qui conduit à rejeter l'hypothèse de normalité des résidus.

## 4.4 Estimation de modèle GARCH(1,1) Univarié

Le tableau 4 représente le résultat d'estimation du modèle  $GARCH(1, 1)$  univarié

Series	Estimate	Coefficient	Std.Error	Z-Statistic	prob
DLogwti	$\omega$	0.004485	0.004069	1.102050	0.270400
	$\omega$	0.001424**	0.000679	2.095656	0.036100
	$\alpha$	0.232790*** <sup>2</sup>	0.063362	3.673971	0.000200
	$\beta$	0.565363***	0.138289	4.088261	0.000000
DLogcac_40	$\omega$	0.006516**	0.002953	2.206805	0.027300
	$\omega$	0.000395*** <sup>3</sup>	0.000166	2.380379	0.017300
	$\alpha$	0.188655***	0.066264	2.847028	0.004400
	$\beta$	0.677769***	0.098287	6.895826	0.000000
Dsp_500	$\omega$	3.945604***	1.217499	3.240746	0.001200
	$\omega$	8.357726	6.708267	1.245885	0.212800
	$\alpha$	0.201430***	0.052514	3.835738	0.000100
	$\beta$	0.820619***	0.041103	19.96483	0.000000
Dnikkei_225	$\omega$	7.307306	45.25027	0.161486	0.871700
	$\omega$	21700.750* <sup>4</sup>	13237.98	1.639280	0.101200
	$\alpha$	0.122979***	0.029467	4.173457	0.000000
	$\beta$	0.856890***	0.030310	28.27050	0.000000

**Table 4.** Résultats d'estimation du modèle GARCH(1,1) univarié

Pour le pétrole brut et pour chacun des trois indices boursiers, la plupart de  $\alpha, \beta$  sont statistiquement significatives.

<sup>2</sup>\*\*\* dénote les statistiques significatives au seuil de 1%.

<sup>3</sup>\*\* dénote les statistiques significatives au seuil de 5%.

<sup>4</sup>\* dénote les statistiques significatives au seuil de 10%.

L'estimation du modèle  $GARCH(1, 1)$  de la volatilité conditionnelle des différentes séries nous donne les résultats suivants :

$$D \log WTI_t = X_t,$$

$$\sigma_t^2 = 0.001424 + 0.232790X_{t-1}^2 + 0.565363\sigma_{t-1}^2.$$

$$D \log CAC\_40_t = 0.006516 + X_t,$$

$$\sigma_t^2 = 0.000395 + 0.188655X_{t-1}^2 + 0.677769\sigma_{t-1}^2.$$

$$DSP\_500_t = 3.945604 + X_t,$$

$$\sigma_t^2 = 8.357726 + 0.201430X_{t-1}^2 + 0.820619\sigma_{t-1}^2.$$

$$DNikkei\_225_t = X_t,$$

$$\sigma_t^2 = 21700.750 + 0.122979X_{t-1}^2 + 0.856890\sigma_{t-1}^2.$$

Au regard des résultats obtenus, le modèle  $GARCH(1, 1)$  semble adéquat pour modéliser la volatilité conditionnelle des quatre séries, car tous les coefficients de l'équation de la variance sont significatifs.

Series	AdjR2	SEE	BIC	RMSE	MAE	MAPE
Dlogwti	1,00000	2.47E-12	-2.287103	0.83E-12	0.61E-13	11.62E-10
Dlogcac_40	1,00000	1.28E-10	-3.653820	0.15E-11	0.24E-12	2.85E-8
Dsp_500	1,00000	6.93E-10	-9.041710	5.23E-10	4.74E-10	1.19E-8
Dnikkei_225	1,00000	3.15E-08	-13.43601	6.52E-7	8.34E-8	3.11E-7

**Table 5.** Quelques métriques d'ajustement  $GARCH(1,1)$  pour les différentes variables stationnaires.

## Discussion

Les modèles  $GARCH$  non linéaires, qui sont capables de capturer la volatilité à longue mémoire et /ou asymétrique, présentent une plus grande précision de prévision que les modèles linéaires, en particulier dans la prévision de volatilité sur des horizons temporels plus longs, tels que cinq ou vingt jours. Nos résultats indiquent que le degré de persistance

de la volatilité a été réduit en intégrant les changements de variance dans le modèle de volatilité.

De plus, on peut voir que les modèles GARCH expliquent une partie de la non-normalité de la distribution des prix de l'énergie. Quand ils le font, le terme d'erreur présente encore l'asymétrie et la leptokurtose. À des niveaux de confiance plus élevés, les estimations de VAR de l'énergie basées sur les modèles GARCH normales sont légèrement meilleures que celles habituellement utilisées par les sociétés d'énergie. Tenir compte de la distribution non-gaussienne des rendements énergétiques et des changements de la volatilité, en utilisant le modèle GARCH stable.

## **4.5 Estimation de modèle GARCH(1,1) Bivarié**

### **4.5.1 BEKK-GARCH(1,1) Bivarié**

GARCH(P=1,Q=1,MV=BEKK,DIST=GED)

Dlogwti-Dlogcac_40				Dlogwti-Dsp_500		
Estimate	coefficient	T-stat	Signif	coefficient	T-stat	Signif
C(1,1)	0.067089***	8.374270	0.000000	0.022404***	4.002340	0.000062
C(2,1)	0.009988* <sup>7</sup>	1.822840	0.068328	-6.318450**	-1.894650	0.058139
C(2,2)	0.011716	0.827540	0.407929	0.000001	1.71e-008	0.999999
A(1,1)	0.672141***	9.313670	0.000000	0.397405***	6.977120	0.000000
A(1,2)	-0.050247	-0.878850	0.379480	-41.387997	-1.340760	0.179998
A(2,1)	0.041396	0.511260	0.609166	0.000146	1.054080	0.291847
A(2,2)	0.630759***	8.288900	0.000000	0.499761***	11.01996	0.000000
B(1,1)	0.355319*	1.781810	0.074779	0.781639***	24.86544	0.000000
B(1,2)	-0.220792** <sup>7</sup>	-2.041210	0.041229	265.71584***	14.23243	0.000000
B(2,1)	-0.093463	-0.661930	0.508019	-0.001009***	-12.50726	0.000000
B(2,2)	0.720523*** <sup>7</sup>	10.456380	0.000000	0.683105***	19.00782	0.000000

Estimate	coefficient	T-stat	Signif
C(1,1)	0.069010***	12.071910	0.000000
C(2,1)	330.59121***	4.626390	0.000003
C(2,2)	0.000022	1.68e-007	0.999999
A(1,1)	0.497504***	8.171140	0.000000
A(1,2)	294.33651	0.497970	0.618506
A(2,1)	0.000016***	4.129530	0.000036
A(2,2)	0.456446***	12.16280	0.000000
B(1,1)	0.394015***	3.805410	0.000141
B(1,2)	-100.90630	-0.132590	0.894514
B(2,1)	-0.000025***	-6.057940	0.000000
B(2,2)	0.860288***	32.43009	0.000000

Table.6 Résultats BEEK-GARCH(1,1) bivarié

<sup>5</sup>\* dénote les statistiques significatives au seuil de 10%.

<sup>6</sup>\*\* dénote les statistiques significatives au seuil de 5%.

<sup>7</sup>\*\*\* dénote les statistiques significatives au seuil de 1%.

À partir du tableau 6, la plupart des variables estimées sont statistiquement significatives.

L'estimation du modèle *BEKK – GARCH(1,1)* peut être obtenue ci-dessous

– **Dlogwti – Dlogcac\_40**

$$A = \begin{bmatrix} 0.672141 & 0 \\ 0 & 0.630759 \end{bmatrix}, \quad B = \begin{bmatrix} 0.355319 & -0.220792 \\ 0 & 0.720523 \end{bmatrix}, \quad C = \begin{bmatrix} 0.067089 & 0 \\ 0.009988 & 0 \end{bmatrix}.$$

– **Dlogwti – Dsp\_500**

$$A = \begin{bmatrix} 0.397405 & 0 \\ 0 & 0.499761 \end{bmatrix}, \quad B = \begin{bmatrix} 0.781639 & 265.71584 \\ -0.001009 & 0.683105 \end{bmatrix}, \quad C = \begin{bmatrix} 0.022404 & 0 \\ -6.318450 & 0 \end{bmatrix}.$$

– **Dlogwti – Dnikkei\_225**

$$A = \begin{bmatrix} 0.497504 & 0 \\ 0.000016 & 0.456446 \end{bmatrix}, \quad B = \begin{bmatrix} 0.394015 & 0 \\ -0.000025 & 0.860288 \end{bmatrix}, \quad C = \begin{bmatrix} 0.069010 & 0 \\ 330.59121 & 0 \end{bmatrix}.$$

## 4.5.2 DCC-GARCH(1,1)Bivarié

GARCH(P=1,Q=1,MV=DCC,DIST=GED)

Estimate	Dlogwti-Dlogcac_40			Dlogwti-Dsp_500		
	coefficient	T-stat	Signif	coefficient	T-stat	Signif
C(1)	0.009192*** <sup>8</sup>	72.66540	0.000000	0.063723***	281.99049	0.000000
C(2)	-0.000003***	-3.68341	0.000230	275.03683***	561.96769	0.000000
A(1)	0.821568***	42.05718	0.000000	0.306941***	783.82073	0.000000
A(2)	2.037555***	86.37950	0.000000	0.500475***	1020.11591	0.000000
B(1)	0.003781***	179.22273	0.000000	-0.904574***	-254.60201	0.000000
B(2)	0.004930***	201.33666	0.000000	0.716348***	266.54864	0.000000
DCC(A)	0.020402***	5.06341	0.000000	0.022167***	628.56565	0.000000
DCC(B)	0.902565***	4.83036	0.000001	0.917991***	1054.18464	0.000000

Dlogwti-Dnikkei\_225

Estimate	coefficient	T-stat	Signif
C(1)	0.006897***	18.52578	0.000000
C(2)	949660.80***	14.46775	0.000000
A(1)	0.383113***	12.69728	0.000000
A(2)	0.332151***	12.67729	0.000000
B(1)	0.875396***	98.69942	0.000000
B(2)	0.745604***	24.50621	0.000000
DCC(A)	0.304008***	15.07005	0.000000
DCC(B)	0.628507***	42.15344	0.000000

Table7. Résultats DCC-GARCH(1,1) bivarié

Les résultats de l'estimation du modèle  $DCC - GARCH(1,1)$  sont présentés au tableau 7, les coefficients de tous les paramètres sont significatifs, ce qui indique qu'il existe un effet de débordement mutuel entre le pétrole brut et les trois marchés boursiers.

<sup>8</sup>\*\*\* dénote les statistiques significatives au seuil de 1%.

Ensuite, le modèle DCC estimé est le suivant,

– **Dlogwti – Dlogcac\_40**

$$h_{11,t} = 0.009192 + 0.821568\varepsilon_{1,t-1}^2 + 0.003781h_{11,t-1}$$

$$h_{22,t} = -0.000003 + 2.037555\varepsilon_{2,t-1}^2 + 0.004930h_{22,t-1}$$

$$Q_t = (1 - 0.020402 - 0.902565)\bar{Q} + 0.020402\varepsilon_{t-1}\varepsilon'_{t-1} + 0.902565Q_{t-1}$$

– **Dlogwti – Dsp\_500**

$$h_{11,t} = 0.063723 + 0.306941\varepsilon_{1,t-1}^2 - 0.904574h_{11,t-1}$$

$$h_{22,t} = 275.03683 + 0.500475\varepsilon_{2,t-1}^2 + 0.716348h_{22,t-1}$$

$$Q_t = (1 - 0.022167 - 0.917991)\bar{Q} + 0.022167\varepsilon_{t-1}\varepsilon'_{t-1} + 0.917991Q_{t-1}$$

– **Dlogwti – Dnikkei\_225**

$$h_{11,t} = 0.006897 + 0.383113\varepsilon_{1,t-1}^2 + 0.875396h_{11,t-1}$$

$$h_{22,t} = 949660.80 + 0.332151\varepsilon_{2,t-1}^2 + 0.745604h_{22,t-1}$$

$$Q_t = (1 - 0.304008 - 0.628507)\bar{Q} + 0.304008\varepsilon_{t-1}\varepsilon'_{t-1} + 0.628507Q_{t-1}$$

# Conclusion et Perspectives

Nous avons vu les limites des modèles habituels face aux événements imprévisibles en finance; sans faire l'erreur de vouloir tout prévoir, nous sommes capables, avec les meilleurs modèles GARCH, d'affiner notre compréhension des risques pour mieux les anticiper. Des modèles récents permettent d'aller encore plus loin. Par exemple, il est probable que la corrélation ait, elle aussi, une réaction plus importante aux chocs négatifs communs qu'aux chocs positifs communs. Des modèles tels que celui de Cappiello, Engle et Sheppard (2006), AG-DCC, permettent de prendre cela en compte. D'autres modèles reconnaissent l'importance de la macroéconomie comme moteur sous-jacent des fluctuations financières.

La présente étude examine les corrélations dynamiques entre les prix des actions à savoir SP500, NIKKEI225, CAC40 et le pétrole brut (WTI). Plus précisément, nous utilisons un modèle GARCH (1,1) univarié et les modèles BEKK GARCH (1,1), DCC GARCH (1,1) bivariés, durant la période d'août 1987 à octobre 2016, en concentrant sur les corrélations conditionnelles dynamiques estimées entre les indices boursiers et le prix du pétrole. Cette approche permet d'étudier les dynamiques des moments de second ordre des prix des indices boursiers et le pétrole brut en prenant en compte le comportement de dépendance à long terme, les asymétries et l'effet de levier.

Les modèles BEKK GARCH (1, 1), DCC GARCH (1, 1) sont identifiés comme étant la meilleure spécification et la meilleure flexibilité pour modéliser l'hétéroscédasticité conditionnelle de séries temporelles individuelles. Nous avons ensuite étendu les modèles GARCH univariés ci-dessus à un cadre bivarié avec une paramétrisation de corrélation conditionnelle dynamique afin d'étudier l'interaction entre les prix du pétrole et les indices boursiers.

Les résultats ci-dessus pourraient être importants pour mieux comprendre la relation entre les prix du pétrole et les indices boursiers et pourraient être utiles aux

investisseurs et autres participants du marché tels que les gestionnaires financiers, les analystes et les entreprises pour gérer leurs investissements et minimiser leurs risques.

Dans les perspectives d'avenir, ce travail peut généraliser sur les données de gaz naturel et d'électricité. En outre, nous pouvons utiliser d'autres modèles GARCH multivariés.

# Bibliographie

- [1] Adouka, L. Boucheta, Y. Chenini A,E. and Kchirid E, M. GARCH models : Application to the Algerian Exchange Rate. *International Journal of Innovation and Applied Studies*. ISSN 2028-9324 Vol. 11, pp. 699-715. (2015).
- [2] Arouri, M. E. H., & Nguyen, D. K. Oil prices, stock markets and portfolio investment : Evidence from sector analysis in Europe over the last decade. *Energy Policy*, 38,4528-4539 (2010).
- [3] Arouri, M. E. H., & Rault, C. Oil prices and stock markets in GCC countries : Empirical evidence from panel analysis. *International Journal of Finance and Economics*. doi :10.1002/ijfe.443 (2011).
- [4] Baillie R.T., Bollerslev, T., and Mikkelsen, H.O. Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* 74, 3-30 (1996).
- [5] Bera, A. K., et Higgins, M. L. , "ARCH Models : Properties, Estimation and Testing", *Journal of Economic Surveys*, 7(4), 307-366.(1993).
- [6] Berkes I., Horváth L., & Kokoszka P. (2003) GARCH processes : Structure and estimation. *Bernoulli* 9, 201–227.
- [7] Bollerslev, T. Engle R.F and Wooldridge, J. A Capital Asset Pricing Model with Time-Varying Covariances. *Journal of Political Economy*,1988, vol.96, issue 1, 116-31.

- [8] Bollerslev, T. Generalized Autoregressive Conditional Heteroskedasticity, *Journal of Econometrics*, 31, 307-327, (1986).
- [9] Bollerslev, T. Modeling the Coherence in Short-run Nominal Exchange Rates : A Multivariate Generalized ARCH Model, (1990) *Review of Economics and Statistics*, Vol. 72, pp. 498-505.
- [10] Boubacar Mainassara, Y. et Francq, C. (2009), 'Estimating structural VARMA models with uncorrelated but non-independent error terms'. Document de travail, Université Lille III, France.
- [11] Bougerol, Ph and Picard, N. Strict Stationarity of Generalized Autoregressive Processes. *The Annals of Probability* Vol. 20, No. 4 (Oct., 1992), pp. 1714-1730.
- [12] Bühlmann, P. and McNeil, A.J. An algorithm for nonparametric GARCH modelling. *Journal of Computational Statistics & Data Analysis* 40, 665-683 (2002).
- [13] Brooks, C. and S. P. Burke. Forecasting exchange rate volatility using conditional variance models selected by information criteria. *Economics Letters* 61 : 273-278 (1998).
- [14] C. FRANCO, and J.M. ZAKOÏAN "Modèles GARCH et à volatilité stochastique", GREMARS-EQUIPPE Université Lille 3, (14 Dec 2009).
- [15] Cliff, A. D. et Ord, J. (1981), *Spatial Processes, Models and Applications*. Methuen : New York.
- [16] Cobbaut, R. , *Théorie Financière, Economica*. [5] Colletaz G. "Condhet.src : Estimation de Modèles ARCH sous RATS", Document de Recherche LEO, 2002-24.
- [17] Davidson, R. (2003). "Maximum Likelihood Estimation by Artificial Regression," Working paper.
- [18] El-Sharif I, Brown D, Burton B, Nixon B, Russell A. Evidence on the nature and extent of the relationship between oil prices and equity values in the UK. *Energy Econom*;27 :819-830 (2005).

- [19] Engle, R and Kroner, K. F. Multivariate Simultaneous Generalized ARCH. *Econometric Theory*, 1995, vol. 11, issue 01, 122-150.
- [20] Engle, R. "Autoregressive Conditional Heteroskedasticity with Estimates of United Kingdom Inflation", *Econometrica*, 50, 987-1008 (1982).
- [21] Engle, R. F and Granger, C W J. Co-integration and Error Correction : Representation, Estimation, and Testing. *Econometrica*, Vol. 55, No. 2. (Mar. 1987), pp. 251-276.
- [22] Engle, R. F. Dynamic Conditional Correlation : A Simple Class of Multivariate Generalized Autoregressive Conditional Heteroskedasticity Models, *Journal of Business & Economic Statistics* (2002), Vol. 20, n° 3, pp. 339-350, July.
- [23] Engle, R. & Sokalska, M.E. Forecasting intraday volatility in the US equity market. Multiplicative Component GARCH. *Journal of Financial Econometrics* 10(1), 54–83 (2012).
- [24] Ezzebsa, A. Modélisation des Marchés Financiers et Krachs Boursiers. Thèse de doctorat (2014).
- [25] Guerouah A, Zeghdoudi H, Bouseba F.Z, On Univariate and Multivariate GARCH Models : Oil Price and Stock market returns volatilities. *International Journal of Applied Mathematics and Statistics*<sup>TM</sup>. Vol. 57, No. 1 (2017).
- [26] Goldsheid I.Y. Lyapounov exponents and asymptotic behavior of the product of random matrices.(1991), pp. 23–37. *Lecture Notes in Mathematics* 1486. Springer-Verlag.
- [27] Granger, C.W.J. and Newbold P., (1974) « Spurious regressions in econometrics », *Journal of Econometrics*, Elsevier, II-2, p. 111-120.
- [28] Granger, C.W.J. « Investigating Causal Relations by Econometric Models and Cross-Spectral Methods », (1969), *Econometrica*. vol. 37 no.3, pp. 424-438.
- [29] Huang, R.D, Masulis, R. W, Stoll, H.R Energy stocks and financial markets. *Journal of futures markets*, vol 16(1), p(1-27) (1996).

- [30] Ji and Fan, How does oil price volatility effect non-energy markets ? *Applied Energy* 89, 273-280 (2012).
- [31] Johansen, S. Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models. 1991, *Econometrica*, Vol. 59, No. 6, pp. 1551–1580.
- [32] Johansen, S. Likelihood-Based Inference in Cointegrated Vector Autoregressive Models (New York : Oxford University Press), 1995.
- [33] Jones, Charles M, Kaul, Gautam. Oil and stock markets.*J Finance*; 51 :463-491 (1996).
- [34] Joshua C.C. Chan , Angelia L. Grant . Modeling energy price dynamics : GARCH versus stochastic volatility. CAMA Working Paper (2015).
- [35] Kesten H. & Spitzer F. (1984) Convergence in distribution of products of random matrices. *Probability Theory and Related Fields* 67, 363–386.
- [36] Kingman, J. F. C. Subadditive ergodic theory, *The Annals of Probability*. 1 (1973), 883-889.
- [37] Lallouche, A. Volatilité stochastique : application aux indices boursiers. Thèse de doctorat (2017).
- [38] Lee, S-W, and B. E. Hansen (1994). "Asymptotic theory for the GARCH(1,1) quasimaximum likelihood estimator," *Econometric Theory*, 10, 29-52.
- [39] Olga Efimova, and Apostolos Serletis. Energy markets volatility modelling using GARCH. *Energy Economics*, Volume 43, Pages 264–273(2014).
- [40] PARK, J. W. and Ratti, R. A., “Oil price shocks and stock markets in the U.S. and 13 European countries.” *Energy Economics*, vol. 30, (2008), pp. 2578–2608.
- [41] Phillips, P. C. B. and Ouliaris, S. Asymptotic Properties of Residual Based Tests for Cointegration. *Econometrica*, Vol. 58, No. 1 (Jan., 1990), pp. 165-193.
- [42] Reinsel, G. C. (2005), *Elements of Multivariate Time Series Analysis*. Springer-Verlag : Berlin.

- [43] Sadorsky, P. Oil price shocks and stock market activity. *Energy Economics* 21 : 449-469 (1999).
- [44] Silvennoinen A. et Terasvirta .T (2008). « Multivariate GARCH models », Stockholm School of Economics/ The Economic Research Institute (SSE/ EFI) Working Paper- Series in Economics and Finance, No . 669.
- [45] Weiss, A. A. Asymptotic Theory for ARCH Models : Estimation and Testing. *Econometric Theory*, 1986, vol. 2, issue 01, 107-131.
- [46] Wei,W.W. (1994), *Time Series Analysis, Univariate and Multivariate Methods*.Addison-Wesley Company.
- [47] Yudong Wang , and Chongfeng Wu. Forecasting energy market volatility using GARCH models : Can multivariate models beat univariate models ?. *Energy Economics*, Volume 34, Issue 6, Pages 2167-2181 (2012).
- [48] Zeghdoudi, H. Boudjaada L. Tlaidjia N. Use of the Artificial Neural Networks (ANNs) to Guiding the Financial Decisions : Exchange Rate Volatility of Algerian Dinar. *International Research Journal of Finance and Economics*, Issue 115 October, 87-94 (2013b).
- [49] Zeghdoudi, H. Bousebaa , F,Z . Use of the GARCH Models to Energy Markets : Oil Price Volatility. *Global Journal of Pure and Applied Mathematics*. Vol11, N° 6 pp. 4385-4394 (2015).
- [50] Zeghdoudi, H. Ezzesba, A.Remita, M. Nedjar,S.Around ARCH/GARCH models and their application to exchange rate volatility. *International Journal of Statistics and Economics*, Vol 11 N 2 , 44-60 (2013a).