

وزارة التعليم العالي والبحث العلمي

BADJI MOKHTAR -ANNABA
UNIVERSITY
UNIVERSITE BADJI MOKHTAR
ANNABA



جامعة باجي مختار
- عنابة -

Faculté des Sciences

Année : 2021/2022

Département de Mathématiques

THÈSE

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

**Modélisation des titres financiers avec volatilité
stochastique (Modèle GARCH)**

Option

Probabilités et Statistiques

Par

AMIRI Omrane

DIRECTEUR DE THÈSE: REMITA Mohamed Riad Pr. ENSIA ALGER

Devant le jury

| | | | |
|---------------------|-------------------------------|-------|---------------|
| PRESIDENT: | MERZOUGUI Mouna | Pr. | U.B.M. ANNABA |
| EXAMINATEUR: | ZAGHDOUDI Halim | Pr | U.B.M. ANNABA |
| EXAMINATEUR: | BRAHIMI Brahim | Pr. | Univ. BISKRA |
| EXAMINATEUR: | EZZEBSA Abdelali | M.C.A | Univ. GUELMA |
| EXAMINATEUR: | BENCHETTAH Djaber Chemseddine | M.C.A | ESSG ANNABA |

Dédicace

A mes parents,

Ceux qui m'ont vu réussir et n'ont pas eu le temps de s'en réjouir,
Je leur exprime toute ma reconnaissance et mon infinie gratitude pour tout,

À eux je dédie

Le fruit de ce succès:

À mon père, qui m'a tout offert pendant toute sa vie

Et qui a toujours voulu que j'atteigne les plus hauts niveaux,

J'espère, à travers cela, son âme sera bénie dans son lieu de repos,

À ma mère qui a sacrifié du plus précieux qu'elle a pour que mon rêve soit réalisé,

Que le dieu le tout-puissant la garde et qu'elle soit heureuse dans la vie.

C'est l'occasion pour moi de vous réaffirmer ma grande reconnaissance

En témoignage de vos soucis de me mettre à l'aise.

Tout le plaisir est pour moi de vous dédier ce travail.

Qui est en fait le résultat de vos efforts.

À mes chère enfant Mariem-Ritedj et Taquieddine

Qu'ils soient toujours heureux, brillant et réussi.

Omrane

Remerciement

Je tiens à exprimer toute ma reconnaissance à mon directeur de thèse Prof. REMITA Mohamed Riad de l'Université de Badji-Mokhtar Annaba, Je le remercie de m'avoir encadré, et pour l'intéressant sujet qui m'a accordé.

Qu'il soit aussi remercié pour la compréhension, sa disponibilité et les encouragements.

J'adresse tous mes remerciements aux membres de jury qui me font l'honneur par leur présence à évaluer cette thèse, la présidente Prof. MERZOUGUI Mouna et les examinateurs Prof. ZAGHDOUDI Halim, Prof. BRAHIMI Brahim, Mr. EZZEBSA Abdelali et Mr. BENCHATTAH Djebber Chemseddine que je leur exprime ma gratitude.

Je tiens aussi à remercier tous les membres de laboratoire de recherche LEM (Lille Economie et Management) de l'université de Lille 3 de France notamment les professeurs Sophie Dabo Niang, Laurence Broze et Christian Francq, de m'avoir accueilli au sein du laboratoire pendant mon stage scientifique dans le cadre du programme PNE et pour leur soutien que j'apprécie beaucoup.

Mes sincères remerciements à tous mes collègues et amis pour l'aide que m'ont été apportée

TABLE DES MATIÈRES

| | |
|---|-----------|
| Introduction | 5 |
| 1 Modèles de Box-Jenkins | 10 |
| 1.1 Notions fondamentales | 10 |
| 1.1.1 Processus Stochastiques | 10 |
| 1.1.2 Stationnarité | 11 |
| 1.1.3 Série chronologique | 12 |
| 1.2 Modèles linéaires | 13 |
| 1.2.1 Fonction d'autocovariance | 13 |
| 1.2.2 Fonction d'autocorrélation (ACF) | 14 |
| 1.2.3 Fonction d'autocorrélation partielle (PACF) | 15 |
| 1.2.4 Modèles ARMA (<i>AutoRégressif Moyenne Mobile</i>) | 17 |
| 1.2.5 Modèles ARIMA (<i>AutoRégressif Moyenne Mobile Intégré</i>) | 20 |
| 1.3 Modèles de volatilité | 27 |
| 1.3.1 Séries financières | 27 |
| 1.3.2 Hétéroscédasticité | 28 |
| 1.3.3 Modèles ARCH (<i>AutoRegressive Conditional Heteroskedasticity</i>) | 30 |
| 1.3.4 Modèles GARCH (<i>Generalized ARCH</i>) | 31 |
| 1.3.5 Modèles GARCH exponentiel (EGARCH) | 38 |
| 1.3.6 Modèles GARCH à seuil (TGARCH) | 39 |
| 2 Identification d'un modèle GARCH | 43 |
| 2.1 Vérification de l'absence de l'autocorrélation pour le bruit blanc | 44 |
| 2.1.1 Comportement des autocorrélations empiriques d'un processus GARCH | 45 |
| 2.1.2 Test Portemanteau (Ljung-Box) | 46 |

| | | |
|----------|--|------------|
| 2.1.3 | Autocorrélations partielles empiriques d'un GARCH | 47 |
| 2.2 | Identification des ordres ARMA d'un ARMA-GARCH | 48 |
| 2.2.1 | Autocorrélations empiriques d'un ARMA-GARCH | 49 |
| 2.2.2 | Autocorrélations empiriques de ARMA-GARCH (bruit non symétrique) | 51 |
| 2.2.3 | Identification des ordres (P,Q) | 53 |
| 2.3 | Identification des ordres GARCH d'un ARMA-GARCH | 56 |
| 2.3.1 | Méthode de coin dans le cas GARCH | 56 |
| 2.4 | Test du multiplicateur de Lagrange pour l'homoscédasticité conditionnelle | 57 |
| 2.4.1 | Forme générale du test LM | 57 |
| 2.4.2 | Test LM d'homoscédasticité conditionnelle | 61 |
| 3 | Estimation d'un modèle GARCH | 64 |
| 3.1 | Estimation d'un modèle ARCH par les moindres carrés | 64 |
| 3.1.1 | Estimation de modèle ARCH(q) par MCO | 65 |
| 3.1.2 | Estimation par MCQG des modèles ARCH(q) | 70 |
| 3.1.3 | Estimation par Moindres Carrés Ordinaires contraints | 74 |
| 3.2 | Estimation par Quasi-Maximum de Vraisemblance des modèles GARCH | 78 |
| 3.2.1 | Quasi-vraisemblance conditionnelle | 78 |
| 3.2.2 | Estimation par Quasi-Maximum de Vraisemblance des modèles ARMA- | |
| | GARCH | 87 |
| 4 | Measuring the impact of covid-19 pandemic on the oil prices volatility by | |
| | GARCH modeling | 93 |
| 4.1 | Introduction | 95 |
| 4.2 | Literature review | 97 |
| 4.3 | Research methodologie | 100 |
| 4.3.1 | Non-Linear Models and heteroskedasticity | 100 |
| 4.4 | Analysis of results and discussion | 104 |
| 4.4.1 | Statistical analysis of WTI raw serie's behavior | 104 |
| 4.4.2 | Stationarity study | 105 |
| 4.4.3 | Statistical analysis of leverage effect and ARCH effect | 107 |
| 4.4.4 | Estimation of models and search for the optimal model | 110 |
| 4.4.5 | Forecasting using the formulated model | 113 |
| | Conclusion | 116 |
| | Bibliographie | 121 |

ملخص

في هذه الأطروحة سوف نقترح طريقة لقياس التقلب (التباين الشرطي) من نوع GARCH لسلسلة زمنية معينة تتميز بالتقلبات القوية مثل أسعار النفط والمؤشرات المالية ، مع الأخذ بعين الاعتبار منهجية Box-Jenkins التي تم تكييفها في بدايتها مع نماذج ARMA - التي عرفت محدوديتها فيما يتعلق بهذا النوع من السلاسل الاقتصادية والمالية - وتتكون من المراحل التالية: تحديد النموذج ، وتقدير المعلمات ، والتحقق من الصحة والتنبؤ. وهذا من أجل استخراج النماذج المناسبة غير الخطية لبيانات الأصول ، مثل تلك التي درسناها في الفصل الثالث من الأطروحة ، وهي أسعار النفط الخام. من بين هذه النماذج T-GARCH و E-GARCH ، والتي أظهرنا أنها فعالة للغاية في تمثيل البيانات. البيانات المستخدمة هنا يومية وتشمل الفترة من جانفي 2015 ديسمبر 2021. وشملت هذه الفترة جائحة كورونا الذي أثرت بشكل كبير على نتائج الدراسة التي أجريناها.

كلمات مفتاحية: التباين غير المتجانس ، تأثير ARCH ، اللاتناظر ، النماذج الخطية ، نماذج GARCH ؛ تقلبات أسعار النفط، التنبؤ، التطاير، السلاسل الزمنية ، منهجية Box-Jenkins

Résumé

Dans cette thèse nous proposons une méthode de mesure de la volatilité (variance conditionnelle) de type GARCH pour des séries temporelles qui se caractérisent par des fortes fluctuations comme les prix du pétrole ou les indices boursiers, en tenant compte de la méthode de Box-Jenkins adaptée en premiers temps aux modèles ARMA, et qui consiste des phases suivantes : l'identification du modèle, l'estimation des paramètres, la validation et la prédiction. Nous adopterons cette méthode pour les modèles GARCH afin de chercher le modèle générateur des données de l'actif financier et économique, qui est, dans notre cas, le prix du pétrole brut, noté WTI. Le but de cette thèse est de trouver le modèle adapté aux données WTI, qui est censé un type de GARCH en raison de la nature de la variable caractérisée par des fortes fluctuations. Dans cette analyse, l'existence d'un événement extrême, dû à la récente crise de Coronavirus couvert par la période des observations concernées, est prouvé, et qui a considérablement affecté les résultats de l'étude que nous avons menée, comme illustré dans le quatrième chapitre. Dans cette étude le phénomène d'asymétrie de la volatilité ou l'effet de levier est prouvé ce qui montre la compatibilité des modèles T-GARCH et E-GARCH avec la variable des prix. Dans la partie pratique les données utilisées sont quotidiennes et s'étendent du janvier 2015 à décembre 2021.

Mots clés : Série chronologique ; Méthode Box-Jenkins ; Modèle GARCH ; Volatilité ; Asymétrie ; Prix de pétrole, Préviation.

Abstract

In this thesis we propose a method for measuring volatility (conditional variance) of the GARCH type for time series which are characterized by strong fluctuations such as oil prices or stock market indices, taking into account the Box-Jenkins method which was initially adapted to ARMA models, and which consists of the following phases : model identification, parameter estimation, validation and prediction. We will adopt this method for the GARCH models in order to seek the model generating the data of the financial and economic asset, which is, in our case, the price of crude oil, denoted WTI. The aim of this thesis is to find the model adapted to the WTI data, which is supposed to be a type of GARCH due to the nature of the variable characterized by strong fluctuations. In this analysis, the existence of an extreme event, due to the recent Coronavirus crisis covered by the period of the observations concerned, is proven, and which considerably affected the results of the study that we carried out, as illustrated in the fourth chapter.

In this study the phenomenon of volatility asymmetry or the leverage effect is proven, which shows the compatibility of the T-GARCH and E-GARCH models with the price variable. In the practical part, the data used are daily and extend from January 2015 to December 2021.

Keywords : Time series ; Box-Jenkins method ; GARCH model ; Volatility ; Asymmetry ; Oil Price ; Forecast.

Introduction

L'analyse des séries temporelles, se réfère à la branche de la statistique où les observations du phénomène sont régulièrement espacées dans le temps. Elles sont utilisées pour la première fois en astronomie ('on the periodicity of sunspots', 1906 par Arthur Schuster), et l'étude de l'analyse des séries temporelles est fondamentale quant à leurs utilisations dans de nombreux domaines tels que l'économie, la finance, l'économétrie, la médecine et la biologie, la science politique, la météorologie, etc... Elle se peut considérer parmi les méthodes de statistiques de prévision, très souvent la prévision se fait pour des raisons économiques (prévoir l'évolution des ventes d'un certain produit, prévoir la consommation d'électricité pour ajuster au mieux la production, etc...). Un nombre important d'exemples de séries chronologiques peut être trouvé dans ([11], [12], [13]).

L'utilisation systématique dans tous les domaines de ces dernières, que ce soit économiques, industriels, physiques etc..., est due au développement exponentiellement rapide de l'outil informatique.

Dans tous les domaines précités, les séries chronologiques observées peuvent être considérées comme des réalisations de processus stochastique $\{X_t, t \in T\}$. c'est à dire une suites de variables aléatoires définies sur un même espace Ω , appelé espace fondamental ou espace des états de la nature. On dit qu'une série chronologique est continue si l'ensemble des instants des observations est continu (ou $t \in I \subset \mathbb{R}$), et qu'elle est discrète si l'ensemble de ces instants est discret (ou $t \in I \subset \mathbb{Z}$) ([11]). Dans ce travail, nous intéressons aux séries chronologiques discrètes et univariées où les observations sont réalisées en des intervalles de temps équidistants.

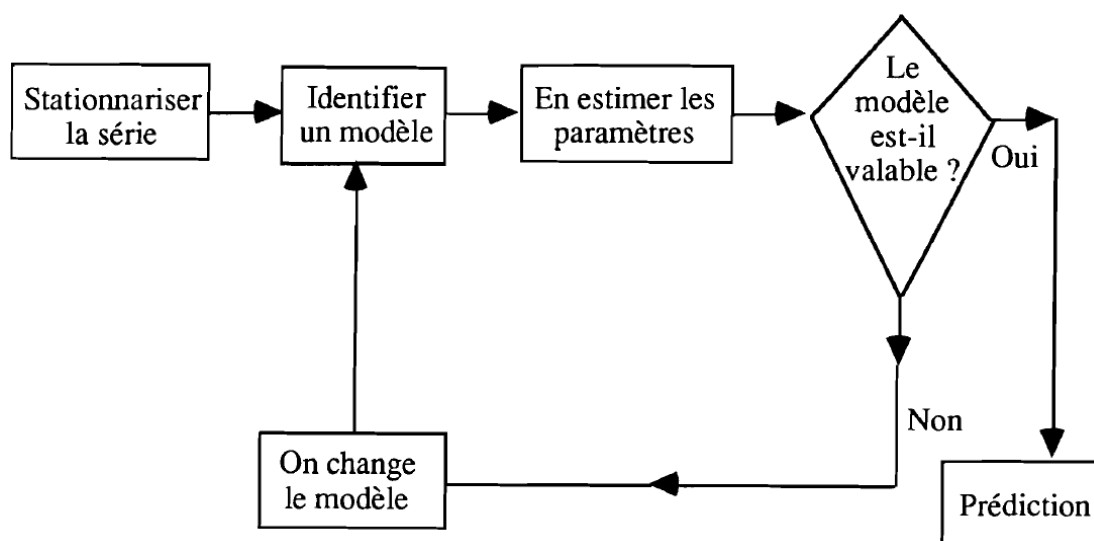
Dans tous ces domaines aussi, il est très utile de pouvoir prédire le futur c'est-à-dire le temps $t + h$, connaissant le passé jusqu'au temps t , t et h étant respectivement l'origine et l'horizon de la prédiction.

La théorie et la pratique de l'analyse des séries temporelles sont développées rapidement et elles semblent avoir atteint la maturité au cours des années 70 où des développements significatifs sont apparus . En 1965, Cooley et Tukey ont beaucoup aidé l'étude spectrale des séries grâce à leur

article ('An algorithm for the machine calculation of complex Fourier series'), introduisant la transformée de Fourier rapide (Fast Fourier Transform : FFT) ([14]). Cet algorithme a permis de calculer rapidement des périodogrammes. À la même époque, en 1970, Box et Jenkins ont publié leur ouvrage ('Time series analysis, forecasting and control ') ([8]), montrant que l'étude des séries temporelles à l'aide du processus de type $ARMA$ (obtenus en associant les écritures des processus AR et MA) s'appliquent à de nombreux domaines, et pouvait être facilement implémentée informatiquement.

Pour rappel, les représentations de AR et MA sont apparus en 1927, respectivement, par l'article de G Udny Yule ('On the method of investigating periodicities in disturbed series with special reference to Wolfer's sunspot numbers ') ([52]) et celui de Eugen Slutsky ('The summation of random causes as the source of cyclical processes ') ([47]).

La méthode de Box et Jenkins liées à la modélisation $ARMA$ est la plus populaires depuis des dizaines d'années. Elle consiste en gros à dégager les tendances évidentes dans le phénomène qu'on observe (croissance, périodicités...) et à se concentrer sur ce qui reste. Son procédure peut être résumée comme suit :



Malgré leurs avantages, les modèles $ARMA(p, q)$ souffrent de la non prise en compte de certaines contraintes structurelles liées au phénomène faisant l'objet de la modélisation. Ces contraintes peuvent traduire le caractère volatile de certaines variables, impliquent parfois le

recours à des modèles non linéaires susceptibles de rendre inadéquat la spécification *ARMA*. Ce qu'exigeait l'émergence d'autres modèles plus qualifiés. Et c'était en effet au début des années 1980 quand l'économètre Robert Engle a généralisé cette méthode de Box et Jenkins dans un article publié dans la revue *Econometrica* ([21]), il s'agit des modèles ARCH (AutoRegressive Conditional Heteroskedasticity). **Cette généralisation est considérée comme le passage des modèles linéaire vers** des autres non linéaires, ce qui permet d'intégrer des propriétés observées empiriquement, telles que la dépendance entre les observations ou la distribution à queue lourde. La spécification d'Engle n'a duré que quatre ans, jusqu'à ce qu'elle soit généralisée par son doctorant Tim Bollerslev en 1986 aux modèles GARCH (Generalised AutoRegressive Conditional Heteroskedasticity) ([4]), ces modèles ont connu un vif succès tant académique qu'auprès des praticiens. Les modèles de GARCH ont conduit à une profonde modification des approches utilisées en finance à travers une spécification particulièrement efficace de la volatilité (c'est à dire de la variabilité) des cours des actifs financiers. Le développement des recherches dans ces modèles s'est vue couronnée par le prix décerné par l'académie Nobel à Engle en 2003. Après cette découverte pionnière, il est devenu courant dans la littérature financière de retenir comme mesure de volatilité (variance conditionnelle) les modèles GARCH.

Plus particulièrement, La modélisation de la volatilité par les modèles GARCH est un enjeu important dans les recherches sur les marchés financiers, économiques et énergétiques pour prédire les risques qui peuvent se produire dans les actifs physiques ou financiers. Parmi ces importants actifs, le pétrole, la cruciale source d'énergie au niveau mondial, et qui représente le moteur de l'économie de plusieurs pays et qui attire toujours l'attention de la communauté internationale en raison des fortes variations affectant ses cours.

Pendant ces dernières années, le prix du baril de brut a fortement fluctué, surtout à la baisse. Notons qu'entre septembre 2014 et Janvier 2015, ce dernier a chuté de 55% suite à la décision de l'OPEP de maintenir inchangé son niveau d'offre qui vient s'ajouter à un ralentissement économique de certains pays émergents notamment la Chine et la montée en croissance de la production du pétrole de schiste par les Etats-Unis. Le même comportement s'est enregistré depuis fin 2019 après que la pandémie de Corona virus ait balayé le monde entier. Bientôt, les prix ont renoué à la hausse inattendue en raison de la reprise après la pandémie et de

l'émergence de conflits en Europe de l'est à la fin de 2021 jusqu'à ce jours.

La question de la volatilité des cours du pétrole brut (tel que WTI : West Texas Intermediate) intéresse et motive en permanence les chercheurs et les décideurs de politiques ainsi les gestionnaires des portefeuilles, notamment pendant les événements imprévisibles, comme la bulle internet de 2002, la crise financière de 2008, la pandémie de Corona virus 2019, les guerres, etc...Des articles de recherche sur ce sujet sont constamment publiés, nous en avons évoqué plusieurs dans notre dernier article paru dans *advances in mathematics : Scientific Journal* en 2022 ([1]).

L'objectif principal de cette thèse sera l'étude de la volatilité dans le cadre des modèles GARCH univariés qui répondent au insuffisance des modèles linéaires ARMA telles que :

- La non stationnarité (surtout en variance).
- La non normalité de la série ou des données sous étude qui est la synonyme de la non-linéarité.

L'analyse de la volatilité des prix dans cette étude suivra la méthode de prévision de Box et Jenkins développée en 1976 ([9]). Cette méthode constitue un outil systématique qui permet de déterminer un modèle susceptible d'être de type GARCH restituant le mieux possible le comportement d'une série chronologique observée. Le choix d'un tel modèle est due à la forte fluctuation qui caractérise ce genre de phénomène comme la plupart des actifs économiques et financières des marchés mondiaux ([24]).

L'analyse d'une série chronologique par cette méthodologie se réalise principalement en quatre phases essentielles : Identification, Estimation, Validation et Prévision.

Pour mener à bien notre objectif, nous avons réparti cette thèse en quatre chapitres résumés comme suit :

Chapitre 1 : Modèles de Box-Jenkins

Au début nous évoquons tous les notions de base nécessaires des processus stochastiques et des séries chronologiques, particulièrement, la stationnarité, bruit blanc, l'autocorrélation etc..., puis nous présentons les différents modèles linéaires de Box-Jenkins : ARMA et leurs caractéristiques, ensuite nous analysons les modèles non linéaires : ARCH, GARCH et les modèles asymétriques tels que TGARCH et EGARCH.

Chapitres 2 : Phase d'identification de Box-Jenkins pour un modèle GARCH

Est réservé à la phase de l'identification de la méthodologie de Box-Jenkins qui consiste à vérifier et à tester les auto-correlations des résidus et le carré des résidus et à estimer en utilisant principalement les tests de Ljang-Box et le test de l'effet ARCH de Multiplicateur de Lagrange (LM). Egalement, à chercher les ordres du modèle en basant généralement sur plusieurs étapes, en utilisant divers outils :

- Analyse des fonctions de l'auto-corrélation empiriques (ACF) et d'auto-corrélation partielle empiriques (PACF) de $\epsilon_1^2, \dots, \epsilon_n^2$.
- Analyse des statistiques des fonctions des auto-covariances empiriques (ACV) de ϵ_t^2 .
- Utilisation des critères d'information (AIC, BIC, ...)
- Test de significativité de certains coefficients (test de LM, ...).
- Analyse des résidus.

Chapitre 3 : Phase d'estimation des modèles GARCH

Le troisième chapitre évoque deux parties essentielles : l'estimation des paramètres (coefficients) du modèle ARCH et l'estimation des paramètres du modèle GARCH par la méthode des moindres carrés et celle de quasi maximum de vraisemblance.

Chapitre 4 : Applications : mesure de volatilité des rendements des prix du pétrole brut WTI

Le quatrième chapitre représente notre contribution de recherche, la propriété d'asymétrie de la volatilité à travers une application par la méthodologie de Box-Jenkins sur la série des cours du pétrole brute WTI (West Texas Intermediate : one of a type of crude oils used as a standard in pricing crude oil and as a commodity for oil futures contracts on the New York Mercantile Exchange (raw materials stock Exchange)). Le travail s'achèvera par une conclusion générale et perspectives.

CHAPITRE

1

Modèles de Box-Jenkins

1.1 Notions fondamentales

1.1.1 Processus Stochastiques

On appelle un processus stochastique la famille des variables aléatoires qui dépendent d'une variable déterministe t .

L'objet de cette section est de préciser la définition des processus stochastiques $X = \{X_t(\omega)\}_{t=1,\dots,n} = (X_1(\omega), \dots, X_n(\omega))'$ et d'examiner leurs propriétés.

Les trois éléments suivants caractérisent le processus stochastique :

Une espérance $\mathbb{E}(X_t)$ à chaque période t .

Une variance $\mathbb{V}(X_t)$ à chaque période t .

Des covariances $\mathbb{Cov}(X_i, X_j)$ pour $1 \leq i, j \leq n$.

La classe des processus aléatoires est très large, les séries chronologiques en font partie à cette classe qui peuvent être stationnaires ou non dans leur nature. En général, dans l'étude des processus, en particulier les séries chronologiques, on doit procéder à leur stationnarité après avoir connu la nature de la non stationnarité.

1.1.2 Stationnarité

Processus Stationnaires

La question de stationnarité joue un rôle majeur dans les processus stochastiques notamment en séries temporelles car elle remplace de manière naturelle l'hypothèse d'observation i.i.d (indépendantes et identiquement distribuées) en statistique standard.

L'étude de la stationnarité est une étape importante avant celle de l'estimation. Le but est de chercher sous quelles conditions le processus étudié est stationnaire. On distingue deux notions de stationnarité, stationnarité stricte et stationnarité au second ordre. Alors on a les deux définitions suivantes.

Définition 1.1.1 (*Stationnarité stricte*)

Le processus (X_t) est dit strictement stationnaire si les vecteurs $(X_1, X_2, \dots, X_k)'$ et $(X_{1+h}, X_{2+h}, \dots, X_{k+h})'$ ont la même loi jointe pour tout $k \in \mathbb{N}$ et tout $h \in \mathbb{N}$.

La notion suivante peut sembler moins exigeante car elle n'impose de contrainte qu'aux deux premiers moments des variables X_t , mais contrairement à la stationnarité stricte.

Définition 1.1.2 (*Stationnarité du second ordre*)

Le processus (X_t) est dit faiblement stationnaire (ou stationnaire au second ordre), si et seulement si :

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

Bruit Blanc

Un bruit blanc (ϵ_t) est une réalisation d'un processus aléatoire dans lequel la densité spectrale de puissance est la même pour toutes les fréquences de la bande passante. Le bruit blanc gaussien est un bruit blanc qui suit une loi normale de moyenne et variance donnée. Des générateurs de signaux aléatoires («signal de bruit») sont utilisés pour des essais de dispositifs de transmission et à faible niveau pour l'amélioration des systèmes numériques par dither. Parmi les processus

aléatoires le bruit blanc revêt une importance particulière parce qu'il représente l'archétype de beaucoup de phénomènes physique.

Définition 1.1.3 (*Bruit Blanc Faible*)

Le processus (ϵ_t) est appelé bruit blanc faible s'il vérifie les hypothèses suivantes, pour une constante positive σ .

- (i) $\mathbb{E}(\epsilon_t) = 0, \forall t \in \mathbb{Z}$,
- (ii) $\mathbb{E}(\epsilon_t^2) = \text{Var}(\epsilon_t) = \sigma^2, \forall t \in \mathbb{Z}$,
- (iii) $\text{Cov}(\epsilon_t, \epsilon_{t+h}) = 0, \forall t, h \in \mathbb{Z}, h \neq 0$.

On notera

$$\epsilon_t \sim BB(0, \sigma^2)$$

Définition 1.1.4 (*Bruit Blanc Fort*).

A noter qu'aucune hypothèse d'indépendance n'est faite dans la définition du bruit blanc faible. Les variables à différentes dates ne sont que décorréllées et la distinction est particulièrement cruciale pour les séries temporelles financières. Hypothèse (iii) les variables ϵ_t et ϵ_{t+h} sont indépendantes et identiquement distribuées.

Le processus (ϵ_t) est alors dit bruit blanc fort.

Le bruit blanc est l'exemple le plus simple d'un processus stationnaire de second ordre. Ce processus est particulièrement important car il permet de construire des processus stationnaires plus complexes.

1.1.3 Série chronologique

Une série chronologique (temporelle) est une réalisation d'un processus stochastique, En fait, une série chronologique est une suite finie des variables aléatoires (X_1, \dots, X_n) de données indexées par le temps. L'indice temps peut être selon les cas la minute, l'heure, le jour, l'année etc.... Le nombre n est appelé la longueur de la série. Il est la plupart du temps bien utile de représenter la série temporelle sur un graphe construit de la manière suivante : en abscisse le temps, en ordonnée la valeur de l'observation à chaque instant. Pour des questions de lisibilité,

les points ainsi obtenus sont reliés par des segments de droite. Le graphe apparaît donc comme une ligne brisée.

1.2 Modèles linéaires

Les modèles autorégressifs moyenne mobile (*ARMA*) est une technique de modélisation économétrique avancée. Ces modèles analysent des données historiques des phénomènes et effectuent des ajustements rétrospectifs pour prendre en compte l'autocorrélation historique (la relation d'une valeur par rapport à une autre valeur dans le temps), la stabilité des données pour corriger les caractéristiques non stationnaires des données. Ces modèles prédictifs apprennent dans le temps en corrigeant ses erreurs de prévisions. Des connaissances économétriques avancées sont généralement requises pour construire de bons modèles prédictifs à l'aide de cette approche.

1.2.1 Fonction d'autocovariance

Un processus ou une série temporelle (X_t) est caractérisée, entre autres, par sa structure d'ordre deux. C'est à dire :

- Sa moyenne qui est un moment d'ordre un.
- Son autocovariance qui est un moment d'ordre deux et qui mesure la dépendance temporelle linéaire entre deux instants.

Définition 1.2.1 (*fonction d'autocovariance*)

Soit $X = \{X_1, \dots, X_n\}$ une série à temps discret. La fonction de covariance (ou d'autocovariance) de X est donnée par

$$\gamma(h) := \text{Cov}(X_t, X_{t+h}) = \mathbb{E}[X_t X_{t+h}] - \mathbb{E}[X_t] \mathbb{E}[X_{t+h}]$$

La fonction d'autocovariance est estimée en général par une fonction d'autocovariance empirique pour une réalisation $\{X_1, \dots, X_n\}$ de X par

$$\hat{\gamma}(h) = \frac{1}{n} \sum_{i=1}^{n-h} (X_i - \bar{X})(X_{i+h} - \bar{X}),$$

où \bar{X} est l'estimateur empirique de la moyenne (globale)

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i.$$

C'est une fonction paire, semi-définie positive, i.e.,

$$\sum_{i=1}^N \sum_{j=1}^N a_i a_j \gamma(t_i - t_j) \geq 0,$$

$$|\gamma(h)| \leq \gamma(0) = \text{Var}(X_t), \quad h \in \mathbb{Z}.$$

1.2.2 Fonction d'autocorrélation (ACF)

Définition 1.2.2 (*Fonction d'autocorrélation*)

De même, on définit une suite $(\rho(h))_{h \in \mathbb{Z}}$ qu'on l'appelle fonction d'autocorrélation de la série (X_t) :

$$\rho(h) := \frac{\text{Cov}(X_t, X_{t+h})}{\sqrt{\text{Var}(X_t) \text{Var}(X_{t+h})}} = \frac{\gamma(h)}{\gamma(0)}.$$

C'est une fonction paire, semi-définie positive, i.e.,

$$\sum_{i=1}^N \sum_{j=1}^N a_i a_j \rho(t_i - t_j) \geq 0,$$

$$\rho(h) \leq \rho(0) = 1, \quad h \in \mathbb{Z}.$$

On appelle coefficient d'autocorrélation d'ordre 1 (resp. d'ordre h) le coefficient de corrélation linéaire $\rho(1)$ (resp $\rho(h)$) obtenu à partir de la série et la même série décalée d'une seule période (resp. h périodes)). Le graphe de la fonction d'autocorrélation est appelé corrélogramme (les coefficients $\rho(h)$, $1 \leq h \leq m$ appelés "les piques" sur ce corrélogramme)

On définit la matrice de corrélation (de dimension $m \times m$) de la manière suivante :

$$R(m) = \begin{pmatrix} 1 & \rho(1) & \rho(2) & \cdots & \rho(m-1) \\ \rho(1) & 1 & \rho(1) & \cdots & \rho(m-2) \\ \rho(2) & \rho(1) & 1 & \cdots & \rho(m-3) \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \rho(m-1) & \rho(m-2) & \cdots & \cdots & 1 \end{pmatrix}.$$

Ainsi l'estimateur de la fonction d'autocorrélation est donné par

$$\hat{\rho}(h) = \frac{\hat{\gamma}(h)}{\hat{\gamma}(0)}.$$

La nullité ou non des coefficients d'autocorrélation d'ordre $h \in \mathbb{Z}$ est une question importante, donc, pour cela on a besoin du test d'hypothèse :

$$\begin{cases} H_0 : \rho_1 = \rho_2 = \dots = \rho_m = 0 \\ H_1 : \rho_1 \neq \rho_2 \neq \dots \neq \rho_m \neq 0. \end{cases}$$

La statistique du test de la nullité est ainsi

$$t = \frac{\hat{\rho}(h)}{\sqrt{\text{Var}(\hat{\rho}(h))}}.$$

Elle suit une loi de Student, où $\text{Var}(\hat{\rho}(h))$ est la variance du coefficient d'autocorrélation

$$\text{Var}(\hat{\rho}(h)) = \frac{1}{T} \sum_{j=-(h-1)}^{h-1} \hat{\rho}^2(j),$$

en utilisant la symétrie de $\rho(j)$, on aura

$$\text{Var}(\hat{\rho}(h)) = \frac{1}{T} \left(1 + 2 \sum_{j=1}^{h-1} \hat{\rho}^2(j) \right),$$

La variance dépend de h , l'intervalle de confiance associé au **corrélogramme** à 95% des AC est donnée par

$$\left[\hat{\rho}(h) \pm 1.96 \sqrt{\frac{1}{T} \left(1 + 2 \sum_{j=1}^{h-1} \hat{\rho}^2(j) \right)} \right]$$

1.2.3 Fonction d'autocorrélation partielle (PACF)

Elle mesure la liaison (linéaire) entre X_t et X_{t-h} une fois retirés les liens transitant par les variables intermédiaires $X_{t-1}, \dots, X_{t-h+1}$. Le coefficient d'autocorrélation partielle d'ordre h , noté $r(h)$, est le coefficient de corrélation entre :

$$X_t - \mathbb{E}(X_t \setminus X_{t-1}, \dots, X_{t-h+1}) \quad \text{et} \quad X_{t-h} - \mathbb{E}(X_{t-h} \setminus X_{t-1}, \dots, X_{t-h+1})$$

On a alors,

$$r(h) = \text{Corr}(X_t, X_{t-h} \setminus X_{t-1}, \dots, X_{t-h+1})$$

C'est donc le coefficient de X_{t-h} dans la regression de X_t sur $X_{t-1}, \dots, X_{t-h+1}, X_{t-h}$.

Si (X_t) est un processus stationnaire centré, la prédiction optimale de X_t sachant son passé jusqu'à $t - h$ est donnée par

$$X_t = \mathbb{E}(X_t \mid X_{t-1}, \dots, X_{t-h+1}) = a_1 X_{t-1} + a_2 X_{t-2} + \dots + a_h X_{t-h+1}$$

en multipliant l'équation précédente par X_{t-i} , $i = 1, \dots, t - h$ et cherchant l'espérance, on a le système suivant :

$$\begin{cases} \gamma(1) = a_1 \gamma(0) + a_2 \gamma(1) + \dots + a_h \gamma(h-1) \\ \gamma(2) = a_1 \gamma(1) + a_2 \gamma(0) + \dots + a_h \gamma(h-2) \\ \dots \\ \gamma(h) = a_1 \gamma(h-1) + a_2 \gamma(h-2) + \dots + a_h \gamma(0) \end{cases}$$

que l'on peut réécrire matriciellement

$$\begin{aligned} \begin{pmatrix} a_1 \\ a_2 \\ \dots \\ a_h \end{pmatrix} &= \begin{pmatrix} \gamma_0 & \gamma_1 & \dots & \gamma_{h-1} \\ \gamma_1 & \gamma_0 & \dots & \gamma_{h-2} \\ \dots & \dots & \dots & \dots \\ \gamma_{h-1} & \gamma_{h-2} & \dots & \gamma_0 \end{pmatrix}^{-1} \begin{pmatrix} \gamma_0 \\ \gamma_1 \\ \dots \\ \gamma_{h-1} \end{pmatrix} \\ &= \begin{pmatrix} 1 & \rho_1 & \dots & \rho_{h-1} \\ \rho_1 & 1 & \dots & \rho_{h-2} \\ \dots & \dots & \dots & \dots \\ \rho_{h-1} & \rho_{h-2} & \dots & 1 \end{pmatrix}^{-1} \begin{pmatrix} \rho_1 \\ \rho_2 \\ \dots \\ \rho_h \end{pmatrix} \end{aligned}$$

Le coefficient d'autocorrélation partielle d'ordre h d'un processus stationnaire est calculé de la manière suivante

$$r(h) = \frac{|R^*(h)|}{|R(h)|}$$

avec

$$R(h) = \begin{pmatrix} 1 & \rho_1 & \dots & \rho_{h-1} \\ \rho_1 & 1 & \dots & \rho_{h-2} \\ \dots & \dots & \dots & \dots \\ \rho_{h-1} & \rho_{h-2} & \dots & 1 \end{pmatrix}$$

et $R^*(h)$ est la matrice obtenue en modifiant la colonne d'ordre h dans $R(h)$ par $(\rho_1, \rho_2, \dots, \rho_h)'$

Ainsi

$$r(1) = \rho(1); \quad r(2) = \frac{\rho(2) - \rho^2(1)}{1 - \rho^2(1)};$$

Les autocorrélations partielles s'estiment en général, soit

- À partir de la régression MCO de X_t sur les h retards et en prenant le dernier coefficient.
- Soit en estimant les autocorrélations simples et en calculant $\hat{\rho}(h)$ à partir de la formule ci-dessus, ces dernières équations sont appelées équations de Yulle-Walker.

Le rôle des autocorrélations et des autocorrélations partielles est primordiale, principalement pour déduire l'ordre du modèle et pour tester la dépendance ou non entre des résidus ou entre les carrés des résidus en utilisant le diagramme du corrélogramme.

1.2.4 Modèles ARMA (*AutoRégressif Moyenne Mobile*) Modèles autorégressifs AR (George Udney Yule, [25])

On exprime souvent l'évolution d'une série en fonction de son passé. Les modèles autorégressifs sont les modèles les plus explicites pour exprimer cette dépendance, nous supposons que ε_t est indépendant de $Y_{t-1}, Y_{t-2}, Y_{t-3}, \dots$

Le modèle Y_t est un modèle autorégressif d'ordre p noté $AR(p)$, s'il satisfait l'équation

$$Y_t = c + \sum_{i=1}^p \varphi_i Y_{t-i} + \varepsilon_t, \quad \text{pour tout } t \in \mathbb{Z},$$

où $\varphi_p \neq 0$ et $\varepsilon_t \sim BB(0, \sigma^2)$.

Cette formule peut s'écrire comme suit

$$\Phi_p(L) Y_t = c + \varepsilon_t,$$

où

$$\Phi_p(L) = 1 - \varphi_1 L - \varphi_2 L^2 - \dots - \varphi_p L^p.$$

Le modèle $AR(p)$ est toujours inversible.

Le modèle $AR(p)$ est stationnaire si toutes les racines de $\Phi(L)$ soient à l'extérieur du cercle unité (i.e, les racines de $\Phi(L)$ en leurs modules sont supérieures à 1).

Modèles moyennes mobiles MA (Eugen Slutsky, [47])

On appelle modèle moyenne mobile d'ordre q , noté $MA(q)$, un modèle stationnaire Y_t vérifiant une relation du type :

$$Y_t = \mu + \varepsilon_t + \sum_{i=1}^q \theta_i \varepsilon_{t-i}, \quad \text{pour tout } t \in \mathbb{Z},$$

où les θ_i sont des réels, $\theta_q \neq 0$ et $\varepsilon_t \sim BB(0, \sigma_\varepsilon^2)$.

Si on introduit l'opérateur moyenne mobile :

$$\Theta(L) = 1 - \theta_1 L - \theta_2 L^2 - \dots - \theta_q L^q.$$

On peut noter de façon équivalente :

$$Y_t = \mu + \Theta(L) \varepsilon_t.$$

Un $MA(q)$ est toujours stationnaire quelles que soient les valeurs de θ , il est de moyenne μ .

Un $MA(q)$ est inversible si les racines de

$$1 - \theta_1 L - \theta_2 L^2 - \dots - \theta_q L^q = 0,$$

sont strictement supérieures à 1 en module.

Modèles ARMA (AutoRegressive Moving Average)

Définition 1.2.3 (*Processus ARMA(p,q)*)

$c, a_1, \dots, a_p, b_1, \dots, b_q$ tel que :

$$\forall t \in \mathbb{Z}, \quad Y_t + \sum_{i=1}^p a_i Y_{t-i} = c + \varepsilon_t + \sum_{j=1}^q b_j \varepsilon_{t-j}$$

$$\Phi(L)Y_t = c + \Theta(L) \varepsilon_t$$

Cette définition implique des contraintes sur les zéros des polynômes autorégressifs

$$\Phi(L) = 1 + \sum_{i=1}^p a_i L^i$$

et moyens mobiles

$$\Theta(L) = 1 + \sum_{i=1}^q b_i L^i.$$

Le principal attrait de ce modèle et des représentations obtenues en inversant successivement les polynômes $a(\cdot)$ et $b(\cdot)$ est qu'il fournit un cadre pour dériver les prédictions linéaires optimales

du processus de manière beaucoup plus simple qu'en supposant seulement la stationnarité du second ordre.

Des nombres séries économiques affichent des tendances ce qui rend l'hypothèse de stationnarité irréaliste.

Ces tendances disparaissent souvent lorsque la série est différenciée une ou plusieurs fois. Soit $\Delta Y_t = Y_t - Y_{t-1}$ la série en différences premières et $\Delta^d Y_t = \Delta(\Delta^{d-1} Y_t)$ avec ($\Delta^0 Y_t = Y_t$) les différences d'ordre d .

Remarque 1.2.1 (*Sur linéarité d'une série par le modèle ARMA*)

Statistiquement, si une série temporelle est adaptée à un modèle ARMA, c'est indispensable que les résidus (ϵ_t) de ce modèle estimé soient i.i.d. et cela veut dire que la série des rendements est linéaire. Cette étape statistique s'effectue par un test très célèbre, le test BDS.

Test d'indépendance des résidus (Test BDS) Le nom de ce test est une abréviation des noms de ses trois inventeurs Brock, Dechert, et Scheinkman (1987), qui est noté "Test BDS".

Son objectif est la détection de la présence de dépendance non linéaire dans une série temporelle.

En général, pour procéder ou déployer ce test :

Premièrement, on fait la différence première des données brut (on obtient la série des rendements (r_t)) pour éliminer la tendance stochastique (souci de la non linéarité).

Deuxièmement, on fait appel à un modèle linéaire qui est le modèle *ARMA* (qui permet d'éliminer la structure linéaire présente dans la série temporelle).

Troisièmement, on vérifie le blanchiment des résidus (ϵ_t) par le test *BDS* qui repose sur une hypothèse nulle :

$$H_0 : \text{Les résidus du modèle ARMA estimé sont i.i.d.}$$

contre une hypothèse H_1 non spécifiée.

Et la statistique de test BDS est définie par

$$W_{m,T}(\epsilon) = \frac{\sqrt{T} [C_m(\epsilon) - (C_1(\epsilon))^m]}{\sigma_{m,T}(\epsilon)},$$

tels que m est la dimension de plongement ou de l'espace des phases, $C_m(\epsilon)$ est l'intégrale de corrélation.

Afin de tester H_0 , on compare la statistique $W_{m,T}(\epsilon)$ avec la valeur critique 1.96 de $\mathcal{N}(0, 1)$ à 5%.

Donc la règle de décision est :

- Si $|W_{m,T}(\epsilon)| > 1.96$, alors H_0 (l'indépendance) est rejetée. Alors le processus générateur de la série est non linéaire, donc ce n'est pas ARMA sélectionné.

1.2.5 Modèles ARIMA (*AutoRégressif Moyenne Mobile Intégré*)

Définition 1.2.4 (*Processus ARIMA(p,d,q)*)

Soit d un entier positif. Le processus (Y_t) est dit processus ARIMA(p, d, q) si pour $k = 1, \dots, d-1$, les processus $(\Delta^k Y_t)$ ne sont pas stationnaires du second ordre et $(\Delta^d Y_t)$ est un processus ARMA(p, q). Le processus ARIMA le plus simple est ARIMA(0, 1, 0) aussi appelé la marche aléatoire satisfaisant $Y_t = Y_{t-1} + \epsilon_t$, $t \geq 1$ où (ϵ_t) est un bruit blanc faible.

1. Processus non stationnaires

La classe des processus non stationnaires est relativement vaste, il existe différents types de non stationnarité. Les processus ARIMA sont non-stationnaires en fait, pour les rendre stationnaires il y a un choix de la méthode qui dépend du type de la non stationnarité : on présente deux classes essentielles des processus non stationnaires : Les processus TS (trend stationary) et les processus DS (difference stationary) :

a/ Processus TS

Définition 1.2.5 (*Processus TS*)

Un processus (Y_t) est non stationnaire de type TS s'il peut présenter sous la forme :

$$Y_t = f(t) + \epsilon_t$$

où $f(t)$ est une fonction polynomiale qui dépend du temps, elle peut être linéaire ou non (le degré 1 le cas le plus couramment retenu), on pourrait aussi considérer entre autre une tendance

quadratique $f(t) = \alpha + \beta t + \gamma t^2$ ou une tendance segmentée $f(t) = a_1 + b_1 t + a_2 I_{t < t_0} + b_2 t I_{t < t_0}$, avec $I_{t < t_0} = 1$ si $t < t_0$ et 0 sinon. $f(t)$ et ϵ_t un processus stationnaire (bruit blanc est un cas particulier).

Le processus TS est le plus répandu, notamment pour les séries moins fluctuées, souvent il est présenté par une fonction polynomiale de degré 1, ce qui lui permet de porter le nom de linéarité, il s'écrit sous la forme :

$$Y_t = \alpha + \beta t + \epsilon_t,$$

ϵ_t représente l'erreur du modèle à la date t . Les caractéristiques de ce processus sont donc

$$\begin{cases} \mathbb{E}(Y_t) = \alpha + \beta t + \mathbb{E}(\epsilon_t) = \alpha + \beta t \\ \mathbb{V}(Y_t) = 0 + \mathbb{V}(\epsilon_t) = \sigma_\epsilon^2 \\ \text{Cov}(Y_t, Y_s) = 0, \quad \text{pour } t \neq s \end{cases}$$

Il présente une non stationnarité de nature déterministe (TS). Ce processus TS n'est pas stationnaire car $\mathbb{E}(Y_t) = \alpha + \beta t$ dépend du temps. Le processus Y_t peut être stationnarisé en retranchant à Y_t la valeur estimée $\hat{\alpha} + \hat{\beta}t$ par la méthode des Moindres Carrés Ordinaires (MCO).

b/ Processus DS

Définition 1.2.6 (*Processus DS*)

Un processus (Y_t) est un DS s'il admet une racine unitaire ou si les racines de son polynôme caractéristique sont supérieures à 1 en valeur absolue. Ces processus se rendent stationnaires par différentiation (ou filtre des différences). Il y a trois types de processus DS :

- Processus DS (Differency Stationary) avec tendance ($\beta \neq 0$) s'exprime comme suit :

$$Y_t = \alpha + \beta t + Y_{t-1} + \epsilon_t,$$

Il présente une non stationnarité de nature stochastique (il y a racine unitaire). Le processus DS avec tendance est non stationnaire. Par récurrence on peut avoir :

$$Y_t = Y_0 + \alpha t + \beta g(t) + \sum_{i=1}^t \epsilon_i, \quad \text{où } \epsilon_i \rightsquigarrow iid(0, \sigma_\epsilon^2)$$

Le processus DS avec dérive est non stationnaire car on a $\mathbb{E}(Y_t) = Y_0 + \alpha t + \beta g(t)$ qui dépend du temps t . Plus $t \rightarrow \infty$ et plus $\mathbb{E}(Y_t) \rightarrow \infty$.

- Processus DS (Differency Stationary) avec dérive ($\alpha \neq 0$) s'exprime comme suit :

$$Y_t = Y_{t-1} + \alpha + \epsilon_t,$$

Le processus DS avec dérive est appelé aussi marche au hasard (ou marche aléatoire) avec dérive.

Il présente une non stationnarité de nature stochastique.

Par récurrence, on obtient (dans le cas avec dérive) :

$$Y_t = Y_0 + \alpha t + \sum_{i=1}^t \epsilon_i, \quad \text{où} \quad \epsilon_i \rightsquigarrow iid(0, \sigma_\epsilon^2)$$

Le processus DS avec dérive est non stationnaire car on a $\mathbb{E}(Y_t) = Y_0 + \alpha t$ qui dépend du temps t . Plus $t \rightarrow \infty$ et plus $\mathbb{E}(Y_t) \rightarrow \infty$.

- Processus DS (Differency Stationary) sans dérive ($\alpha = 0$) s'écrit :

$$Y_t = Y_{t-1} + \epsilon_t.$$

Le processus DS sans dérive est appelé aussi marche au hasard (ou marche aléatoire).

Par récurrence, on obtient (dans le cas sans dérive) :

$$\begin{aligned} Y_1 &= Y_0 + \epsilon_1, \\ Y_2 &= Y_1 + \epsilon_2 = Y_0 + \epsilon_1 + \epsilon_2 \\ &\vdots \\ Y_t &= Y_0 + \sum_{i=1}^t \epsilon_i, \quad \text{où} \quad \epsilon_i \rightsquigarrow iid(0, \sigma_\epsilon^2) \end{aligned}$$

Le processus DS sans dérive est non stationnaire car on a :

$$\text{Var}(Y_t) = \text{Var}\left(\sum_{i=1}^t \epsilon_i\right) = \sum_{i=1}^t \text{Var}(\epsilon_i) = \sum_{i=1}^t \sigma_\epsilon^2 = t\sigma_\epsilon^2.$$

La variance du processus DS sans dérive dépend alors du temps t .

Pour stationnariser le processus DS (avec ou sans dérive), il suffit de le passer en différence première :

$$Y_t - Y_{t-1} = \alpha + \epsilon_t \quad (\text{cas avec dérive})$$

$$Y_t - Y_{t-1} = \epsilon_t \quad (\text{cas sans dérive}).$$

Remarque 1.2.2 *En utilisant le logiciel Eviews, le graphe de la série brute peut nous offrir une idée générale si la série est stationnaire ou non comme aussi le test de Correlogram à l'aide des fonctions d'Autocorrelation et d'autocorrelation Partielle). Le principe du test de Correlogram est le suivant*

$$\begin{cases} H_0 : \text{la série est stationnaire} \\ H_1 : \text{la série n'est pas stationnaire.} \end{cases}$$

Mais pour confirmer le type de la non stationnarité des quatre type cités en hauts, TS et les trois types DS, la stratégie séquentielle du test de Dickey Fuller Augmenté (ADF) ([15]) est mieux à suivre. Dans un logiciel comme Eviews par exemple la stratégie est présentée comme suit :

- La première étape : de commencer par le premier type, de tester si la tendance est significative c'est à dire,

$$Y_t = Y_{t-1} + \underline{\beta}t + \epsilon_t,$$

sur Eviews c'est le modèle M3 "**Trend and intercept**".

- La deuxième étape (si la tendance n'est pas significative) : de tester significativité de la constante et dans l'absence de la tendance (ou modèle M2),

$$Y_t = Y_{t-1} + \underline{\alpha} + \epsilon_t,$$

sur Eviews c'est "**Intercept**".

- La troisième étape (si ni la tendance ni la constante est significative) : de tester l'absence de tendance et de constante,

$$Y_t = Y_{t-1} + \epsilon_t,$$

sur Eviews c'est le modèle M1 "**None**".

2. Tests de stationnarité (ou tests de racine unitaire)

Il existe plusieurs tests de racine unitaire : tests de Dickey-Fuller simple et Dickey-Fuller Augmenté, test de Phillips et Perron, test de Kwiatkowski, Phillips, Schmidt et Shin (test de KPSS). Nous n'étudierons ici que les tests de Dickey-Fuller et de Phillips-Perron.

i) Test de Dickey-Fuller simple Le test de Dickey-Fuller permet de savoir si une série est stationnaire ou non et permet aussi de déterminer la bonne manière de stationnariser la série.

Les hypothèses du test sont les suivantes :

$$\left\{ \begin{array}{l} H_0 : \text{processus non stationnaire, il correspond à une de ces formes de non stationnarité :} \\ \quad \text{[1]} \quad y_t = \phi_1 y_{t-1} + \epsilon_t \\ \quad \text{[2]} \quad y_t = \phi_1 y_{t-1} + c + \epsilon_t \\ \quad \text{[3]} \quad y_t = \phi_1 y_{t-1} + bt + c + \epsilon_t \\ \quad \text{où } \phi_1 = 1 \quad \text{et} \quad \epsilon_t \rightsquigarrow iid(0, \sigma_\epsilon^2) \\ H_1 : |\phi_1| < 1, \quad \text{processus est stationnaire} \end{array} \right.$$

on peut réécrire ces hypothèses sous la forme :

$$\left\{ \begin{array}{l} H_0 : \text{processus non stationnaire, il correspond à une de ces formes de non stationnarité :} \\ \quad \text{[1]} \quad \Delta y_t = (\phi_1 - 1) y_{t-1} + \epsilon_t \\ \quad \text{[2]} \quad \Delta y_t = (\phi_1 - 1) y_{t-1} + c + \epsilon_t \\ \quad \text{[3]} \quad \Delta y_t = (\phi_1 - 1) y_{t-1} + bt + c + \epsilon_t \\ \quad \text{où } (\phi_1 - 1) = 0 \quad \text{et} \quad \epsilon_t \rightsquigarrow iid(0, \sigma_\epsilon^2) \\ H_1 : \phi_1 < 1, \quad \text{processus stationnaire} \end{array} \right.$$

Sous H_0 vraie, la statistique de test pour l'estimateur de ϕ_1 est donnée par :

$$t_{\hat{\phi}_1} = \frac{\hat{\phi}_1 - 1}{\hat{\sigma}_{\hat{\phi}_1}^2}.$$

Dans la stratégie de Dickey-Fuller on commence par étudier le modèle général [3]. On regarde si b est significativement différent de 0 ou non. Si b est non significativement différent de 0, on passe à l'étude du modèle [2] et on cherche à savoir si c est significativement différent de 0 ou pas. Si c est non significativement différent de 0, on étudie le modèle [1].

Il est à noter que sous H_0 vraie, les t de Student de la constante et de la tendance sont à comparer avec les valeurs de la table de Dickey-Fuller. Par exemple : pour une taille d'échantillon

supérieure à 500 observations, les valeurs critiques sont : 2.78 à 5% pour la tendance du modèle [3], 2.52 pour la constante du modèle [2] et -1.95 pour le paramètre ϕ_1) car sous H_0 vraie le processus étudié est non stationnaire ($y_t \rightsquigarrow I(1)$) et l'estimateur de ϕ_1 ne suit pas la loi normale.

Les règles de décision sont les suivantes :

- Si $t_{\phi_1} > t_{DF}$ où t_{DF} désigne la valeur critique donnée par la table de *Dickey – Fuller*, donc on accepte H_1 : le coefficient de la variable explicative est significativement différent de 0. Si on a b significativement différent de 0 pour le modèle [3], le test s'arrête ici, on n'étudie pas les autres modèles. De même que si on arrive au modèle [2] et que l'on a la constante c qui est significativement différente de 0, le test s'arrête au modèle [2].
- Si $t_{\phi_1} < t_{DF}$, donc on accepte H_0 : la série est non stationnaire.

ii) Test de Dickey-Fuller Augmenté (ADF) Dans le test de Dickey-Fuller que nous venons d'étudier, le processus ϵ_t est par hypothèse un bruit blanc. Or il n'y a aucune raison pour que, a priori, l'erreur soit non corrélée. Le test de Dickey-Fuller Augmenté ne suppose pas que ϵ_t est un bruit blanc.

Les hypothèses du test de Dickey-Fuller Augmenté se définissent de la façon suivante :

$$\left\{ \begin{array}{l} H_0 : \text{processus non stationnaire, il correspond à une de ces formes de non stationnarité :} \\ \text{[1]} \quad \Delta y_t = \rho y_{t-1} - \sum_{k=2}^p \gamma_k \Delta y_{t-k+1} + \eta_t \\ \text{[2]} \quad \Delta y_t = \rho y_{t-1} - \sum_{k=2}^p \gamma_k \Delta y_{t-k+1} + c + \eta_t \\ \text{[3]} \quad \Delta y_t = \rho y_{t-1} - \sum_{k=2}^p \gamma_k \Delta y_{t-k+1} + bt + c + \eta_t \\ \text{où } \rho = 0, \quad \phi_1 = 1 \quad \text{et} \quad \eta_t \rightsquigarrow iid(0, \sigma_\epsilon^2) \\ H_1 : \phi_1 < 1, \quad \text{processus stationnaire} \end{array} \right.$$

tel que $\rho = (\phi_1 - 1)(1 - \theta_1 - \theta_2 - \dots - \theta_{p-1})$ et $\gamma_k = \sum_{i=k+1}^p \alpha_i$, où $1 \leq k \leq p-1$ et $\alpha_1 = \phi_1 + \theta_1, \alpha_2 = \theta_2 - \phi_1 \theta_1, \dots, \alpha_p = \theta_p - \phi_1 \theta_{p-1}$

Cette forme de modèles obtenue en faisant les démarches suivantes :

On a $y_t = \phi_1 y_{t-1} + \epsilon_t$ où $\epsilon_t \rightsquigarrow AR(p-1)$, ϵ_t n'est pas un bruit blanc : $\epsilon_t = \sum_{i=1}^{p-1} \theta_i \epsilon_{t-i} + \eta_t$ où $\eta_t \rightsquigarrow iid(0, \sigma^2)$.

On peut écrire ces équations à l'aide de l'opérateur de décalage B tel que $By_t = y_{t-1}$ et $B^i y_t = y_{t-i}$. On obtient alors

$$(1 - \phi_1 B) y_t = \epsilon_t,$$

d'un part, et d'autre part on a

$$\epsilon_t = \theta_1 B \epsilon_t + \theta_2 B^2 \epsilon_t + \dots + \theta_{p-1} B^{p-1} \epsilon_t + \eta_t \iff \left(1 - \sum_{i=1}^{p-1} \theta_i B^i\right) \epsilon_t = \eta_t.$$

En remplaçant, dans ce dernier résultat, ϵ_t par $(1 - \phi_1 B) y_t$, on obtient alors

$$\left(1 - \sum_{i=1}^{p-1} \theta_i B^i\right) (1 - \phi_1 B) y_t = \eta_t$$

alors

$$\left(1 - \sum_{i=1}^{p-1} \theta_i B^i\right) (y_t - \phi_1 y_{t-1}) = \eta_t,$$

cet équation, après un simple calcul, devient

$$y_t = (\phi_1 + \theta_1) y_{t-1} + (\theta_2 - \phi_1 \theta_1) y_{t-2} + \dots + (\theta_{p-1} - \phi_1 \theta_{p-2}) y_{t-p+1} + \phi_1 \theta_{p-1} y_{t-p} + \eta_t$$

ou encore

$$\Delta y_t = (\phi_1 - 1) (1 - \theta_1 - \theta_2 - \dots - \theta_{p-1}) y_{t-1} + \dots + \sum_{k=1}^{p-1} \gamma_k \Delta y_{t-k} + \eta_t$$

iii) Détermination du retard p du test ADF La valeur p est déterminée à l'aide du corrélogramme partiel de la série différenciée Δy_t .

Une fois déterminée la valeur p , on procède de la même façon qu'avec le test de Dickey Fuller simple : on commence par étudier la significativité de b du modèle [3]. La règle de décision est la même que pour le test de DF simple.

La statistique de test pour l'estimateur de ϕ_1 est :

$$t_{\hat{\phi}_1} = \frac{\hat{\phi}_1 - 1}{\hat{\sigma}_{\hat{\phi}_1}^2}.$$

qui est à comparer avec la valeur critique t_{DF} de la table de Dickey-Fuller.

Si $t_{\phi_1} > t_{DF}$, alors on accepte H_0 : la série est non stationnaire (ATTENTION : il faut observer comme dans le cas du test de DF simple que pour $t_{\phi_1} > t_{DF}$, on n'a pas H_1 ! La règle de décision est inversée ici !).

1.3 Modèles de volatilité

1.3.1 Séries financières

La modélisation des séries financières est un problème majeur et difficile. Cette difficulté n'est pas seulement due à la grande variété des séries utilisées (prix d'action, taux d'intérêt, taux de change...), à l'importance de la fréquence d'observation (seconde, minute, heure, jour...) ou à la disponibilité d'échantillon de très grande taille. Elle tient surtout à l'existence des régularités statistiques (faits stylisés) communes à un très grand nombre de séries financières et difficiles à reproduire artificiellement à partir de modèle stochastique.

Plusieurs auteurs mettaient en évidence un ensemble de telles propriétés. Ces régularités empiriques, apparaissent plus ou moins nettement en fonction de la fréquence d'observation de la série et de sa nature. Les propriétés que nous citons ci-dessous sont principalement valables pour des séries quotidiennes de cours d'action boursier.

Soit p_t le cours d'un actif à la date t et $\epsilon_t = \log(p_t/p_{t-1})$ le logarithme de rendement (également appelé rendement ou risque auprès des financiers, les gestionnaires des portefeuilles et les économètres). Notons que la série (ϵ_t) est souvent proche de la série du rendement arithmétique $R_t = (p_t - p_{t-1})/p_{t-1}$ calculé ou utilisé auprès de l'univers des comptables, en général $\epsilon_t \simeq R_t$. Ces propriétés ont été très commentées dans la littérature sont les suivantes :

- i) Non stationnarité de la série des prix : les séries économiques et financières présentent de grandes fluctuations, voir Figure 1 des prix de pétrole brut (WTI) en chapitre 4. Cependant, les séries de rendements ont des trajectoires compatibles avec la stationnarité de second ordre. On constate par exemple sur la Figure 3 que la série $\epsilon_t = \log(p_t/p_{t-1})$, où p_t désigne le cours du pétrole est constituée d'oscillations autour de zéro.
- ii) Absence d'autocorrélation des variations de prix : la série des variations de prix présente de

faibles autocorrélations, elle est proche d'un bruit blanc.

iii) Autocorrélation des carrés des variations de prix : les séries des carrés (ϵ_t^2) ou des valeurs absolues ($|\epsilon_t|$) sont souvent fortement autocorrélées

iv) Regroupement des fortes variations de prix (volatility clustering). Les grandes valeurs de $|\epsilon_t|$ ou les fortes variations de prix, tendent à être suivies de grandes valeurs, et les petites de petites, qui se voit clairement sur le graphe de Figure 3.

v) Queues de distribution épaisses : les distributions de probabilité empiriques de séries de rendements quotidiens, ou de variations de prix, ou encore du logarithme de ces variations de prix, ne correspondent pas à une distribution gaussienne (diagnostique par le test de normalité).

vi) Effet de levier : La propriété d'asymétrie de la volatilité, la propriété découvert par Black (1976) ([3]) : Les valeurs négatives (baisses des cours) tendent à augmenter la volatilité supérieur à celui induit par des valeurs positives (hausse des cours).

iiiv) Saisonnalités : L'effet de saisonnier est très présent pour les séries financières intrajournalières.

1.3.2 Hétéroscédasticité

Les propriétés précédentes illustrent la difficulté de modéliser les séries financières. Les formulations classiques (de types ARMA) sont inappropriées car centrées sur la structure d'autocovariance des processus. Or, la plupart des séries de rendements ne diffèrent guère des bruits blancs faibles.

Le fait que les grandes valeurs des carrés des rendements soient généralement précédées par de grandes valeurs est difficilement compatible avec une variance conditionnelle constante. Ce phénomène est connu sous le nom d'hétéroscédasticité conditionnelle.

$$\text{Var}(\epsilon_t | \epsilon_{t-1}, \epsilon_{t-2}, \dots) \neq \text{cste}.$$

Les modèles économétriques qui prennent en compte les propriétés particulières des séries financières se présentent généralement sous la forme multiplicative :

$$\epsilon_t = \sigma_t \eta_t,$$

où, (η_t) et (σ_t) sont des processus tels que :

- i) σ_t est mesurable par rapport à une tribu, notée \mathcal{F}_{t-1} , contenant le passé de ϵ_t .
- ii) (η_t) est un processus *i.i.d.*, centré, de variance unité et η_t est indépendant de \mathcal{F}_{t-1} .
- iii) $\sigma_t > 0$

La variable aléatoire σ_t est appelée *volatilité* de ϵ_t . Alors le signe de la variation des prix (signe de ϵ_t) est celui de η_t .

Notons que

$$\mathbb{E}(\epsilon_t) = \mathbb{E}(\sigma_t) \mathbb{E}(\eta_t) = 0$$

et

$$\text{Cov}(\epsilon_t, \epsilon_{t-h}) = \mathbb{E}(\eta_t) \mathbb{E}(\sigma_t \epsilon_{t-h}) = 0, \quad \forall h > 0$$

ce qui fait de (ϵ_t) un bruit blanc faible. La série des carrés présente en revanche des autocorrélations non nulles, alors (ϵ_t) n'est donc pas un bruit blanc fort. En fait, l'outil principal pour vérifier cette propriété importante (autocorrélation ou non des carrés des résidus) se fait par le test de Multiplicateur de Lagrange introduit par Robert Engle (1982), qui connut sous le nom "test d'homoscédasticité" ou "test d'hétérosécédasticité". Il a de même principe que le test BDS qui est porté sur les résidus (on dit autocorrélation), mais celui-ci est porté sur les carrés des résidus (on dit hétérosécédasticité), il est présenté comme suit :

Test LM (test d'homoscédasticité)

Il dit aussi "Test de l'effet d'ARCH", basé sur le Multiplicateur de Lagrange (*LM*), il consiste en une régression de ϵ_t^2 sur une constante α_0 et sur ses valeurs retardés des carrés des résidus jusqu'à l'ordre p : $\epsilon_{t-1}^2, \epsilon_{t-2}^2, \dots, \epsilon_{t-p}^2$,

$$\epsilon_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \alpha_2 \epsilon_{t-2}^2 + \dots + \alpha_p \epsilon_{t-p}^2 + e.$$

Nous testons l'hypothèse nulle de l'homoscédasticité H_0 :

$$\alpha_0 = \alpha_1 = \alpha_2 = \dots = \alpha_p = 0.$$

La statistique du test est :

$$LM(p) = nR^2 \stackrel{H_0}{\rightsquigarrow} \chi^2(p),$$

où n est la taille de l'échantillon étudié.

La règle de décision est : - Si $LM(p) < \chi_{1-\alpha/2}^2(p)$, alors H_0 est acceptée, il y a donc absence d'hétéroscédasticité, i.e., pas d'autocorrélation entre les carrés des résidus.

- Si non, si $LM(p) > \chi_{1-\alpha/2}^2(p)$, alors H_0 est rejetée et il y a présence d'hétéroscédasticité, i.e., les autocorrélations entre les carrés des résidus, le modèle choisi alors peut être retenu.

Le résultat de ce test peut conclure par la probabilité associée, comparée à 0.05 en général.

Remarque 1.3.1 Avant de procéder au test, le corrélogramme des carrés des résidus (*Correlogram of Residuals Squared*) peut nous donner une idée sur la présence de l'autocorrélation et l'ordre p à travers la colonne de fonction d'autocorrélation partielle (*Partial Correlation*).

1.3.3 Modèles ARCH (*AutoRegressive Conditional Heteroskedasticity*)

Le modèle *ARCH* a été présenté pour la première fois par Robert Engle en 1982 pour modéliser la série des résidus (ϵ_t) du modèle ARMA suivant, quand ces résidus ne sont pas un bruit blanc (propriété peut vérifier par le test *BDS* dont H_0 : les résidus du modèle ARMA(P, Q) *estimé* sont *i.i.d.*),

$$r_t = \mathbb{E}(r_t | I_{t-1}) + \epsilon_t.$$

ou

$$r_t = \lambda + \sum_{i=1}^P \phi_i r_{t-i} + \sum_{j=1}^Q \theta_j \epsilon_{t-j} + \epsilon_t \quad (1.1)$$

où le comportement de ϵ_t peut être décrit par

$$\epsilon_t = \sigma_t \eta_t,$$

avec $\eta_t \rightsquigarrow i.i.d. (0, 1)$ (bruit blanc).

La représentation ARMA considère que l'espérance conditionnelle (qui est l'objectif de la modélisation Box-Jenkins) des rendements r_t de la série observée varie au cours du temps alors que la variance conditionnelle ne change pas (homoscédasticité). Par contre, le modèle ARCH

considère que la variance conditionnelle (σ_t^2) de ϵ_t d'ARMA est variable au cours du temps (hétéroscédasticité) et c'est une fonction des carrés des erreurs ou des innovations passés jusqu'au retard q tant que donnée par la représentation autorégressive (1.2) de la définition suivante :

Définition 1.3.1 (Processus ARCH d'ordre q)

On dit que (ϵ_t) est un processus ARCH (q) si ses deux premiers moments conditionnels vérifient :

(i) $\mathbb{E}(\epsilon_t/\epsilon_{t-1}) = 0, \forall t \in \mathbb{Z},$

(ii) Il existe des constantes $\omega > 0, \alpha_i \geq 0, i = 1, \dots, q$ et $\beta_j \geq 0, j = 1, \dots, p$, telles que :

$$\mathbb{V}ar(\epsilon_t/\epsilon_{t-1}, \epsilon_{t-2}, \dots) = \sigma_t^2 = \omega + \sum_{i=1}^q \alpha_i \epsilon_{t-i}^2 \quad (1.2)$$

où ω est une constante strictement positive, α_i (paramètres) $\geq 0, i = 1, \dots, q, q$ est le retard et σ_t est la volatilité comme mentionné ci-dessus (dans la sous-section précédente). On a les propriétés suivantes :

$$\mathbb{E}(\epsilon_t/I_{t-1}) = 0; \quad \omega > 0,$$

Théorème 1.3.1 (Stationnarité au second ordre du ARCH(q))

Un processus (ϵ_t) de définition 1.3.1 est stationnaire si la condition suivant est vérifiée :

$$\sum_{i=1}^q \alpha_i < 1.$$

Cette stationnarité est de second ordre.

1.3.4 Modèles GARCH (Generalized ARCH)

A. Définitions et présentations

Le modèle GARCH est une extension du modèle ARCH présenté ci-dessus. Ce processus a les mêmes propriétés et les mêmes fondements que le processus mais sa formule qui lui caractérise est plus générale que celle d'ARCH. La volatilité est en plus une fonction des carrées de ses valeurs passées.

Définition 1.3.2 (Processus GARCH(p, q) semi fort)

On dit que (ϵ_t) est un processus $GARCH(p, q)$ si :

(i) $\mathbb{E}(\epsilon_t/\epsilon_{t-1}) = 0, \forall t \in \mathbb{Z},$

(ii) Il existe des constantes $\omega > 0, \alpha_i \geq 0, i = 1, \dots, q$ et $\beta_j \geq 0, j = 1, \dots, p,$ telles que :

$$\mathbb{V}ar(\epsilon_t/\epsilon_{t-1}) = \sigma_t^2 = \omega + \sum_{i=1}^q \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2. \quad (1.3)$$

Définition 1.3.3 L'équation (1.3) peut être écrite de manière symbolique plus compacte

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

où B est l'opérateur de retard vérifiant $(B^i \epsilon_t^2 = \epsilon_{t-i}^2$ et $B^j \sigma_t^2 = \sigma_{t-j}^2)$, et

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

Si $\beta(B) = 0$, le processus est applé $ARCH(q)$.

L'innovation du processus ϵ_t^2 est par définition la variable $v_t = \epsilon_t^2 - \sigma_t^2$. En remplaçant, dans l'équation (1.3), les variables σ_{t-j}^2 par $\epsilon_{t-j}^2 - v_{t-j}$ on obtient :

$$\epsilon_t^2 = \omega + \sum_{i=1}^r (\alpha_i + \beta_i) \epsilon_{t-i}^2 + v_t - \sum_{j=1}^p \beta_j v_{t-j}, \quad t \in \mathbb{Z} \quad (1.4)$$

où $r = \max(p, q)$. Cette structure ressemble le processus $ARMA$. Donc on peut confirmer que sous certaines conditions si (ϵ_t) est un $GARCH(p, q)$ alors ϵ_t^2 est un $ARMA(r, p)$. Ces représentations de $ARMA$ seront utiles pour **l'estimation** et **l'identification** des processus $GARCH$.

Remarque 1.3.2 (Corrélation des carrés d'un $GARCH$)

Il est discuté dans la sous-section 1.3.1 qu'une propriété caractéristique des séries financières que le fait les rendements ϵ_t sont non corrélés, tandis que les carrés des rendements sont bien autocorrélés. En effet, les ϵ_t sont non corrélés vu (i) de définition 1.3.2 et la représentation (1.4) montre que les processus $GARCH$ sont bien adaptés à la prise en compte de cette propriété empirique. En utilisant cette formule pour donner la suite des autocorrélations d'ordre h de ϵ_t^2

vérifie une relation de récurrence caractéristique des modèles *ARMA*. Prenons pour simplifier un modèle *GARCH* (1, 1). Son carré (ϵ_t^2) est alors un *ARMA* (1, 1) son autocorrélation alors est

$$\text{Corr}(\epsilon_t^2, \epsilon_{t-h}^2) = K(\alpha_1 + \beta_1)^h, \quad h > 1 \text{ et } K \text{ est une constante.}$$

Cette autocorrélation tend vers zéro quand h est grand.

Définition 1.3.4 (Processus *GARCH*(p, q) au sens fort)

Un processus (ϵ_t) est dit processus *GARCH* au sens fort s'il vérifie

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

où $\omega > 0$, $\alpha_i \geq 0$ et $\beta_j \geq 0$. (η_t) une suite de variables *i.i.d.*

En remplaçant ϵ_{t-i} par $\sigma_{t-i} \eta_{t-i}$ dans l'équation (2.5) on obtient

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

Dans l'équation (1.5), si on remplace ϵ_{t-i} par $\sigma_{t-i} \eta_{t-i}$ on aboutit à la représentation (1.6)

$$\sigma_t^2 = \omega + \sum_{i=1}^r a_i (\eta_{t-i}) \sigma_{t-i}^2, \quad (1.6)$$

où $a_i (\eta_{t-i}) = \alpha_i \eta_{t-i}^2 + \beta_i$, $i = 1, \dots, r = \max(p, q)$. Cette équation montre que, dans le cas d'un *GARCH* (p, q) fort, le processus de volatilité (σ_t^2) vérifie une équation autorégressive *AR* (r), mais avec des coefficients aléatoires.

B. Étude de stationnarité :

Dans le cas du processus *GARCH*(p, q) fort, l'écriture vectorielle suivante sera très utile. On a

$$z_t = \underline{b}_t + A_t z_{t-1}, \quad (1.7)$$

où

$$\underline{b}_t = \underline{b}_t(\eta_t) = \begin{pmatrix} \alpha_0 \eta_t^2 \\ 0 \\ \vdots \\ \omega \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in \mathbb{R}^{p+q}, \quad z_t = \begin{pmatrix} \epsilon_t^2 \\ \vdots \\ \epsilon_{t-q+1}^2 \\ \sigma_t^2 \\ \vdots \\ \sigma_{t-q+1}^2 \end{pmatrix} \in \mathbb{R}^{p+q}$$

et

$$A_t = \begin{pmatrix} \alpha_1 \eta_t^2 & \cdots & \alpha_q \eta_t^2 & \beta_1 \eta_t^2 & \cdots & \beta_p \eta_t^2 \\ 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 1 & 0 & 0 & \cdots & 0 \\ \alpha_1 & \cdots & \alpha_q & \beta_1 & \cdots & \beta_p \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 1 & 0 \end{pmatrix} \quad (1.8)$$

est une matrice de dimension $(p+q) \times (p+q)$. L'équation (1.7) constitue un modèle vectoriel autorégressif d'ordre un, avec coefficients positifs et *i.i.d.* 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 (1.7) on obtient

$$\underline{z}_t = \underline{b}_t + \sum_{k=1}^{\infty} A_t A_{t-1} \cdots A_{t-k+1} \underline{b}_{t-k} \quad (1.9)$$

B.1 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

$$\lim_{t \rightarrow \infty} \frac{1}{t} \log \|A^t\| = \log \rho(A) \quad (1.10)$$

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

$$\lim_{t \rightarrow \infty} \frac{1}{t} \mathbb{E}(\log \|A_t A_{t-1} \cdots A_1\|) = \gamma = \inf_{t \in \mathbb{N}^*} \frac{1}{t} \mathbb{E}(\log \|A_t A_{t-1} \cdots A_1\|) \quad (1.11)$$

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

$$\gamma = \lim_{t \rightarrow \infty} p.s. \frac{1}{t} \log \|A_t A_{t-1} \cdots A_1\| \quad (1.12)$$

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

$$\lim_{t \rightarrow \infty} p.s. \|A_0 \dots A_{-t}\| = 0 \Rightarrow \gamma < 0 \quad (1.13)$$

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

Théorème 1.3.2 (Stationnarité stricte du modèle GARCH(p, q))

Une condition nécessaire et suffisante d'existence d'un processus GARCH(p, q) strictement stationnaire, solution du modèle (1.5) est que

$$\gamma < 0 \quad (1.14)$$

où γ est le plus grand exposant de Lyapounov de la suite $\{A_t, t \in \mathbb{Z}\}$ définie par (1.8).

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

Preuve 1.3.1 En cette preuve utilisons la norme définie par $\|A\| = \sum_i |a_{ij}|$. Remarquons que les variables η_t étant de variance finie, tous les termes de la matrice A_t sont intégrables. On a donc

$$\mathbb{E} \log^+ \|A_t\| \leq \mathbb{E} \|A_t\| < \infty.$$

Supposons d'abord que $\gamma < 0$. Donc la suite

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

converge presque sûrement lorsque N tend vers l'infini, vers une limite notée \tilde{z}_t . On a en effet, en utilisant la multiplicité de la norme,

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

et

$$\begin{aligned} \|A_t A_{t-1} \dots A_{t-n}\|^{\frac{1}{n}} \|\underline{b}_{t-n-1}\|^{\frac{1}{n}} &= \exp \left[\frac{1}{n} \log \|A_t A_{t-1} \dots A_{t-n}\| + \frac{1}{n} \log \|\underline{b}_{t-n-1}\| \right] \\ &\xrightarrow{p.s.} e^\gamma < 1. \end{aligned}$$

Pour montrer que $\frac{1}{n} \log \|\underline{b}_{t-n-1}\| \longrightarrow 0$, p.s. on a

$$\mathbb{E} \log \|\underline{b}_{t-n-1}\| \leq \log^- \omega + \mathbb{E} \log^+ \|\underline{b}_{t-n-1}\| \leq \log^- \omega + \mathbb{E} \|\underline{b}_{t-n-1}\| < \infty$$

Par suite, par la règle de Cauchy, \tilde{z}_t est bien défini dans $(\mathbb{R}^{*+})^{p+q}$ et coïncide avec le nombre de droite de (1.9). Soit $\tilde{z}_{q+1,t}$ la $q+1$ -ème composante de \tilde{z}_t . En posant $\epsilon_t = \sqrt{\tilde{z}_{q+1,t}}$ on définit une solution du modèle (1.5). Cette solution est non anticipative car d'après (1.9), ϵ_t s'exprime en fonction mesurable de $\eta_t, \eta_{t-1}, \dots$. Cette solution est strictement stationnaire et ergodique puisque (η_t) est ergodique.

L'unicité se démontre par le même raisonnement que dans le cas $p = q = 1$. Soit la solution strictement stationnaire de (1.5), ou bien une solution (z_t) strictement stationnaire de (1.7). Pour tout $N \geq 1$,

$$z_t = \tilde{z}_t(N-1) + A_t \dots A_{t-N} z_{t-N-1}.$$

Alors

$$\|z_t - \tilde{z}_t\| \leq \|\tilde{z}_t(N) - \tilde{z}_t\| + \|A_t \dots A_{t-N}\| \|\tilde{z}_{t-N-1}\|$$

Le premier terme à droite de l'inégalité tend vers 0 p.s. quand $N \longrightarrow \infty$. Par ailleurs, puisque la série définissant \tilde{z}_t converge p.s., on a $\|A_t \dots A_{t-N}\| \longrightarrow 0$ avec probabilité 1 quand $N \longrightarrow \infty$. De plus $\|\tilde{z}_{t-N-1}\|$ est indépendante de N par stationnarité. Par suite $\|A_t \dots A_{t-N}\| \|\tilde{z}_{t-N-1}\| \longrightarrow 0$ en probabilité lorsque $N \longrightarrow \infty$. Donc on a montré que $z_t - \tilde{z}_t \longrightarrow 0$ en probabilité si $N \longrightarrow \infty$. Ce terme est indépendant de N , on a donc nécessairement $\tilde{z}_t = z_t$ pour tout t p.s.

On montre finalement la condition nécessaire du théorème. D'après le lemme 1.3.1, il suffit d'établir la partie gauche de (1.12). Nous allons montrer que pour $1 \leq i \leq p+q$

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

où e_i est le i -ème élément de la base canonique de \mathbb{R}^{p+q} . Soit (ϵ_t) une solution strictement stationnaire de (1.5) et soit (z_t) défini par (1.7). 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}_t et \underline{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 pour $k \rightarrow \infty$. Or $\underline{b}_{-k-1} = \omega \eta_{-k-1}^2 e_1 + \omega e_{q+1}$. Donc $A_0 \dots A_{-k} \underline{b}_{-k-1}$ se décompose en deux termes positifs et on a

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

Puisque $\omega \neq 0$, la **limite** " $\lim_{t \rightarrow \infty} A_0 \dots A_{-t} e_i = 0$, p.s." est alors vraie pour $i = q + 1$. On utilise la relation

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

avec par convention $e_{p+q+1} = 0$. Pour $i = 1$ cette relation donne

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

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

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

est une récurrence ascendante. Donc, le théorème est démontré. □

B.2 Stationnarité au second ordre :

Théorème 1.3.3 (Stationnarité au second ordre du modèle GARCH(p, q))

S'il existe un processus GARCH(p, q), au sens de la définition 1.3.2, stationnaire au second ordre et non anticipatif, et si $\omega > 0$, alors

$$\sum_{i=1}^q \alpha_i + \sum_{j=1}^p \beta_j < 1. \quad (1.15)$$

Inversement, si (1.15) 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.

1.3.5 Modèles GARCH exponentiel (EGARCH)

A. Définitions et présentations :

Définition 1.3.5 (*Processus EGARCH(p,q)*)

Soit (η_t) une suite de variables iid, centrées et réduites ($\mathbb{E}(\eta_t) = 0$ et $\text{Var}(\eta_t) = 1$). Le processus (ϵ_t) est dit un processus GARCH exponentiel (exponential GARCH) s'il vérifie une équation de la forme

$$\begin{cases} \epsilon_t = \sigma_t \eta_t \\ \log \sigma_t^2 = \omega + \sum_{i=1}^q \alpha_i g(\eta_{t-i}) + \sum_{j=1}^p \beta_j \log \sigma_{t-j}^2. \end{cases} \quad (1.16)$$

où

$$g(\eta_{t-i}) = \theta \eta_{t-i} + h(|\eta_{t-i}| - \mathbb{E}|\eta_{t-i}|)$$

$\omega > 0$, α_i , β_j , θ et h sont des réels.

1. Ce type de modèle est connu sous le nom d'asymétrie. Cependant, la positivité des coefficients n'est pas exigée ce qui rend EGARCH plus général que GARCH.
2. Une autre différence par rapport GARCH classique ou symétrique réside dans l'écriture de la variance conditionnelle comme fonction des innovations normalisées (divisées par leur écart-type conditionnel), plutôt que fonction des innovations passées.

B. Stationnarité

Théorème 1.3.4 (*Stationnarité de processus EGARCH(p,q)*)

On suppose que $g(\eta_t)$ n'est pas presque sûrement nulle et que $\alpha(z) = \sum_{i=1}^q \alpha_i z^i$ et $\beta(z) = 1 - \sum_{j=1}^p \beta_j z^j$ n'ont pas de racines communes, avec $\alpha(z)$ non identiquement nulle. Alors, le modèle EGARCH(p,q) défini par (1.16) admet une solution strictement stationnaire et non anticipative si et seulement si les racines de $\beta(z)$ sont de modules strictement supérieure à 1. Cette solution vérifie $\mathbb{E}(\log \epsilon_t^2)^2 < \infty$ dès que $\mathbb{E}(\log \eta_t^2)^2 < \infty$ et $\mathbb{E}(g^2(\eta_t)) < \infty$. Si de plus

$$\prod_{i=1}^{\infty} \mathbb{E} \exp(\lambda_i g(\eta_t)) < \infty \quad (1.17)$$

tels que les λ_i sont définies par $\frac{\alpha(L)}{\beta(L)} = \sum_{i=1}^{\infty} \lambda_i L^i$, alors (ϵ_t) est un bruit blanc de variance

$$\mathbb{E}(\epsilon_t^2) = \mathbb{E}(\sigma_t^2) = e^{\omega^*} \prod_{i=1}^{\infty} (g_{\eta}(\lambda_i))$$

où $\omega^* = \frac{\omega}{\beta(1)}$ et $g_{\eta}(x) = \mathbb{E} \exp [xg(\eta_t)]$.

Preuve 1.3.2 On a $\log \epsilon_t^2 = \log \sigma_t^2 + \log \eta_t^2$. Puisque $\log \sigma_t^2$ suit un modèle ARMA($p, q-1$) dont le polynôme AR est β , les hypothèses sur les polynômes retards sont nécessaires et suffisantes pour exprimer de manière unique $\log \sigma_t^2$ en fonction de η_{t-1} et de son passé infini sous forme moyenne mobile infinie:

$$\log \sigma_t^2 = \omega^* + \sum_{i=1}^{\infty} \lambda_i g(\eta_{t-i}), \quad p.s.$$

Les processus $(\log \epsilon_t^2)$ et $(\log \sigma_t^2)$ sont par suite strictement stationnaires. Le processus $\log \sigma_t^2$ est stationnaire au second ordre et également $(\log \epsilon_t^2)$ sous l'hypothèse $E(\log \eta_t^2)^2 < \infty$. De plus d'après le développement précédent

$$\epsilon_t^2 = \sigma_t^2 \eta_t^2 = e^{\omega^*} + \prod_{i=1}^{\infty} \exp [\lambda_i g(\eta_{t-i})] \eta_t^2, \quad p.s.$$

En utilisant le fait que le processus $g(\eta_t)$ est iid, on obtient le résultat sur l'espérance de ϵ_t^2 .

□

1.3.6 Modèles GARCH à seuil (TGARCH)

A. Définitions et présentations :

Le modèle de GARCH à seuil (threshold GARCH) découvert par Glosten, Jagannathan et Runkle (1993) ([25]) est une façon d'introduire l'asymétrie et de spécifier la variance conditionnelle (volatilité) en fonction des composantes positives et négatives des innovations passées.

Notons

$$\epsilon_t^+ = \max(\epsilon_t, 0), \quad \text{et} \quad \epsilon_t^- = \min(\epsilon_t, 0)$$

ces composantes, en remarquant que $\epsilon_t = \epsilon_t^+ + \epsilon_t^-$. La classe suivante introduit un effet de seuil dans la volatilité.

Définition 1.3.6 (*Processus TGARCH(p,q)*)

Soit (η_t) une suite de variables iid, centrées et réduites ($\mathbb{E}(\eta_t) = 0$ et $\text{Var}(\eta_t) = 1$). Le processus (ϵ_t) est dit un processus GARCH à seuil (Threshold GARCH(p,q) s'il vérifie une équation de la forme

$$\begin{cases} \epsilon_t = \sigma_t \eta_t \\ \sigma_t = \omega + \sum_{i=1}^q [\alpha_{i,+} \epsilon_{t-i}^+ - \alpha_{i,-} \epsilon_{t-i}^-] + \sum_{j=1}^p \beta_j \sigma_{t-j}. \end{cases} \quad (1.18)$$

où ω , $\alpha_{i,+}$, $\alpha_{i,-}$ et β_j sont des réels.

Cette représentation s'interprète comme l'écart-type conditionnelle de ϵ_t . Dans tous les cas, l'écart-type conditionnelle de ϵ_t est $|\sigma_t|$. Il n'est pas nécessaire d'imposer la positivité de σ_t comme le cas des modèles GARCH classiques fondés sur la modélisation de σ_t^2 . Mais on peut revenir à mettre au carré les variables aléatoires de l'équation (1.18), c'est une variante où on pose

$$\begin{cases} \epsilon_t = \sigma_t \eta_t \\ \sigma_t^2 = \omega + \sum_{i=1}^q \left(\alpha_i \epsilon_{t-i}^2 - \gamma_{i,-} \epsilon_{t-i}^2 I_{[\epsilon_{t-i}^- > 0]} \right) + \sum_{j=1}^p \beta_j \sigma_{t-j}^2. \end{cases} \quad (1.19)$$

Dans le cas de positivité de tous les coefficients du (1.18) on a

$$\epsilon_t^+ = \sigma_t \eta_t^+, \quad \epsilon_t^- = \sigma_t \eta_t^- \quad (1.20)$$

ce qui permet d'écrire l'écart-type conditionnel sous la forme

$$\sigma_t = \omega + \sum_{i=1}^{\max(p,q)} a_i(\eta_{t-i}) \sigma_{t-i} \quad (1.21)$$

où $a_i(\eta_{t-i}) = \alpha_{i,+} \eta_{t-i}^+ - \alpha_{i,-} \eta_{t-i}^- + \beta_i$, $i = 1, \dots, \max(p, q)$. La dynamique de σ_t est donc autorégressive à coefficients aléatoires.

B. Stationnarité :

L'étude de la stationnarité repose sur une représentation analogue à (1.7) obtenue en remplaçant dans le vecteur \underline{z}_{t-i} les variables ϵ_{t-i}^2 par $\begin{pmatrix} \epsilon_{t-i}^+ \\ -\epsilon_{t-i}^- \end{pmatrix}$ et les σ_{t-i}^2 par σ_{t-i} et en modifiant de manière adéquate \underline{b}_t et A_t . Plus précisément, en utilisant (1.20) on obtient

$$\underline{z}_t = \underline{b}_t + A_t \underline{z}_{t-1} \quad (1.22)$$

où

$$\underline{b}_t = \underline{b}(\eta_t) = \begin{pmatrix} \omega\eta_t^{+-} \\ -\omega\eta_t^{--} \\ 0 \\ \vdots \\ \omega \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in \mathbb{R}^{p+2q}, \quad \underline{z}_t = \begin{pmatrix} \epsilon_t^+ \\ -\epsilon_t^- \\ \vdots \\ \epsilon_{t-q+1}^+ \\ \epsilon_{t-q+1}^- \\ \sigma_t \\ \vdots \\ \sigma_{t-p+1} \end{pmatrix} \in \mathbb{R}^{p+2q}$$

et

$$A_t = \begin{pmatrix} \eta_t^{+-}\alpha_{1:q-1} & \alpha_{q,+}\eta_t^{+-} & \alpha_{q,-}\eta_t^{+-} & \eta_t^{+-}\beta_{1:p-1} & \beta_p\eta_t^{+-} \\ -\eta_t^{--}\alpha_{1:q-1} & -\alpha_{q,+}\eta_t^{--} & -\alpha_{q,-}\eta_t^{--} & -\eta_t^{--}\beta_{1:p-1} & -\beta_p\eta_t^{--} \\ I_{2q-2} & 0_{2q-2} & 0_{2q-2} & 0_{2q-2 \times p-1} & 0_{2q-2} \\ \alpha_{1:q-1} & \alpha_{q,+} & \alpha_{q,-} & \beta_{1:p-1} & \beta_p \\ 0_{p-1 \times 2q-2} & 0_{p-1} & 0_{p-1} & I_{p-1} & 0_{p-1} \end{pmatrix} \quad (1.23)$$

est une matrice de dimension $(p+2q) \times (p+2q)$,

$$\alpha_{1:q-1} = (\alpha_{1,+}, \alpha_{1,-}, \dots, \alpha_{q-1,+}, \alpha_{q-1,-}) \in \mathbb{R}^{2q-2},$$

$$\beta_{1:p-1} = (\beta_1, \dots, \beta_{p-1}) \in \mathbb{R}^{p-1}.$$

Le résultat suivant est analogue à celui obtenu pour la stationnarité stricte des modèles $GARCH(p, q)$.

Théorème 1.3.5 (*Stationnarité stricte du modèle TGARCH(p,q)*)

Le modèle TGARCH(p,q) définie par (1.18) – (1.19) admet une solution strictement stationnaire, non anticipative si

$$\gamma < 0.$$

γ le plus grand exposant de Lyapounov de la suite $\{A_t, t \in \mathbb{Z}\}$ définie par (1.23). Cette solution est unique et ergodique.

Preuve 1.3.3 *La condition suffisante de ce théorème est immédiate et similaire à celle du théorème 1.2.3. Pour la partie nécessaire on remarque que les termes des matrices A_t , \underline{b}_t et \underline{z}_t sont positifs. Ceci permet d'obtenir comme précédemment que $A_0 \dots A_{-k}\underline{b}_{-k-1}$ tend presque sûrement vers 0 quand $k \rightarrow \infty$. Or $\underline{b}_{-k-1} = \omega\eta_{-k-1}^+e_1 - \omega\eta_{-k-1}^-e_2 + \omega e_{2q+1}$. Donc, en utilisant*

la positivité, on a

$$\begin{aligned} \lim_{k \rightarrow \infty} A_0 \dots A_{-k} \omega \eta_{-k-1}^+ e_1 &= \lim_{k \rightarrow \infty} A_0 \dots A_{-k} \omega \eta_{-k-1}^- e_2 \\ &= \lim_{k \rightarrow \infty} A_0 \dots A_{-k} \omega e_{2q+1} \quad p.s. \end{aligned}$$

On en déduit que $\lim_{k \rightarrow \infty} A_0 \dots A_{-k} e_i = 0$, p.s. pour $i = 1, \dots, 2q+1$ par récurrence, de manière similaire au cas GARCH.

□

CHAPITRE

2

Identification d'un modèle GARCH

Cette étape consiste à déterminer dans la famille des modèles ARIMA ou GARCH l'ordre (ou le nombre de décalage optimal k) du modèle générateur des données par principe de parcimonie. L'identification des premières valeurs des ordres p et q est basée sur l'étude des corrélogrammes simple et partiel des fonctions d'autocorrélation AC et PAC. Par exemple,

- (i) Pour un AR , l'ordre correspond au nombre des coefficients d'autocorrélation partiel (PAC) non nuls (différents de 0), tandis que ;
- (ii) Pour un MA , l'ordre correspond au nombre des coefficients d'autocorrélation simple (AC) non nuls (différents de 0) ;
- (iii) Si les fonctions d'autocorrélation simple et partiel ne paraissent pas tronquées, il s'agit d'un processus $ARMA$.

Il est à noter que le modèle optimal parmi ceux-ci sélectionnés est choisi par tâtonnement.

Dans ce chapitre, nous intéressons à la sélection d'un modèle $GARCH$ ou $ARMA - GARCH$ approprié pour des observations (données) X_1, \dots, X_n d'un processus stationnaire centré. Une grande partie de la théorie de la finance repose sur l'hypothèse que les prix suivent une marche aléatoire. Le processus des variations des prix $X = (X_t)$ devrait donc constituer une séquence

de différence de martingale et devrait coïncider avec son processus d'innovation, $\epsilon = (\epsilon_t)$. La première question abordée dans ce chapitre à la section 2.1, sera de tester cette propriété, ou du moins une conséquence de celle-ci : l'absence de corrélation. Si des autocorrélations empiriques significativement non nulles sont détectées dans les variations des prix, ou si l'hypothèse de marche aléatoire ne tient pas, la première idée est d'ajuster un $ARMA(P, Q)$ aux données avant d'utiliser un modèle $GARCH(p, q)$ pour les résidus du modèle $ARMA$

$$X_t - \sum_{i=1}^P a_i X_{t-i} = \epsilon_t - \sum_{j=1}^Q b_j \epsilon_{t-j} \quad (2.1)$$

où (ϵ_t) est un bruit blanc GARCH vérifiant

$$\begin{cases} \epsilon_t = \sigma_t \eta_t \\ \sigma_t^2 = \omega + \sum_{i=1}^q \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2 \end{cases} \quad (2.2)$$

avec (η_t) une suite de variables iid centrées de variance unité, $\omega > 0$, $\alpha_i \geq 0$ ($i = 1, \dots, q$) et $\beta_j \geq 0$ ($j = 1, \dots, p$). En fait, parmi les outils de la méthode de Box-Jenkins est la détermination (identification) des ordres du modèle approprié (l'objet de ce chapitre) avant de passer à l'estimation des différents coefficients qui est l'objet du chapitre suivant. L'identification des ordres (P, Q) sera examinée dans la section 2.2, et celle des ordres (p, q) dans la section 2.3.

2.1 Vérification de l'absence de l'autocorrélation pour le bruit blanc

Cette section est consacrée à l'étude de l'ordre d'un modèle GARCH a priori choisi aux observations soumises à l'étude en utilisant les fonctions d'autocorrélations simples et partielles.

Considérons le modèle $GARCH(p, q)$ de (2, 2), on sait que quelques soient les ordres p et q , la solution non anticipative stationnaire du second ordre de (2.2) est un bruit blanc, c'est-à-dire un processus centré et les autocorrélations théoriques ($ACRTs$)

$$\rho(h) = E\epsilon_t \epsilon_{t+h} / E\epsilon_t^2 \quad (2.3)$$

satisfait $\rho(h) = 0$ pour tout $h \neq 0$.

Etant donné les observations $\epsilon_1, \dots, \epsilon_n$, les autocorrélations théoriques d'un processus centré (ϵ_t) sont généralement estimés par les autocorrélations empiriques (*ACREs*)

$$\hat{\rho}(h) = \frac{\hat{\gamma}(h)}{\hat{\gamma}(0)}, \quad \hat{\gamma}(h) = \hat{\gamma}(-h) = n^{-1} \sum_{t=1}^{n-h} \epsilon_t \epsilon_{t+h}, \quad (2.4)$$

pour $h = 0, \dots, n-1$.

Bien que défini pour $|h| < n$ la fonction d'autocovariance empirique évidemment un estimateur très pauvre de $\gamma(h)$ pour des valeurs de h proche de n . D'après Christian Francq ([24], p. 3), Box, Jenkins et Reinsel ([10], p. 32) recommandent à titre indicatif de n'utiliser ces quantités que si $n > 50$ et pour $h \leq \frac{n}{4}$.

Si (ϵ_t) est une suite de variables aléatoires centrées *iid* à variance finie alors la valeur de l'estimateur d'une *ACREs* est normale telle que :

$$\sqrt{n}\hat{\rho}(h) \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1), \quad (2.5)$$

pour tout $h \neq 0$. Pour un bruit blanc fort les *ACREs* se situent donc entre les bornes de confiance $\pm 1.96/\sqrt{n}$ avec une probabilité d'environ 95% lorsque n est grand. Dans un logiciel standard ces bornes au niveau de 5% sont généralement présentées par des lignes pointillées. Ces bandes de significativité ne sont pas valables pour un bruit blanc faible, en particulier pour un processus GARCH.

Mais, il est souvent utile et important (notamment pour la sélection des modèles) de déterminer si les autocovariances empiriques sont significativement différentes de zéro au-delà d'un certain rang. Pour cela il est primordial d'estimer la structure de ces autocorrélations empiriques. Cette structure sera étudiée dans la section suivante.

2.1.1 Comportement des autocorrélations empiriques d'un processus GARCH

Soient $\hat{\rho}_m = (\hat{\rho}_m(1), \dots, \hat{\rho}_m(m))'$ le vecteur des m premières *ACRE* de n observations du processus *GARCH*(p, q) défini par (2.2) et soit $\hat{\gamma}_m = (\hat{\gamma}(1), \dots, \hat{\gamma}(m))'$ le vecteur d'*ACVE* (Autocovariances Empiriques).

Théorème 2.1.1 (*Distributions asymptotiques des ACVE et ACRE d'un GARCH*)

Si (ϵ_t) est non anticipative et est une solution stationnaire du modèle $GARCH(p, q)$ de (2.2) et si $E\epsilon_t^4 < \infty$, alors quand $n \rightarrow \infty$,

$$\sqrt{n}\hat{\gamma}_m \xrightarrow{\mathcal{L}} \mathcal{N}(0, \Sigma_{\hat{\gamma}_m}) \quad \text{et} \quad \sqrt{n}\hat{\rho}_m \xrightarrow{\mathcal{L}} \mathcal{N}(0, \Sigma_{\hat{\rho}_m} := \{E\epsilon_t^2\}^{-2}\Sigma_{\hat{\gamma}_m}) \quad (2.6)$$

où

$$\Sigma_{\hat{\gamma}_m} = \begin{pmatrix} E\epsilon_t^2\epsilon_{t-1}^2 & E\epsilon_t^2\epsilon_{t-1}\epsilon_{t-2} & E\epsilon_t^2\epsilon_{t-1}\epsilon_{t-m} \\ E\epsilon_t^2\epsilon_{t-1}\epsilon_{t-2} & E\epsilon_t^2\epsilon_{t-2}^2 & \cdot \\ \cdot & \cdot & \dots \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ E\epsilon_t^2\epsilon_{t-1}\epsilon_{t-m} & E\epsilon_t^2\epsilon_{t-2}\epsilon_{t-m} & E\epsilon_t^2\epsilon_{t-m}^2 \end{pmatrix} \quad (2.7)$$

est non singulière (inversible). Si la loi de η_t est symétrique alors $\Sigma_{\hat{\gamma}_m}$ est diagonale.

Notons que, $\Sigma_{\hat{\rho}_m} = I_m$ lorsque (ϵ_t) est un bruit blanc fort d'après le Théorème 1.1 ([24]).

2.1.2 Test Portemanteau (Ljung-Box)

Le test de portemanteau standard pour vérifier que les données sont une réalisation d'un bruit blanc fort est celui de Ljung et Box (1978) ([31]), ayant les hypothèses présentées comme suit :

$$\begin{cases} H_0 : \text{absence d'autocorrélation} \\ H_1 : \text{existence d'autocorrélation} \end{cases}$$

Il s'agit de calculer la statistique Q_m^{LB} :

$$Q_m^{LB} := n(n+2) \sum_{i=1}^m \frac{\hat{\rho}^2(i)}{n-i} \quad (2.8)$$

où n est la taille de l'échantillon,

$\hat{\rho}(i)$ est l'autocorrélation au retard i et m est le nombre de retards testés. Les hypothèses peuvent se transmettre par :

$$\begin{cases} H_0 : \hat{\rho}(1) = \hat{\rho}(2) = \dots = \hat{\rho}(m) = 0 \\ H_1 : \text{il existe au moins } \hat{\rho}(i) \neq 0 \end{cases}$$

Sous H_0 , la statistique Q_m^{LB} est asymptotiquement distribuée selon une loi de χ^2 à m degré de liberté.

La règle de décision est :

- Si $Q_m^{LB} \geq \chi_m^2 (1 - \alpha)$ (est supérieure au $(1 - \alpha)$ -quantile d'une loi χ_m^2) on rejette l'hypothèse de bruit blanc fort H_0

- Si $Q_m^{LB} \leq \chi_m^2 (1 - \alpha)$ on accepte H_0

En utilisant les probabilités correspondantes ça peut mener à la même décision (Corrélogramme des résidus).

Les tests de Portemanteau sont construits pour chercher la non-corrélation, mais la distribution asymptotique des statistiques n'est plus χ_m^2 lorsque la série s'écarte de l'hypothèse de bruit blanc fort.

Par exemple, ces tests ne sont pas robustes à l'hétéroscédasticité conditionnelle. Dans le cadre *GARCH*, on peut souhaiter tester simultanément la nullité des m premières *ACRT* en utilisant des statistiques de portemanteau.

Théorème 2.1.2 (*Test du portemanteau corrigé en présence d'ARCH*)

Sous les hypothèses du théorème 2.1.1, la statistique de *Ljung – Box*

$$Q_m = n \hat{\rho}_m' \hat{\Sigma}_m^{-1} \hat{\rho}_m \quad (2.9)$$

a une distribution χ_m^2 asymptotique.

Preuve

Il suffit d'utiliser le Théorème 2.1.1 et le résultat suivant : si $X_n \longrightarrow \mathcal{N}(0, \Sigma)$ avec Σ non singulière et si $\hat{\Sigma}_n \longrightarrow \Sigma$ en probabilité, alors

$$X_n' \hat{\Sigma}_n^{-1} X_n \xrightarrow{\mathcal{L}} \chi_m^2.$$

Le test de *Ljung – Box* au seuil asymptotique α basé sur les m premières *ACRE* consiste à rejeter l'hypothèse que les données sont générées par un processus *GARCH*, si Q_m est supérieur au $(1 - \alpha)$ -quantile d'un χ_m^2 .

2.1.3 Autocorrélations partielles empiriques d'un *GARCH*

Soit r_m le vecteur des m premières autocorrélations partielles (*ACP*) et \hat{r}_m le vecteur des autocorrélations partielles empiriques (*ACPE*) du processus (ϵ_t) . Pour un bruit blanc faible, les *ACRE* et les *ACPE* ont la même distribution asymptotique d'après le théorème suivant

Théorème 2.1.3 (*Loi des ACPE et ACRE pour un bruit faible*)

Si (ε_t) est un bruit blanc faible et

$$\sqrt{n}\hat{\rho}_m \xrightarrow{\mathcal{L}} \mathcal{N}(0, \Sigma_{\hat{\rho}_m}) \quad (2.10)$$

alors

$$\sqrt{n}\hat{r}_m \xrightarrow{\mathcal{L}} \mathcal{N}(0, \Sigma_{\hat{\rho}_m}). \quad (2.11)$$

Preuve (Voir [24], p. 435)

Le résultat de ce théorème peut s'appliquer en particulier à un processus *GARCH*.

Par conséquent, sous l'hypothèse d'un bruit blanc *GARCH* avec un moment fini d'ordre 4, les estimateurs cohérents de $\sum \hat{r}_m$ sont

$$\hat{\Sigma}_{\hat{r}_m}^{(1)} = \hat{\Sigma}_{\hat{\rho}_m} \quad \text{ou} \quad \hat{\Sigma}_{\hat{r}_m}^{(2)} = \hat{J}_m \hat{\Sigma}_{\hat{\rho}_m} \hat{J}_m' \quad (2.12)$$

où \hat{J}_m est la matrice obtenue en remplaçant $\rho_X(1), \dots, \rho_X(m)$ par $\hat{\rho}_X(1), \dots, \hat{\rho}_X(m)$ dans la matrice jacobienne J_m de l'application $\rho_m \mapsto r_m$, et $\hat{\Sigma}_{\hat{\rho}_m}$ est l'estimateur consistant de $\Sigma_{\hat{\rho}_m}$ défini à la suite du Théorème 2.1.1.

On peut tester la nullité simultanée de plusieurs autocorrélations partielles théoriques en utilisant des tests portemanteau basés sur les statistiques

$$Q_m^{r,BP} = n\hat{r}_m' \hat{r}_m \quad \text{et} \quad Q_m^r = n\hat{r}_m' (\hat{\Sigma}_{\hat{\rho}_m}^{(i)})^{-1} \hat{r}_m. \quad (2.13)$$

2.2 Identification des ordres ARMA d'un ARMA-GARCH

Supposons que les outils développés dans la section 2.1 conduisent à rejeter l'hypothèse selon laquelle les données sont une réalisation d'un processus GARCH pur. Il est alors judicieux de chercher un modèle *ARMA*(P, Q) avec innovations GARCH. Le problème est alors de choisir (ou d'identifier) des ordres plausibles pour le modèle de la forme (2.1) suivant :

$$X_t - \sum_{i=1}^P a_i X_{t-i} = \varepsilon_t - \sum_{i=1}^Q b_i \varepsilon_{t-i}$$

sous des hypothèses standard (les polynômes AR et MA n'ayant pas de racines communes et ayant racines hors du disque unité avec $a_P b_Q \neq 0$, $E\epsilon_t^4 < \infty$), où (ϵ_t) est un bruit blanc $GARCH$ du formule (2.2).

2.2.1 Autocorrélations empiriques d'un ARMA-GARCH

Rappelons qu'un $MA(Q)$ satisfait $\rho_X(h) = 0$ pour tout $h > Q$, alors qu'un $AR(P)$ satisfait $r_X(h) = 0$ pour tout $h > P$.

Les $ACRE$ et les $ACEP$ jouent ainsi un rôle important dans l'identification des ordres P et Q .

Invalidité de la formule de Bartlett standard et de la formule modifiée

La validité de la formule habituelle de Bartlett repose sur des hypothèses incluant le bruit blanc fort, hypothèses qui sont évidemment incompatibles avec les erreurs $GARCH$. Nous verrons que cette formule conduit à sous-estimer les variances des $ACRE$ et des $ACPE$, et donc à des ordres $ARMA$ erronés.

Nous supposons tout au long que la loi de η_t est symétrique. Et le comportement asymptotique des $ACRE$ est déterminé par la formule de Bartlett généralisée suivante

$$\lim_{n \rightarrow +\infty} nCov(\hat{\rho}_X(i), \hat{\rho}_X(j)) = v_{ij} + v_{ij}^*, \quad i > 0, \quad j > 0$$

où

$$v_{ij} = \sum_{\ell=1}^{\infty} \omega_i(\ell)\omega_j(\ell), \quad v_{ij}^* = (\kappa_\epsilon - 1) \sum_{\ell=1}^{\infty} \rho_{\epsilon^2}(\ell)\omega_i(\ell)\omega_j(\ell) \quad (2.14)$$

Cette formule implique les autocorrélations théoriques de (X_t) et (ϵ_t^2) , ainsi que le rapport $\kappa_\epsilon - 1 = \gamma_{\epsilon^2}(0)/\gamma_\epsilon^2(0)$.

$$\omega_i(\ell) = \{2\rho_X(i)\rho_X(\ell) - \rho_X(\ell+i) - \rho_X(\ell-i)\}.$$

Le résultat suivant montre que la formule standard de Bartlett sous-estime toujours la variance asymptotique des autocorrélations de l'échantillon en présence d'erreurs $GARCH$.

Proposition 2.2.1 *Sous les hypothèses du théorème concernant la formule de Bartlett pour un processus linéaire faible, si le processus d'innovation linéaire (ϵ_t) est un processus $GARCH$ avec*

η_t a une loi symétrique, alors

$$v_{ii}^* \geq 0 \text{ pour tout } i > 0.$$

Si de plus

$$\alpha_1 > 0, \text{ Var}(\eta_t^2) > 0 \text{ et } \sum_{-h=\infty}^{\infty} \rho_X(h) \neq 0,$$

alors

$$v_{ii}^* > 0 \text{ pour tout } i > 0.$$

Preuve

D'après la proposition 2.2 (Voir [24], p. 62), on a $\rho_{\epsilon^2}(\ell) \geq 0$ pour tout ℓ , avec inégalité stricte quand $\alpha_1 > 0$.

Il résulte donc immédiatement de (2.14) que $v_{ii}^* > 0$. Lorsque $\alpha_1 > 0$ cette inégalité est stricte sauf si $\kappa_\epsilon = 1$ ou $\omega_i(l) = 0$ pour tout $l \geq 1$, c'est-à-dire :

$$2\rho_X(i)\rho_X(\ell) = \rho_X(\ell + i) + \rho_X(\ell - i).$$

Supposons que cette relation soit vérifiée et notons qu'elle est également satisfaite pour tout $\ell \in \mathbb{Z}$. De plus, en sommant sur ℓ , on obtient

$$2\rho_X(i) \sum_{-\infty}^{\infty} \rho_X(\ell) = \sum_{-\infty}^{\infty} \rho_X(\ell + i) + \rho_X(\ell - i) = 2 \sum_{-\infty}^{\infty} \rho_X(\ell).$$

Comme la somme de toutes les autocorrélations est supposée non nulle, nous avons donc $\rho_X(i) = 1$. En prenant $\ell = i$ dans la relation précédente, on trouve donc que $\rho_X(2i) = 1$.

L'itération de cet argument donne $\rho_X(ni) = 1$, et laisser n tendre vers l'infini donne une contradiction. Enfin, on ne peut pas avoir $\kappa_\epsilon = 1$, car :

$$\text{Var}(\epsilon_t^2) = (E\epsilon_t^2)^2(\kappa_\epsilon - 1) = Eh_t^2 \text{Var}(\eta_t^2) + \text{Var}(h_t) \geq \omega^2 \text{Var}(\eta_t^2) > 0.$$

Algorithme de calcul

En pratique on ne dispose pas des ACRT de (X_t) en (ϵ_t^2) et des autres quantités théoriques qui interviennent dans la formule de Bartlett généralisée.

L'algorithme utilisé pour estimer ces quantités est le suivant :

- i) Ajustement d'un $AR(p_0)$ à la suite X_1, \dots, X_n en utilisant un critère d'information pour sélectionner l'ordre p_0 ;
- ii) Calcul des ACRT $\rho_1(h)$, $h = 1, 2, \dots$ de ce modèle $AR(p_0)$;
- iii) Calcul des résidus e_{p_0+1}, \dots, e_n d'ajustement du modèle $AR(p_0)$;
- iv) Ajustement d'un $AR(p_1)$ à la série des carrés des résidus $e_{p_0+1}^2, \dots, e_n^2$ en utilisant également un critère d'information pour sélectionner p_1 ;
- v) Calcul des ACRT $\rho_2(h)$, $h = 1, 2, \dots$ de ce modèle $AR(p_1)$;
- vi) Estimation de $\lim_{n \rightarrow \infty} nCov\{\hat{\rho}(i), \hat{\rho}(j)\}$ par v_{ii} et v_{ii}^* où

$$v_{ii} = \sum_{\ell=-\ell_{\max}}^{\ell_{\max}} \rho_1(\ell) [2\rho_1(i)\rho_1(j)\rho_1(\ell) - 2\rho_1(i)\rho_1(\ell+j) - 2\rho_1(j)\rho_1(\ell+i) + \rho_1(\ell+j-i) + \rho_1(\ell-j-i)],$$

$$v_{ii}^* = \frac{\hat{\gamma}_{\epsilon^2}(0)}{\hat{\gamma}_{\epsilon}^2(0)} \sum_{\ell=-\ell_{\max}}^{\ell_{\max}} \rho_2(\ell) [2\rho_1(i)\rho_1(j)\rho_1^2(\ell) - 2\rho_1(j)\rho_1(\ell)\rho_1(\ell+i) - 2\rho_1(i)\rho_1(\ell)\rho_1(\ell+j) + \rho_1(\ell+i)\{\rho_1(\ell+j) + \rho_1(\ell-j)\}],$$

$$\hat{\gamma}_{\epsilon^2}(0) = \frac{1}{n-p_0} \sum_{\ell=-p_0+1}^n e_t^4 - \hat{\gamma}_{\epsilon}^2(0) \quad ; \quad \hat{\gamma}_{\epsilon}^2(0) = \frac{1}{n-p_0} \sum_{\ell=-p_0+1}^n e_t^2,$$

et ℓ_{\max} est un paramètre de troncature déterminé numériquement de sorte que $|\rho_1(\ell)|$ et $|\rho_2(\ell)|$ soient inférieurs à une certaine tolérance (par ex. 10^{-5}) pour tout $\ell > \ell_{\max}$.

2.2.2 Autocorrélations empiriques de ARMA-GARCH (bruit non symétrique)

La formule de Bartlett généralisée ci-dessus est vérifiée sous la condition

$$E\epsilon_t^4 = \kappa_{\epsilon} (E\epsilon_t^2)^2 < \infty \quad \text{et}$$

$$E\epsilon_{t_1}\epsilon_{t_2}\epsilon_{t_3}\epsilon_{t_4} = 0 \quad \text{si} \quad t_1 \neq t_2 \quad t_1 \neq t_3 \quad t_1 \neq t_4$$

qui peut ne pas être satisfaite si la distribution du bruit η_t de l'équation *GARCH* n'est pas symétrique. Nous allons voir quel est le comportement asymptotique des *ACVE* et *ACRE* dans le cas très général d'un processus linéaire dont l'innovation (ϵ_t) est un bruit blanc faible. La

propriété suivante permet d'interpréter la variance asymptotique des ACVE comme la densité spectrale en 0 d'un processus multivarié. Soit $\hat{\gamma}_{0:m} = (\hat{\gamma}(0), \dots, \hat{\gamma}(m))'$.

Théorème 2.2.1 (*Variance asymptotique des ACVE d'un processus linéaire faible*)

Soit $(X_t)_{t \in \mathbb{Z}}$ un processus stationnaire réel vérifiant :

$$X_t = \sum_{j=-\infty}^{\infty} \Psi_j \epsilon_{t-j}, \quad \sum_{j=-\infty}^{\infty} |\Psi_j| < \infty,$$

où $(\epsilon_t)_{t \in \mathbb{Z}}$ est un bruit blanc faible tel que $E\epsilon_t^4 < \infty$. On pose :

$$Y_t = X_t(X_t, X_{t+h}, \dots, X_{t+m})', \quad \Gamma_Y(h) = EY_t^* Y_{t+h}^{*'} \quad \text{et} \quad f_{Y^*}(\lambda) = \frac{1}{2\pi} \sum_{h=-\infty}^{+\infty} e^{-ih\lambda} \Gamma_Y(h),$$

la densité spectrale du processus $Y_t^* = (Y_t^*)$, $Y_t^* = Y_t - EY_t$. Alors,

$$\lim_{n \rightarrow \infty} nVar_{\hat{Y}_{0:m}} := \Sigma_{\hat{Y}_{0:m}} = 2\pi f_{Y^*}(0). \quad (2.15)$$

Preuve

Par stationnarité et application du théorème de convergence dominée de Lebesgue,

$$\begin{aligned} nVar_{\hat{Y}_{0:m}} + o(1) &= nCov\left(\frac{1}{n} \sum_{t=1}^n Y_t^*, \frac{1}{n} \sum_{s=1}^n Y_s^*\right) \\ &= \sum_{h=-n+1}^{n-1} \left(1 - \frac{|h|}{n}\right) Cov(Y_t^*, Y_{t+h}^*) \\ &\longrightarrow \sum_{h=-\infty}^{+\infty} \Gamma_Y(h) = 2\pi f_{Y^*}(0) \end{aligned}$$

lorsque $n \rightarrow \infty$.

La matrice $\Sigma_{\hat{Y}_{0:m}}$ impliqué dans (2.15) est appelé la variance de long terme dans la littérature économétrique, pour rappeler qu'il s'agit de la variance limite d'une moyenne empirique. Plusieurs méthodes peuvent être envisagées pour l'estimation de la variance à long terme.

(i) L'estimateur naïf basé sur le remplacement de $\Gamma_Y(h)$ par $\hat{\Gamma}_Y(h)$ dans $f_{Y^*}(0)$ ne converge pas. Cependant, un estimateur convergent peut être obtenu en pondérant les $\hat{\Gamma}_Y(h)$, en utilisant un poids proche de 1 quand h est très petit par rapport à n , et un poids proche de 0 quand h est grand.

(ii) Un estimateur convergent de $f_{Y^*}(0)$ peut également être obtenu à l'aide du périodogramme lissé (Brockwell et Davis 1991, [11]).

(iii) Pour un $AR(r)$ multivarié,

$$\mathcal{A}_r(B)Y_t = Y_t - \sum_{i=1}^r A_i Y_{t-i} = Z_t,$$

(Z_t) bruit blanc avec variance Σ_Z , la densité spectrale en 0 est

$$f_Y(0) = \frac{1}{2\pi} \mathcal{A}_r(1)^{-1} \Sigma_Z \mathcal{A}_r'(1)^{-1}.$$

Un modèle AR multivarié est facilement ajusté, même un AR d'ordre élevé, en utilisant une version multivariée de l'algorithme de Durbin-Levinson ([11], p. 422). Pour cela une méthode peut ainsi être proposée :

1. Ajuster les modèles $AR(r)$, avec $r = 0, 1, \dots, R$, aux données $Y_1 - \bar{Y}_n, \dots, Y_{n-m} - \bar{Y}_n$ où $\bar{Y} = (n - m)^{-1} \sum_{t=1}^{n-m} Y_t$.
2. Sélectionner une valeur r_0 en minimisant un critère d'information, par exemple le BIC .
3. Posons

$$\hat{\Sigma}_{\hat{Y}_{0:m}} = \hat{\mathcal{A}}_{r_0}(1)^{-1} \hat{\Sigma}_{r_0} \hat{\mathcal{A}}_{r_0}(1)^{-1}$$

avec une notation évidente.

2.2.3 Identification des ordres (P,Q)

La détermination des ordres d'un modèle $ARMA(P, Q)$ par les autocorrélations empiriques (ACRE) et des autocorrélations partielles (ACPE) n'est pas une tâche facile. Plusieurs méthodes sont utilisées telles que la méthode du coin (1980), méthode de l'épsilon-algorithme ou encore la méthode de Glasbey(1982) qui définissent des statistiques plus commodes. Dans la section suivante la méthode du coin est présentée.

Méthode du coin :

Notons $D(i, j)$ la matrice de Toeplitz $j \times j$

$$D(i, j) = \begin{pmatrix} \rho_X(i) & \rho_X(i-1) & \dots & \rho_X(i-j+1) \\ \rho_X(i+1) & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \rho_X(i+j-1) & \rho_X(i+1) & \dots & \rho_X(i) \end{pmatrix} \quad (2.16)$$

et notons $\Delta(i, j)$ son déterminant. Puisque $\rho_X(h) = \sum_{i=1}^P a_i \rho_X(h-i) = 0$, pour tout $h > Q$, il est clair que $D(i, j)$ n'est pas une matrice de plein rang si $i > Q$ et $j > P$. Plus précisément, P et Q sont les ordres minimaux (c'est-à-dire que (X_t) n'admet pas une représentation $ARMA(P', Q')$ avec $P' < P$ ou $Q' < Q$) si et seulement si :

$$\begin{cases} \Delta(i, j) = 0 & \forall i > Q \text{ et } \forall j > P, \\ \Delta(i, P) \neq 0 & \forall i \geq Q, \\ \Delta(Q, j) \neq 0 & \forall j \geq P. \end{cases} \quad (2.17)$$

Les ordres minimaux P et Q sont donc caractérisés par le tableau suivant :

| | | | | | | |
|-----------------|----------|----------|--|----------|--------------|---------|
| $i \setminus j$ | 1 | 2 | | Q | $Q + 1$ | ... |
| 1 | ρ_1 | ρ_2 | | ρ_Q | ρ_{Q+1} | ... |
| \vdots | | | | | | |
| P | | | | × | × | × |
| $P + 1$ | | | | × | 0 | 0 0 0 0 |
| \vdots | | | | × | 0 | 0 0 0 0 |

(2.18)

où $\Delta(j, i)$ est à l'intersection de la ligne i et de la colonne j , et \times désigne un élément non nul. Les ordres P et Q sont ainsi caractérisés par un coin de zéros dans le tableau (T1), d'où le terme "méthode des coins". Les entrées de ce tableau sont facilement obtenues en utilisant la récursivité sur j donnée par

$$\Delta(i, j)^2 = \Delta(i + 1, j)\Delta(i - 1, j) + \Delta(i, j + 1)\Delta(i, j - 1) \quad (2.19)$$

et soit $\Delta(i, 0) = 1$, $\Delta(i, 1) = \rho_X(|i|)$.

Notons $\hat{D}(i, j)$, $\hat{\Delta}(i, j)$, $(\hat{T}1)$, ... les termes obtenus en remplaçant $\{\rho_X(h)\}$ par $\{\hat{\rho}_X(h)\}$ dans $D(i, j)$, $\Delta(i, j)$, (T1), ...

En pratique, on a qu'un nombre fini de *ACRE* $\hat{\rho}_X(1), \dots, \hat{\rho}_X(K)$, ce qui permet de calculer $\hat{\Delta}(j, i)$ pour $i \geq 1, j \geq 1$ et $i + j \leq K + 1$. Le tableau $(\hat{T}1)$ est donc triangulaire. Comme les $\hat{\Delta}(j, i)$ estiment de manière consistente les $\Delta(j, i)$, les ordres P et Q sont caractérisés par un coin de petites valeurs dans le tableau $(\hat{T}1)$. Cependant, la notion de « petite valeur » dans $(\hat{T}1)$ n'est pas assez précis.

Il est préférable de considérer la statistique studentisée définie pour $i = -K, \dots, K$ et $j = 0, \dots, K - |i| + 1$, par

$$t(i, j) = \sqrt{n} \frac{\hat{\Delta}(i, j)}{\hat{\sigma}_{\hat{\Delta}(i, j)}}, \quad \hat{\sigma}_{\hat{\Delta}(i, j)}^2 = \frac{\partial \hat{\Delta}(i, j)}{\partial \rho'_K} \hat{\Sigma}_{\rho'_K} \frac{\partial \hat{\Delta}(i, j)}{\partial \rho_K} \quad (2.20)$$

où $\hat{\Sigma}_{\rho'_K}$ est un estimateur convergent de la matrice de covariance asymptotique des K premières *ACRE*, que l'on peut obtenir par l'algorithme de la partie 2.2.1 ou par celui de la partie 2.2.2, et où le Jacobien :

$$\frac{\partial \hat{\Delta}(i, j)}{\partial \rho'_K} = \left(\frac{\partial \hat{\Delta}(i, j)}{\partial \rho_X(1)}, \dots, \frac{\partial \hat{\Delta}(i, j)}{\partial \rho_X(K)} \right)$$

est obtenu à partir de la dérivation de (2.19) :

$$\begin{aligned} \frac{\partial \hat{\Delta}(i, 0)}{\partial \rho_X(k)} &= 0, \quad \text{pour } i = -K - 1, \dots, K - 1 \quad \text{et } k = 1, \dots, K; \\ \frac{\partial \hat{\Delta}(i, 1)}{\partial \rho_X(k)} &= \mathbf{1}_{\{k\}}(|i|), \quad \text{pour } i = -K, \dots, K \quad \text{et } k = 1, \dots, K; \\ \frac{\partial \hat{\Delta}(i, j + 1)}{\partial \rho_X(k)} &= \frac{2\hat{\Delta}(i, j) \frac{\partial \hat{\Delta}(i, j)}{\partial \rho_X(k)} - \hat{\Delta}(i + 1, j) \frac{\partial \hat{\Delta}(i - 1, j)}{\partial \rho_X(k)} - \hat{\Delta}(i - 1, j) \frac{\partial \hat{\Delta}(i + 1, j)}{\partial \rho_X(k)}}{\hat{\Delta}(i, j - 1)} \\ &\quad - \frac{\left\{ \hat{\Delta}(i, j)^2 - \hat{\Delta}(i + 1, j) \hat{\Delta}(i - 1, j) \right\} \frac{\partial \hat{\Delta}(i, j - 1)}{\partial \rho_X(k)}}{\hat{\Delta}^2(i, j - 1)}, \end{aligned} \quad (2.21)$$

pour $k = 1, \dots, K, i = -K + j, \dots, K - j$ et $j = 1, \dots, K$.

Lorsque $\Delta(i, j) = 0$ la statistique $t(i, j)$ suit asymptotiquement une loi $\mathcal{N}(0, 1)$ (particulièrement, sous réserve d'existence de EX_t^4). Si, au contraire, $\Delta(i, j) \neq 0$ alors $(1/\sqrt{n})|t(i, j)| \rightarrow \infty$ *p.s.* quand $n \rightarrow \infty$. On peut rejeter l'hypothèse $\Delta(i, j) = 0$ au niveau α si $|t(i, j)|$ est supérieure de $(1 - \alpha/2)$ -quantile d'un $\mathcal{N}(0, 1)$.

On peut aussi détecter automatiquement un coin de petites valeurs dans le tableau (\tilde{T}_1) des $t(i, j)$, si aucune valeur de ce coin n'est supérieure à ce $(1 - \alpha/2)$ -quantile en valeur absolue. Cette pratique ne correspond à aucun test formel au niveau α , mais admet un petit nombre de valeurs plausibles à sélectionner pour les ordres P et Q .

2.3 Identification des ordres GARCH d'un ARMA-GARCH

Pour le problème d'identification, on peut supposer que les observations sont tirées d'un modèle GARCH(p, q) pur. On peut également admettre que (ϵ_t^2) sont un ARMA(\tilde{P}, \tilde{Q}) où $\tilde{P} = \max(p, q)$ et $\tilde{Q} = p$. Le choix d'un petit nombre valeurs plausibles des ordres p et q peut être obtenu en plusieurs étapes à l'aide de divers outils :

- (i) Inspection des autocorrélations empiriques et des autocorrélations partielles empiriques de $\epsilon_1^2, \dots, \epsilon_n^2$.
- (ii) Inspection des statistiques qui sont des fonctions des autocovariances empiriques de ϵ_t^2 (méthode des coins, algorithme epsilon, ...).
- (iii) Utilisation de critères d'information (AIC, BIC, \dots).
- (iv) tests de significativité de certains coefficients.
- (v) Analyse des résidus.

Les étapes (iii) et (v) et dans une large mesure l'étape (iv) nécessitent l'estimation de modèles et sont utilisé pour les valider ou les modifier. L'étape (i) repose sur la représentation ARMA pour le carré d'un GARCH traité. En particulier si (ϵ_t) est un processus ARCH(q) alors l'autocorrélation partielle théorique fonction $r_{\epsilon^2}(\cdot)$ de (ϵ_t^2) satisfait $r_{\epsilon^2}(h) = 0, \forall h > q$.

Pour les modèles mixtes la méthode des coins peut être utilisée.

2.3.1 Méthode de coin dans le cas GARCH

Pour identifier les ordres d'un processus GARCH(p, q) on peut utiliser le fait que (ϵ_t^2) suit un ARMA(\tilde{P}, \tilde{Q}) avec $\tilde{P} = \max(p, q)$ et $\tilde{Q} = p$. Dans le cas d'un GARCH pur, $(\epsilon_t) = (X_t)$ est observé. La variance asymptotique des ACRE de $\epsilon_t^2, \dots, \epsilon_n^2$ peut être estimée par la méthode décrit dans la partie 2.2.2. Le tableau des statistiques studentisées pour la méthode du coin

s'en déduit comme décrit dans la partie précédente. Le problème est alors de détecter au moins un coin des valeurs normales à partir de la ligne $\tilde{P} + 1$ et de la colonne $\tilde{Q} + 1$ du tableau sous les contraintes $\tilde{P} \geq 1$ (car $\max(p, q) \geq q \geq 1$) et $\tilde{P} \geq \tilde{Q}$. Cela conduit à la sélection des modèles $GARCH(p, q)$ tels que :

$$(p, q) = (\tilde{P}, \tilde{Q}) \quad \text{quand} \quad \tilde{P} < \tilde{Q} \quad (2.22)$$

et

$$(p, q) = (\tilde{Q}, 1), \quad (p, q) = (\tilde{Q}, 2), \quad \dots, \quad (p, q) = (\tilde{Q}, \tilde{P}) \quad \text{quand} \quad \tilde{Q} \geq \tilde{P}. \quad (2.23)$$

Dans le cas $ARMA - GARCH$ les ϵ_t ne sont pas observés mais peuvent être approchés par les résidus de $ARMA$. Alternativement pour éviter l'estimation $ARMA$, on peut les approcher par les résidus d'ajustement d'un AR , comme décrit dans les étapes 1 et 3 de l'algorithme de la section 2.2.1.

2.4 Test du multiplicateur de Lagrange pour l'homoscédasticité conditionnelle

Pour tester des restrictions linéaires sur les paramètres d'un modèle, les tests les plus utilisés sont les tests de Wald, test de multiplicateur de Lagrange (LM) et le test du rapport de vraisemblance (LR). Le test de multiplicateur de Lagrange (LM) également appelé test du score, ou test de Rao, est attrayant car il ne nécessite qu'une estimation du modèle contraint, qui est souvent beaucoup plus facile que d'estimer le modèle non contraint. La forme générale du test LM est donnée dans la section suivante. Ensuite, nous présentons un test LM pour l'homoscédasticité conditionnelle dans la section 2.4.2.

2.4.1 Forme générale du test LM

Considérons un modèle paramétrique, avec une vraie valeur de paramètre $\theta_0 \in \mathbb{R}^d$, et une hypothèse nulle

$$H_0 : R\theta_0 = r$$

où R est une matrice $s \times d$ donnée de plein rang s , et r est un vecteur $s \times 1$ donné. Cette formule permet de tester, par exemple, si les s premières composantes de θ_0 sont nulles (il suffit

de poser $R = [I_s : 0_{s \times (d-s)}]$ et $r = 0_s$).

Soit $l_n(\theta)$ la log-vraisemblance des observations X_1, \dots, X_n .

Nous supposons l'existence des estimateurs du maximum de vraisemblance non contraints et contraints (par H_0), satisfaisant respectivement

$$\hat{\theta} = \arg \sup_{\theta} l_n(\theta) \quad \text{et} \quad \hat{\theta}^c = \arg \sup_{\theta: R\theta=r} l_n(\theta).$$

Sous certaines hypothèses de régularité, le vecteur score satisfait un théorème central limite et nous avons

$$\frac{1}{\sqrt{n}} \frac{\partial}{\partial \theta} l_n(\theta_0) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \mathcal{J}) \quad \text{et} \quad \sqrt{n}(\hat{\theta} - \theta_0) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \mathcal{J}^{-1}) \quad (2.24)$$

où \mathcal{J} est la matrice d'information de Fisher. Pour déterminer l'estimateur contraint, nous introduisons le Lagrangien

$$\mathcal{L}(\theta, \lambda) = l_n(\theta) - \lambda'(R\theta - r).$$

Nous avons

$$(\hat{\theta}^c, \hat{\lambda}) = \arg \sup_{(\theta, \lambda)} \mathcal{L}(\theta, \lambda).$$

Les conditions du premier ordre donnent : $R' \hat{\lambda} = \frac{\partial}{\partial \theta} l_n(\hat{\theta}^c)$ et $R \hat{\theta}^c = r$.

La deuxième convergence dans (2.24) montre donc que sous H_0 ,

$$\sqrt{n}R(\hat{\theta} - \hat{\theta}^c) = \sqrt{n}R(\hat{\theta} - \theta_0) \xrightarrow{\mathcal{L}} \mathcal{N}(0, R\mathcal{J}^{-1}R'). \quad (2.25)$$

En utilisant la convention $a \stackrel{c}{=} b$ pour $a = b + c$, des développements limités impliquent, sous des hypothèses de régularité habituelle :

$$0 = \frac{1}{\sqrt{n}} \frac{\partial}{\partial \theta} l_n(\hat{\theta}) \stackrel{op(1)}{=} \frac{1}{\sqrt{n}} \frac{\partial}{\partial \theta} l_n(\theta_0) - \mathcal{J} \sqrt{n}(\hat{\theta} - \theta_0),$$

$$\frac{1}{\sqrt{n}} \frac{\partial}{\partial \theta} l_n(\hat{\theta}^c) \stackrel{op(1)}{=} \frac{1}{\sqrt{n}} \frac{\partial}{\partial \theta} l_n(\theta_0) - \mathcal{J} \sqrt{n}(\hat{\theta}^c - \theta_0)$$

ce qui donne par soustraction :

$$\sqrt{n}(\hat{\theta} - \hat{\theta}^c) \stackrel{op(1)}{=} \mathcal{J}^{-1} \frac{1}{\sqrt{n}} \frac{\partial}{\partial \theta}(\hat{\theta}^c) = \mathcal{J}^{-1} \frac{1}{\sqrt{n}} R' \hat{\lambda}. \quad (2.26)$$

Enfin, (2.25) et (2.26) impliquent

$$\frac{1}{\sqrt{n}} \hat{\lambda} = (R\mathcal{J}^{-1}R')^{-1} \sqrt{n}R(\hat{\theta} - \theta_0) \xrightarrow{\mathcal{L}} \mathcal{N}\{0, (R\mathcal{J}^{-1}R')^{-1}\}$$

et puis

$$\frac{1}{\sqrt{n}} \frac{\partial}{\partial \theta} \ell_n(\hat{\theta}^c) = R' \frac{1}{\sqrt{n}} \hat{\lambda} \xrightarrow{\mathcal{L}} \mathcal{N}\{0, R'(R\mathcal{J}^{-1}R')^{-1}R\}.$$

Ainsi, sous H_0 , la statistique du test

$$LM_n = \frac{1}{n} \hat{\lambda}' R \hat{\mathcal{J}}^{-1} R' \hat{\lambda} = \frac{1}{n} \frac{\partial}{\partial \theta'} \ell_n(\hat{\theta}^c) \hat{\mathcal{J}}^{-1} \frac{\partial}{\partial \theta} \ell_n(\hat{\theta}^c) \quad (2.27)$$

suit asymptotiquement un χ_s^2 , à condition que $\hat{\mathcal{J}}$ soit un estimateur convergeant en probabilité vers \mathcal{J} . En général on peut prendre

$$\hat{\mathcal{J}} = -\frac{1}{n} \frac{\partial^2 \ell_n(\hat{\theta}^c)}{\partial \theta \partial \theta'} \quad \text{ou} \quad \hat{\mathcal{J}} = -\frac{1}{n} \frac{\partial \ell_n(\hat{\theta}^c)}{\partial \theta} \frac{\partial \ell_n(\hat{\theta}^c)}{\partial \theta'}$$

La région critique du test LM au niveau asymptotique α est $\{LM_n > \chi_s^2(1 - \alpha)\}$.

Cas où la statistique LM_n prend la forme nR^2

Le teste LM peut être mis en œuvre de manière extrêmement simple. Considérons un modèle non linéaire conditionnellement homoscédasticité de la forme $Y_t = F_{\theta_0}(W_t) + \epsilon_t$, où $W_t = (X_t, Y_{t-1}, \dots)$ dépend d'un vecteur de variables exogènes X_t et de valeurs passées de la variable endogène, et où ϵ_t est $iid(0, \sigma_0^2)$, de sorte que W_t et ϵ_t sont indépendants.

L'hypothèse que l'on cherche à tester est définie par

$$H_0 : \psi_0 = 0 \quad \text{où} \quad \theta_0 = \begin{pmatrix} \beta_0 \\ \psi_0 \end{pmatrix}, \quad \beta_0 \in \mathbb{R}^{d-s}, \quad \psi_0 \in \mathbb{R}^s.$$

Pour se ramener au cadre de la partie précédente, on peut poser $R = [0_{s \times d-s} : I_s]$ et remarquer que

$$\frac{\partial \ell_n(\hat{\theta}^c)}{\partial \psi} = R \frac{\partial \ell_n(\hat{\theta}^c)}{\partial \theta} = RR' \hat{\lambda} = \hat{\lambda} \quad \text{et} \quad \frac{1}{\sqrt{n}} \hat{\lambda} \xrightarrow{\mathcal{L}} \mathcal{N}(0, \Sigma_\lambda)$$

où $\sum_{\lambda} = (\mathcal{J}^{22})^{-1}$ et $\mathcal{J}^{22} = R\mathcal{J}^{-1}R'$ est le bloc en bas à droite de \mathcal{J}^{-1} . On suppose que $\sigma_0^2 = \sigma^2(\beta_0)$ ne dépend pas de ψ_0 . Avec une vraisemblance gaussienne, on a

$$\frac{1}{\sqrt{n}}\hat{\lambda} = \frac{1}{\sqrt{n}\hat{\sigma}^{c2}} \sum_{t=1}^n \epsilon_t(\hat{\theta}^c) \frac{\partial}{\partial \psi} F_{\hat{\theta}^c}(W_t) = \frac{1}{\sqrt{n}\hat{\sigma}^{c2}} F'_{\psi} \hat{U}^c,$$

où $\epsilon_t(\theta) = Y_t - F_{\theta}(W_t)$, $\hat{\sigma}^{c2} = \sigma^2(\hat{\beta}^c)$, $\hat{\epsilon}_t^c = \epsilon_t(\hat{\theta}^c)$, $\hat{U}^c = (\hat{\epsilon}_1^c, \dots, \hat{\epsilon}_n^c)$ et

$$F'_{\psi} = \left(\frac{\partial F_{\hat{\theta}^c}(W_1)}{\partial \psi}, \dots, \frac{\partial F_{\hat{\theta}^c}(W_n)}{\partial \psi} \right).$$

On écrit \mathcal{J} par bloc sous la forme

$$\mathcal{J} = \begin{pmatrix} \mathcal{J}_{11} & \mathcal{J}_{12} \\ \mathcal{J}_{21} & \mathcal{J}_{22} \end{pmatrix}$$

où \mathcal{J}_{11} et \mathcal{J}_{22} sont carrées, de tailles respectivement $d - s$ et s . Sous l'hypothèse que la matrice d'information \mathcal{J} est bloc-diagonale, c'est-à-dire $\mathcal{J}_{12} = 0$. On a $\mathcal{J}^{22} = \mathcal{J}_{22}^{-1}$ où $\mathcal{J}_{22} = R\mathcal{J}R'$, ce qui donne $\sum_{\lambda} = \mathcal{J}_{22}$. On peut alors prendre

$$\sum_{\lambda} = \frac{1}{\hat{\sigma}^{c4}} \frac{1}{n} \hat{U}^{c'} \hat{U}^c \frac{1}{n} F'_{\psi} F_{\psi} \stackrel{0p(I)}{=} \frac{1}{\hat{U}^{c'} \hat{U}^c} F'_{\psi} F_{\psi}$$

comme estimateur convergent de \sum_{λ} . On obtient alors

$$LM_n = n \frac{\hat{U}^{c'} F_{\psi} (F'_{\psi} F_{\psi})^{-1} F'_{\psi} \hat{U}^c}{\hat{U}^{c'} \hat{U}^c}, \tag{2.28}$$

qui n'est autre que n multiplié par le coefficient de détermination non centré dans la regression de $\hat{\epsilon}_t^c$ sur les variables $\partial F_{\hat{\theta}^c}(W_t)/\partial \psi_i$ pour $i = 1, \dots, s$.

Test quasi-LM

Lorsque $\ell_n(\theta)$ n'est plus supposée être la log-vraisemblance, mais simplement la quasi log-vraisemblance les équations peuvent en général être remplacées par

$$\frac{1}{\sqrt{n}} \frac{\partial}{\partial \theta} \ell_n(\theta_0) \xrightarrow{\mathcal{L}} \mathcal{N}(0, I) \quad \text{et} \quad \sqrt{n} (\hat{\theta} - \theta_0) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \mathcal{J}^{-1} I \mathcal{J}^{-1}) \tag{2.29}$$

où

$$I = \lim_{n \rightarrow \infty} \frac{1}{n} Var \frac{\partial}{\partial \theta} \ell_n(\theta_0), \quad \mathcal{J} = \lim_{n \rightarrow \infty} \frac{1}{n} \frac{\partial^2}{\partial \theta \partial \theta'} \ell_n(\theta_0) \quad p.s.$$

Il convient alors de remplacer (2.27) par l'expression suivante, plus complexe mais plus robuste

$$\begin{aligned} LM_n &= \frac{1}{n} \hat{\lambda}' R \hat{\mathcal{J}}^{-1} R' \hat{\lambda} \left(R \hat{\mathcal{J}}^{-1} \hat{I} \hat{\mathcal{J}}^{-1} R' \right)^{-1} R \hat{\mathcal{J}}^{-1} R' \hat{\lambda} \\ &= \frac{1}{n} \frac{\partial}{\partial \theta'} \ell_n(\hat{\theta}^c) \hat{\mathcal{J}}^{-1} R' \left(R \hat{\mathcal{J}}^{-1} \hat{I} \hat{\mathcal{J}}^{-1} R' \right)^{-1} R \hat{\mathcal{J}}^{-1} \frac{\partial}{\partial \theta} \ell_n(\hat{\theta}^c) \end{aligned} \quad (2.30)$$

où \hat{I} et $\hat{\mathcal{J}}^{-1}$ sont des estimateurs convergents de I et \mathcal{J} . Un estimateur convergent de \mathcal{J} est obtenu de manière évidente comme moyenne empirique.

2.4.2 Test LM d'homoscédasticité conditionnelle

On considère le test d'homoscédasticité conditionnelle

$$H_0 : \alpha_{01} = \alpha_{02} = \dots = \alpha_{0q} = 0$$

dans le modèle ARCH(q) suivant

$$\begin{cases} \epsilon_t = \sigma_t \eta_t & \eta_t \text{ iid}(0, 1) \\ \sigma_t^2 = \omega_0 + \sum_{i=1}^q \alpha_{0i} \epsilon_{t-i}^2 & \omega_0 > 0 \quad \alpha_{0i} \geq 0 \end{cases}$$

En un point $\theta = (\omega, \alpha_1, \alpha_2, \dots, \alpha_q)$, à une constante additive près, la quasi log-vraisemblance s'écrit

$$\ell_n(\theta) = -\frac{1}{2} \sum_{t=1}^n \left\{ \frac{\epsilon_t^2}{\sigma_t^2(\theta)} + \log \sigma_t^2(\theta) \right\} \quad \sigma_t^2(\theta) = \omega + \sum_{i=1}^q \alpha_i \epsilon_{t-i}^2$$

avec la convention $\epsilon_{t-i}^2 = 0$ pour $t \leq 0$. L'estimateur de quasi maximum de vraisemblance contraint est

$$\hat{\theta}^c = (\hat{\omega}^c, 0, 0, \dots, 0) \quad \text{où} \quad \hat{\omega}^c = \sigma_t^2(\theta^c) = \frac{1}{n} \sum_{i=1}^q \epsilon_t^2.$$

En $\theta_0 = (\omega_0, 0, 0, \dots, 0)$, le vecteur des scores satisfait :

$$\begin{aligned} \frac{1}{\sqrt{n}} \frac{\partial}{\partial \theta} \ell_n(\theta_0) &= \frac{1}{2\sqrt{n}} \sum_{t=1}^n \frac{1}{\sigma_t^2(\theta_0)} \{ \epsilon_t^2 - \sigma_t^2(\theta_0) \} \frac{\partial \sigma_t^2(\theta_0)}{\partial \theta} \\ &= \frac{1}{2\sqrt{n}} \sum_{t=1}^n \frac{1}{\omega_0} (\eta_t^2 - 1) \begin{pmatrix} 1 \\ \epsilon_{t-1}^2 \\ \vdots \\ \epsilon_{t-q}^2 \end{pmatrix} \\ \xrightarrow{\mathcal{L}} \mathcal{N}(0, I), \quad I &= \frac{\kappa_\eta - 1}{4\omega_0^2} \begin{pmatrix} 1 & \boldsymbol{\omega}_0 \\ \boldsymbol{\omega}'_0 & I_{22} \end{pmatrix} \end{aligned}$$

sous H_0 , avec $\boldsymbol{\omega}_0 = (\omega_0, \omega_0, \dots, \omega_0)' \in \mathbb{R}^q$ et I_{22} une matrice dont les éléments diagonaux sont $\omega_0^2 \kappa_\eta$ et $\kappa_\eta = E\eta_t^4$ et les autres éléments sont ω_0^2 .

Le bloc en bas à droite de I^{-1} est alors

$$I^{22} \frac{4\omega_0^2}{\kappa_\eta - 1} (I_{22} - \boldsymbol{\omega}_0 \boldsymbol{\omega}'_0)^{-1} = \frac{4\omega_0^2}{\kappa_\eta - 1} \{ (\kappa_\pi - 1) \omega_0^2 I_q \}^{-1} = \frac{4}{(\kappa_\eta - 1)^2} I_q \quad (2.31)$$

On a de plus

$$\begin{aligned} \frac{1}{n} \frac{\partial^2 \ell_n(\theta_0)}{\partial \theta \partial \theta'} &= \frac{1}{2n} \sum_{t=1}^n \frac{1}{\sigma_t^6(\theta_0)} \{ 2\epsilon_t^2 - \sigma_t^2(\theta_0) \} \frac{\partial \sigma_t^2(\theta_0)}{\partial \theta} \frac{\partial \sigma_t^2(\theta_0)}{\partial \theta'} \\ &= \frac{1}{2n} \sum_{t=1}^n \frac{1}{\omega_0^2} (2\eta_t^2 - 1) \begin{pmatrix} 1 \\ \epsilon_{t-1}^2 \\ \vdots \\ \epsilon_{t-q}^2 \end{pmatrix} (1, \epsilon_{t-1}^2, \dots, \epsilon_{t-q}^2) \\ \longrightarrow J &= \frac{2}{\kappa_\eta - 1} I, \quad p.s. \end{aligned}$$

D'après (2, 30) si on utilise des estimateurs de I et J tels que $\hat{J} = \frac{2}{\kappa_\eta - 1} \hat{I}$, on obtient

$$LM_n = \frac{2}{\hat{\kappa}_\eta - 1} \frac{1}{n} \hat{\lambda}' R \hat{J}^{-1} R' \hat{\lambda} = \frac{1}{n} \frac{\partial}{\partial \theta'} \ell_n(\hat{\theta}^c) \hat{J}^{-1} \frac{\partial}{\partial \theta} \ell_n(\hat{\theta}^c).$$

En utilisant (2, 31) et de remarquer que

$$\frac{\partial}{\partial \theta} \ell_n(\hat{\theta}^c) = \begin{pmatrix} 0 \\ \frac{\partial}{\partial \alpha} \ell_n(\hat{\theta}^c) \end{pmatrix}, \quad \frac{\partial}{\partial \alpha} \ell_n(\hat{\theta}^c) = \frac{1}{2n} \sum_{t=1}^n \frac{1}{\hat{\omega}^c} \left(\frac{\epsilon_t^2}{\hat{\omega}^c} - 1 \right) \begin{pmatrix} \epsilon_{t-1}^2 \\ \vdots \\ \epsilon_{t-q}^2 \end{pmatrix}$$

on obtient

$$\begin{aligned}
 LM_n &= \frac{1}{n} \frac{\partial}{\partial \alpha'} \ell_n(\hat{\theta}^c) \hat{I}^{22} \frac{\partial}{\partial \alpha'} \ell_n(\hat{\theta}^c) \\
 &= \frac{1}{n} \sum_{h=1}^q \left\{ \frac{1}{\hat{\kappa}_\eta - 1} \sum_{t=1}^n \left(\frac{\epsilon_t^2}{\hat{\omega}^c} - 1 \right) \frac{\epsilon_{t-h}^2}{\hat{\omega}^c} \right\}.
 \end{aligned} \tag{2.32}$$

Equivalence avec un test portemanteau

En utilisant

$$\sum_{t=1}^n \left(\frac{\epsilon_t^2}{\hat{\omega}^c} - 1 \right) = 0 \quad \text{et} \quad \sum_{t=1}^n \left(\frac{\epsilon_t^2}{\hat{\omega}^c} - 1 \right)^2 = \hat{\kappa}_\eta - 1,$$

la relation (2, 32) donne

$$LM_n = n \sum_{h=1}^q \hat{\rho}_{\epsilon^2}(h) \tag{2.33}$$

ce qui permet de conclure que le test LM est équivalent à un test portemanteau sur les carrés.

Estimation d'un modèle GARCH

3.1 Estimation d'un modèle ARCH par les moindres carrés

La méthode d'estimation la plus simple pour les modèles *ARCH* est celle des moindres carrés ordinaires (*MCO*). Cette procédure d'estimation a pour avantage la simplicité numérique, mais présente deux inconvénients :

- (i) L'estimateur *MCO* n'est pas efficace en comparaisant par les méthodes basées sur la vraisemblance ou quasi-vraisemblance qui seront présentées dans les prochains sections.
- (ii) Afin de fournir des estimateurs asymptotiquement normaux, la méthode nécessite des moments d'ordre 8 du processus observé.

Une extension de la méthode *MCO*, est la méthode des Moindres Carrés Quasi-Généralisés (*MCQG*), qui permet d'éliminer le premier inconvénient et de réduire le second, elle fournit des estimateurs asymptotiquement aussi précis que le quasi-maximum de vraisemblance sous des hypothèses de moments d'ordre 4. Notons que les méthodes des moindres carrés sont intéressantes en pratique car elles fournissent des estimateurs initiaux pour la procédure d'optimisation

utilisée dans la méthode du quasi-maximum de vraisemblance.

Nous commençons par les estimateurs *MCO* et *MCQG* non contraints. Nous verrons ensuite comment tenir compte des contraintes de positivités sur les paramètres.

3.1.1 Estimation de modèle ARCH(q) par MCO

Dans cette section, nous considérons l'estimation par *MCO* du modèle *ARCH*(q) :

$$\begin{cases} \epsilon_t = \sigma_t \eta_t \\ \sigma_t^2 = \omega_0 + \sum_{i=1}^q \alpha_{0i} \epsilon_{t-i}^2 \end{cases} \quad (3.1)$$

avec $\omega_0 > 0$, $\alpha_{0i} \geq 0$, $i = 1, \dots, q$, (η_t) est une suite *iid*, $\mathbb{E}(\eta_t) = 0$, $\text{Var}(\eta_t) = 1$.

La méthode *MCO* utilise la représentation *AR* sur les carrés du processus observé. Aucune hypothèse ne s'est faite sur la loi de η_t .

La vraie valeur du vecteur des paramètres est notée $\theta_0 = (\omega_0, \alpha_{01}, \dots, \alpha_{0q})'$ et nous notons θ une valeur quelconque des paramètres.

De (3.1) on obtient la représentation *AR*(q)

$$\epsilon_t^2 = \omega_0 + \sum_{i=1}^q \alpha_{0i} \epsilon_{t-i}^2 + u_t \quad (3.2)$$

où $u_t = \epsilon_t^2 - \sigma_t^2 = (\eta_t^2 - 1)\sigma_t^2$.

La suite $(u_t, \mathcal{F}_t)_t$ constitue une différence de martingale lorsque $\mathbb{E}_t^2 \epsilon_t = \sigma_t^2 < \infty$, telle que \mathcal{F}_t le σ -algèbre engendré par $\{\epsilon_s : s \leq t\}$.

Supposons que l'on dispose des observations $\epsilon_1, \dots, \epsilon_n$, de longueur n du processus (ϵ_t) , et soit $\epsilon_0, \dots, \epsilon_{1-q}$ les valeurs initiales. Par exemple, les valeurs initiales peuvent être choisies nulles.

Introduisons le vecteur

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

d'après (3.2) on obtient le système

$$\epsilon_t^2 = Z'_{t-1} \theta_0 + u_t, \quad t = 1, \dots, n, \quad (3.3)$$

qui peut s'écrire comme

$$Y = X\theta_0 + U,$$

définissons la matrice X de type $n \times q$

$$X = \begin{pmatrix} 1 & \epsilon_0^2 & \dots & \epsilon_{1-q}^2 \\ \vdots & & & \\ 1 & \epsilon_{n-1}^2 & \dots & \epsilon_{n-q}^2 \end{pmatrix} = \begin{pmatrix} Z'_0 \\ \vdots \\ Z'_{n-1} \end{pmatrix},$$

et puis les vecteurs $n \times 1$

$$Y = \begin{pmatrix} \epsilon_1^2 \\ \vdots \\ \epsilon_n^2 \end{pmatrix}, \quad U = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}.$$

Supposons que la matrice $X'X$ est inversible, ou de manière équivalente que X de plein rang colonne (nous allons voir que c'est toujours le cas asymptotiquement, et donc pour n assez grand). On en déduit l'estimateur *MCO* de θ_0 :

$$\hat{\theta}_n := \arg \min_{\theta} \|Y - X\theta\|^2 = (X'X)^{-1}X'Y. \quad (3.4)$$

Sous les hypothèses **MCO1** et **MCO2** ci-dessous, la variance de u_t existe et est constante.

L'estimateur des *MCO* de $\sigma_0^2 = \text{Var}_{\theta_0}(u_t)$ est

$$\hat{\sigma}^2 = \frac{1}{n-q-1} \|Y - X\hat{\theta}_0\|^2 = \frac{1}{n-q-1} \sum_{t=1}^n \left\{ \epsilon_t^2 - \hat{\omega} - \sum_{i=1}^q \hat{\alpha}_i \epsilon_{t-i}^2 \right\}^2.$$

Remarque 3.1.1 (*Estimateur MCO d'un GARCH*)

Un estimateur *MCO* peut également être défini pour un modèle *GARCH*(p, q), mais l'estimateur n'est pas explicite, car ϵ_t^2 ne satisfait pas un modèle *AR* quand $p \neq 0$.

Pour établir la convergence des estimateurs *MCO* de θ_0 et σ_0^2 , nous devons considérer les hypothèses suivantes.

MCO1 : (ϵ_t^2) est la solution strictement stationnaire non anticipative du modèle (3.1), et $\omega_0 > 0$.

MCO2 : $\mathbb{E}\epsilon_t^4 < \infty$.

MCO3 : $\mathbb{P}(\eta_t^2 = 1) \neq 1$.

Théorème 3.1.1 (*Convergence de l'estimateur MCO d'un modèle ARCH*)

Soit $(\hat{\theta}_n)$ une suite d'estimateurs satisfaisant (3.4). Sous les hypothèses **MCO1-MCO3**, presque sûrement

$$\hat{\theta}_n \longrightarrow \theta_0, \quad \hat{\sigma}_n^2 \longrightarrow \sigma_0^2, \quad \text{qand } n \longrightarrow \infty.$$

Preuve

La preuve comporte plusieurs étapes.

(i) Nous savons que pour une telle représentation du modèle $GARCH(p, q)$ semi fort, (ϵ_t) , l'unique solution stationnaire non anticipative du modèle, est ergodique. Le processus (Z_t) est aussi ergodique car Z_t est une fonction mesurable de $\{\epsilon_{t-i}, i \geq 0\}$. Le théorème ergodique appliqué aux composantes de la matrice $Z_{t-1}Z'_{t-1}$ implique alors que

$$\frac{1}{n}X'X = \frac{1}{n} \sum_{t=1}^n Z_{t-1}Z'_{t-1} \longrightarrow \mathbb{E}(Z_{t-1}Z'_{t-1}), \quad p.s. \quad \text{quand } n \longrightarrow \infty. \quad (3.5)$$

L'existence de l'espérance est garantie par l'hypothèse **MCO3**. Notons que les valeurs initiales n'interviennent que dans un nombre fixe de termes de la somme, et n'ont donc pas d'effet sur le résultat asymptotique. De même, nous avons

$$\frac{1}{n}X'U = \frac{1}{n} \sum_{t=1}^n Z_{t-1}u_t \longrightarrow \mathbb{E}(Z_{t-1}u_t), \quad \text{quand } n \longrightarrow \infty.$$

(ii) Montrons par l'absurde l'inversibilité de la matrice $\mathbb{E}Z_{t-1}Z'_{t-1} = \mathbb{E}Z_tZ_t$. Supposons qu'il existe un vecteur non nul c de \mathbb{R}^{q+1} tel que $c'\mathbb{E}Z_tZ'_t c = 0$. Ainsi $\mathbb{E}\{c'Z_t\}^2 = 0$, et il s'ensuit que $c'Z_t = 0$ p.s. Il existe donc une combinaison linéaire presque sûrement égale à une constante des variables $\epsilon_t^2, \dots, \epsilon_{t-q+1}^2$. On peut supposer sans perte de généralité que, dans cette combinaison linéaire, le coefficient de $\epsilon_t^2 = \eta_t^2 \sigma_t^2$ vaut 1. Ainsi η^2 est p.s. s'exprime comme fonction mesurable des variables $\epsilon_{t-1}, \dots, \epsilon_{t-q}$. Cependant, la solution étant non anticipative, η_t^2 est indépendante de ces variables.

Ceci implique que η_t^2 est p.s. égale à une constante. Cette constante est nécessairement égale à 1, mais cela conduit à une contradiction avec **MCO3**. Ainsi $\mathbb{E}(Z_{t-1}Z'_{t-1})$ est inversible.

(iii) L'innovation de ϵ_t^2 étant $u_t = \epsilon_t^2 - \sigma_t^2 = \epsilon_t^2 - \mathbb{E}(\epsilon_t^2 | \mathcal{F}_{t-1})$, on a les relations d'orthogonalité

$$\mathbb{E}(u_t) = \mathbb{E}(u_t \epsilon_{t-1}^2) = \dots = \mathbb{E}(u_t \epsilon_{t-q}^2) = 0$$

c'est-à-dire

$$\mathbb{E}(Z_{t-1}u_t) = 0.$$

(iv) La partie (ii) montre que $n^{-1}X'X$ est p.s. inversible, pour n assez grand et que, p.s. quand $n \rightarrow \infty$,

$$\hat{\theta}_n - \theta_0 = \left(\frac{X'X}{n} \right)^{-1} \frac{X'U}{n} \rightarrow \{ \mathbb{E}(Z_{t-1}Z'_{t-1}) \}^{-1} \mathbb{E}(Z_{t-1}u_t) = 0.$$

Donc d'après (ii) et (iii) $\hat{\theta}_n$ converge p.s. vers θ_0 .

La convergence de $\hat{\sigma}_n^2$ vers σ_0^2 s'en déduit. □

Pour la normalité asymptotique de l'estimateur MCO, nous avons besoin de l'hypothèse supplémentaire suivante.

MCO4 : $\mathbb{E}(\epsilon_t^8) < +\infty$.

Considérons les matrices $(q+1) \times (q+1)$

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

L'inversibilité de A a été établie dans la preuve du théorème 3.1.1, et l'inversibilité de B est montré par le même argument, notant que $c'\sigma_t^2 Z_{t-1} = 0$ si et seulement si $c'Z_{t-1} = 0$ car $\sigma_t^2 > 0$.

Le résultat suivant établit la normalité asymptotique de l'estimateur MCO.

Soit $\kappa_\eta = \mathbb{E}\eta_t^4$.

Théorème 3.1.2 (Normalité asymptotique de l'estimateur MCO)

Sous les hypothèses **MCO1-MCO4**,

$$\sqrt{n}(\hat{\theta}_n - \theta_n) \xrightarrow{\mathcal{L}} \mathcal{N}(0, (\kappa_\eta - 1)A^{-1}BA^{-1}).$$

Preuve

D'après (3.3), on a

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

Ainsi

$$\sqrt{n}(\hat{\theta}_n - \theta_n) = \left(\frac{1}{n} \sum_{t=1}^n Z_{t-1} Z'_{t-1} \right)^{-1} \left\{ \frac{1}{\sqrt{n}} \sum_{t=1}^n Z_{t-1} u_t \right\} \quad (3.6)$$

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

$$\begin{aligned} \text{Var}(\lambda' Z_{t-1} u_t) &= \lambda' E(Z_{t-1} Z'_{t-1} u_t^2) \lambda \\ &= \lambda' E \{ Z_{t-1} Z'_{t-1} (u_t^2 - 1)^2 \sigma_t^4 \} \lambda \\ &= (\kappa_\eta - 1) \lambda' B \lambda \end{aligned}$$

On obtient avec TCL que, pour tout $\lambda \neq 0$,

$$\frac{1}{\sqrt{n}} \sum_{t=1}^n \lambda' Z_{t-1} u_t \xrightarrow{\mathcal{L}} \mathcal{N}(0, (\kappa_\eta - 1) \lambda' B \lambda).$$

Par suite en appliquant la propriété de Cramèr-Wold, il s'ensuit que

$$\frac{1}{\sqrt{n}} \sum_{t=1}^n Z_{t-1} u_t \xrightarrow{\mathcal{L}} \mathcal{N}(0, (\kappa_\eta - 1) B). \quad (3.7)$$

La conclusion découle de (3.5), (3.6) et (3.7), c'est-à-dire que $\sqrt{n}(\hat{\theta}_n - \theta_n)$ est asymptotiquement normal de moyenne le vecteur nul et de variance la matrice du théorème.

□

Remarque 3.1.2 (*Estimation des matrices d'information*)

Les estimateurs convergents \hat{A} et \hat{B} des matrices A et B obtenus en remplaçant les moments théoriques par leurs homologues empiriques,

$$\hat{A} = \frac{1}{n} \sum_{t=1}^n Z_{t-1} Z'_{t-1}, \quad \hat{B} = \frac{1}{n} \sum_{t=1}^n \hat{\sigma}_t^4 Z_{t-1} Z'_{t-1},$$

où $\hat{\sigma}_t^2 = Z_{t-1} \hat{\theta}_n$. Le moment du quatrième ordre du processus $\eta_t = \epsilon_t / \sigma_t$ est également estimé de manière convergente par $\hat{\mu}_4 = n^{-1} \sum_{t=1}^n (\epsilon_t / \hat{\sigma}_t)^4$. On en déduit finalement un estimateur convergent de la variance asymptotique de l'estimateur MCO, défini par

$$\hat{\text{Var}}_{p.s} \left\{ \sqrt{n}(\hat{\theta}_n - \theta_0) \right\} = (\hat{\mu}_4 - 1) \hat{A}^{-1} \hat{I} \hat{A}^{-1}.$$

3.1.2 Estimation par MCQG des modèles ARCH(q)

Dans un modèle de régression linéaire lorsque, conditionnellement aux variables exogènes, les erreurs sont hétéroscédastiques, l'estimateur *MCQG* est asymptotiquement plus précis que l'estimateur *MCO*. Notons que dans (3.3) les erreurs u_t sont, conditionnellement à Z_{t-1} , hétéroscédastiques de variance conditionnelle $Var(u_t|Z_{t-1}) = (\kappa_\eta - 1)\sigma_t^4$. Notons aussi, pour tout $\theta = (\omega, \alpha_1, \dots, \alpha_q)$,

$$\sigma_t^2(\theta) = \omega + \sum_{i=1}^q \alpha_i \epsilon_{t-i}^2 \quad \text{et} \quad \hat{\Omega} = \text{diag}(\sigma_1^{-4}(\hat{\theta}_n), \dots, \sigma_n^{-4}(\hat{\theta}_n)).$$

L'estimateur *MCQG* est défini par

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

Théorème 3.1.3 (*Propriétés asymptotiques de l'estimateur MCQG*)

Sous les hypothèses MCO1-MCO3 et si $\alpha_{0i} > 0$, $i = 1, \dots, q$,

$$\tilde{\theta}_n \longrightarrow \theta_n, \quad p.s. \quad \sqrt{n}(\tilde{\theta}_n - \theta_n) \xrightarrow{\mathcal{L}} \mathcal{N}(0, (\kappa_\eta - 1)J^{-1}),$$

où $J = \mathbb{E}(\sigma_t^{-4} Z_{t-1} Z'_{t-1})$ est définie positive.

Preuve

On peut montrer que J est définie positive par l'argument utilisé dans le théorème 3.1.1

1. Nous avons

$$\begin{aligned} \tilde{\theta}_n &= \left(\frac{1}{n} \sum_{t=1}^n \sigma_t^{-4}(\hat{\theta}_n) Z_{t-1} Z'_{t-1} \right)^{-1} \left(\frac{1}{n} \sum_{t=1}^n \sigma_t^{-4}(\hat{\theta}_n) Z_{t-1} \epsilon_t^2 \right) \\ &= \left(\frac{1}{n} \sum_{t=1}^n \sigma_t^{-4}(\hat{\theta}_n) Z_{t-1} Z'_{t-1} \right)^{-1} \left\{ \frac{1}{n} \sum_{t=1}^n \sigma_t^{-4}(\hat{\theta}_n) Z_{t-1} (Z'_{t-1} \theta_0 + u_t) \right\} \\ &= \theta_0 + \left(\frac{1}{n} \sum_{t=1}^n \sigma_t^{-4}(\hat{\theta}_n) Z_{t-1} Z'_{t-1} \right)^{-1} \left\{ \frac{1}{n} \sum_{t=1}^n \sigma_t^{-4}(\hat{\theta}_n) Z_{t-1} u_t \right\}. \end{aligned} \quad (3.8)$$

Un développement de Taylor autour de θ_0 donne, avec $\sigma_t^2 = \sigma_t^2(\theta_0)$,

$$\sigma_t^{-4}(\hat{\theta}_n) = \sigma_t^{-4} - 2\sigma_t^{-6}(\theta^*) \frac{\partial \sigma_t^2}{\partial \theta'}(\theta^*)(\hat{\theta}_n - \theta_0), \quad (3.9)$$

où θ^* est compris entre $\hat{\theta}_n$ et θ_0 . Notons que, pour tout θ , $\frac{\partial \sigma_t^2}{\partial \theta'}(\theta) = Z_{t-1}$. Il s'ensuit que

$$\frac{1}{n} \sum_{t=1}^n \sigma_t^{-4}(\hat{\theta}_n) Z_{t-1} Z'_{t-1} = \frac{1}{n} \sum_{t=1}^n \sigma_t^{-4} Z_{t-1} Z'_{t-1} - \frac{2}{n} \sum_{t=1}^n \sigma_t^{-6}(\theta^*) Z_{t-1} Z'_{t-1} \times_{t-1} Z'_{t-1} (\hat{\theta}_n - \theta_0).$$

Le premier terme du membre de droite de l'égalité converge p.s. vers J par le théorème ergodique. Le second terme converge p.s. vers 0 parce que l'estimateur *MCO* est convergent et

$$\begin{aligned} \left\| \frac{1}{n} \sum_{t=1}^n \sigma_t^{-6}(\theta^*) Z_{t-1} \times_{t-1} Z'_{t-1} (\hat{\theta}_n - \theta_0) \right\| &\leq \left(\frac{1}{n} \sum_{t=1}^n \|\sigma_t^{-2}(\theta^*) Z_{t-1}\|^3 \right) \|\hat{\theta}_n - \theta_0\| \\ &\leq K \|\hat{\theta}_n - \theta_0\|, \end{aligned}$$

pour n grand. La borne constante K est obtenue en remarquant que les composantes de $\hat{\theta}_n$, et donc celles de θ^* , sont strictement positives pour n assez grand (car $\hat{\theta}_n \rightarrow \theta_0$ p.s.). Ainsi, nous avons $\sigma_t^{-2}(\theta^*) \epsilon_{t-i}^2 < 1/\theta_i^*$, pour $i = 1, \dots, q$, et finalement $\|\sigma_t^{-2}(\theta^*) Z_{t-1}\|$ est borné. Nous avons donc montré que p.s.

$$\left(\frac{1}{n} \sum_{t=1}^n \sigma_t^{-4}(\hat{\theta}_n) Z_{t-1} Z'_{t-1} \right)^{-1} \longrightarrow J^{-1}. \quad (3.10)$$

Pour le terme entre accolades dans (3.8) on a

$$\begin{aligned} &\frac{1}{n} \sum_{t=1}^n \sigma_t^{-4}(\hat{\theta}_n) Z_{t-1} u_t \\ &= \frac{1}{n} \sum_{t=1}^n \sigma_t^{-4} Z_{t-1} u_t - \frac{2}{n} \sum_{t=1}^n \sigma_t^{-6}(\theta^*) Z_{t-1} u_t \times_{t-1} Z'_{t-1} (\hat{\theta}_n - \theta_0) \\ &\longrightarrow 0, \quad p.s. \end{aligned} \quad (3.11)$$

par les mêmes arguments précédents, en notant que $\mathbb{E}(\sigma_t^{-4} Z_{t-1} u_t) = 0$ et

$$\begin{aligned} &\left\| \frac{2}{n} \sum_{t=1}^n \sigma_t^{-6}(\theta^*) Z_{t-1} u_t \times_{t-1} Z'_{t-1} (\hat{\theta}_n - \theta_0) \right\| \\ &= \left\| \frac{2}{n} \sum_{t=1}^n \sigma_t^{-6}(\theta^*) Z_{t-1} \sigma_t^{-6}(\theta_0) (\eta_t^2 - 1) \times_{t-1} Z'_{t-1} (\hat{\theta}_n - \theta_0) \right\| \\ &\leq K \left(\frac{1}{n} \sum_{t=1}^n |\eta_t^2 - 1| \right) \|\hat{\theta}_n - \theta_0\| \longrightarrow 0. \end{aligned}$$

Ainsi, nous avons montré que $\tilde{\theta}_n \longrightarrow \theta_0$ p.s.

En utilisant (3.11), (3.8) et (3.10), on aura

$$\begin{aligned} \sqrt{n}(\tilde{\theta}_n - \theta_0) &= (J^{-1} + R_n) \left\{ \frac{1}{\sqrt{n}} \sum_{t=1}^n \sigma_t^{-4} Z_{t-1} u_t \right\} \\ &\quad - \frac{2}{n} (J^{-1} + R_n) \sum_{t=1}^n \sigma_t^{-6} (\theta^*) Z_{t-1} u_t \times Z'_{t-1} \sqrt{n}(\hat{\theta}_n - \theta_0), \end{aligned}$$

où $R_n \longrightarrow 0$, p.s. Un nouveau développement autour de θ_0 donne

$$\sigma_t^{-6}(\theta^*) = \sigma_t^{-6} - 3\sigma_t^{-8}(\theta^{**}) Z'_{t-1}(\hat{\theta}_n - \theta_0), \quad (3.12)$$

où θ^{**} est compris entre θ^* et θ_0 . Il s'ensuit que

$$\begin{aligned} &\sqrt{n}(\tilde{\theta}_n - \theta_0) \\ &= (J^{-1} + R_n) \left\{ \frac{1}{\sqrt{n}} \sum_{t=1}^n \sigma_t^{-2} Z_{t-1} (\eta_t^2 - 1) \right\} \\ &\quad - \frac{2}{n} (J^{-1} + R_n) \sum_{t=1}^n \sigma_t^{-4} Z_{t-1} (\eta_t^2 - 1) \times Z'_{t-1} \sqrt{n}(\hat{\theta}_n - \theta_0) \\ &\quad + \frac{6}{n^{3/2}} (J^{-1} + R_n) \sum_{t=1}^n \sigma_t^{-8} (\theta^{**}) Z_{t-1} u_t \times \left\{ Z'_{t-1} \sqrt{n}(\hat{\theta}_n - \theta_0) \right\} \times \left\{ Z'_{t-1} \sqrt{n}(\theta^* - \theta_0) \right\} \\ &: = S_{n1} + S_{n2} + S_{n3}. \end{aligned} \quad (3.13)$$

Le théorème central limite (*TCL*) appliqué à la différence de martingale stationnaire, ergodique et de carré intégrable $\sigma_t^{-2} Z_{t-1} (\eta_t^2 - 1)$ montre que S_{n1} converge en loi vers un vecteur gaussien de moyenne nulle et de variance

$$J^{-1} \mathbb{E} \left\{ \sigma_t^{-4} (\eta_t^2 - 1)^2 Z_{t-1} Z'_{t-1} \right\} J^{-1} = (\kappa_\eta - 1) J^{-1}$$

En outre,

$$\begin{aligned} &\frac{1}{n} \sum_{t=1}^n \sigma_t^{-4} Z_{t-1} (\eta_t^2 - 1) \times Z'_{t-1} \sqrt{n}(\hat{\theta}_n - \theta_0) \\ &= \left\{ \frac{1}{n} \sum_{t=1}^n \sigma_t^{-4} Z_{t-1} (\eta_t^2 - 1) \right\} \sqrt{n}(\hat{\omega}_n - \omega_0) \\ &\quad + \sum_{j=1}^q \left\{ \frac{1}{n} \sum_{t=1}^n \sigma_t^{-4} Z_{t-1} (\eta_t^2 - 1) \epsilon_{t-j}^2 \right\} \sqrt{n}(\hat{\alpha}_{nj} - \alpha_{0j}). \end{aligned}$$

Les deux termes entre accolades tendent vers 0 p.s. par le théorème ergodique. De plus, les termes $\sqrt{n}(\hat{\omega}_n - \omega_0)$ et $\sqrt{n}(\hat{\alpha}_{nj} - \alpha_{0j})$ sont bornés en probabilité, ainsi que $J^{-1} + R_n$. Il s'ensuit que S_{n2} tend vers 0 en probabilité. Finalement, par des arguments déjà utilisés et parce que θ^* est compris entre $\hat{\theta}_n$ et θ_0 ,

$$\|S_{n3}\| \leq \frac{K}{n^{1/2}} \|J^{-1} + R_n\| \left\| \sqrt{n}(\hat{\theta}_n - \theta_0) \right\|^2 \frac{1}{n} \sum_{t=1}^n |\eta_t^2 - 1| \longrightarrow 0,$$

en probabilité. En utilisant (3.12), nous avons montré la convergence en loi du théorème. \square

La normalité asymptotique de l'estimateur des *MCQG* d'un ARCH ne nécessite, que la condition de moment $\mathbb{E}(\epsilon_t^4) < \infty$. Pour l'estimateur *MCO* nous avons la condition plus restrictive $\mathbb{E}(\epsilon_t^8) < \infty$. En outre, lorsque ce moment d'ordre 8 existe, le résultat suivant montre que l'estimateur *MCO* est asymptotiquement moins précis que l'estimateur *MCQG*.

Théorème 3.1.4 (*Variances asymptotiques MCO contre MCQG*)

Sous les hypothèses MCO1 – MCO4, la matrice

$$A^{-1}BA^{-1} - J^{-1}$$

est semi-définie positive.

Preuve

Soit $D = \sigma_t^2 A^{-1} Z_{t-1} - \sigma_t^2 J^{-1} Z_{t-1}$. Alors

$$\begin{aligned} E(DD') &= A^{-1}E(\sigma_t^{-4} Z_{t-1} Z'_{t-1})A^{-1} + J^{-1}E(\sigma_t^{-4} Z_{t-1} Z'_{t-1})J^{-1} \\ &\quad - A^{-1}E(Z_{t-1} Z'_{t-1})J^{-1} - J^{-1}E(Z_{t-1} Z'_{t-1})A^{-1} \\ &= A^{-1}BA^{-1} - J^{-1} \end{aligned}$$

est semi-défini positif, ce qui prouve le résultat. \square

3.1.3 Estimation par Moindres Carrés Ordinaires contraints

Les composantes négatives ne sont pas exclues et ils sont présentes dans l'estimateur *MCO* $\hat{\theta}_n$ défini par (3.4).

Lorsque l'estimation a des composantes négatives, les prédictions de la volatilité peuvent être négatives. Afin d'éviter ce problème, nous considérons l'estimateur *MCO* contraint défini par

$$\begin{aligned}\hat{\theta}_n^c &= \arg \min_{\theta \in [0, \infty[^{q+1}} Q_n(\theta), \\ Q_n(\theta) &= \frac{1}{n} \|Y - X\theta\|^2.\end{aligned}$$

L'existence de $\hat{\theta}_n^c$ contraint est garantie par la continuité de la fonction Q_n et le fait que

$$\{nQ_n(\theta)\}^{1/2} \geq \|X\theta\| - \|Y\| \longrightarrow \infty$$

lorsque $\|\theta\| \longrightarrow \infty$ et $\theta \geq 0$, dès que X n'a pas de colonne nulle. Il est à noter que cette dernière condition est satisfaite au moins pour n suffisamment grand.

Propriétés de l'estimateur MCO contraint

Le théorème suivant donne une condition d'égalité entre estimateurs contraint et non contraint. Le théorème est énoncé dans le cas *ARCH* mais il reste valide dans un cadre plus général.

Théorème 3.1.5 (*Égalité entre MCO contraint et non contraint*)

Si X est de rang $q + 1$, les estimateurs contraint et non contraint coïncident, $\hat{\theta}_n^c = \hat{\theta}_n$, si et seulement si $\hat{\theta}_n \in [0, \infty[^{q+1}$.

Preuve

Puisque $\hat{\theta}_n$ et $\hat{\theta}_n^c$ sont obtenus en minimisant la même fonction $Q_n(\cdot)$, et puisque $\hat{\theta}_n^c$ minimise cette fonction sur un ensemble plus petit, on a $Q_n(\hat{\theta}_n) \leq Q_n(\hat{\theta}_n^c)$. De plus, $\hat{\theta}_n^c \in [0, \infty[^{q+1}$, et on a $Q_n(\theta) \geq Q_n(\hat{\theta}_n^c)$, pour tout $\theta \in [0, \infty[^{q+1}$.

Supposons que l'estimation sans constraint $\hat{\theta}_n$ appartient à $[0, \infty[^{q+1}$. Dans ce cas $Q_n(\hat{\theta}_n) = Q_n(\hat{\theta}_n^c)$. Comme la solution non constraint est unique, $\hat{\theta}_n^c = \hat{\theta}_n$.

La réciproque est triviale.

□

Le théorème suivant présente un moyen d'obtenir l'estimateur contraint à l'aide de l'estimateur non contraint.

Théorème 3.1.6 (*MCO contraint est projection de MCO*)

Si X est de rang $q + 1$, l'estimateur $\hat{\theta}_n^c$ est la projection orthogonale de $\hat{\theta}_n$ sur $[0, \infty[^{q+1}$ par rapport à la métrique $X'X$, c'est-à-dire

$$\hat{\theta}_n^c = \arg \min_{\theta \in [0, \infty[^{q+1}} (\hat{\theta}_n - \theta)' X' X (\hat{\theta}_n - \theta). \quad (3.14)$$

Preuve

Si on note P le projecteur orthogonal sur les colonnes de X , et $M = I_n - P$, on a

$$\begin{aligned} nQ(\theta) &= \|Y - X\theta\|^2 = \|P(Y - X\theta)\|^2 + \|M(Y - X\theta)\|^2 \\ &= \|X(\hat{\theta}_n - \theta)\|^2 + \|MY\|^2, \end{aligned}$$

en utilisant les propriétés des projecteurs, le théorème de Pythagore et $PY = X\hat{\theta}_n$. L'estimation contrainte $\hat{\theta}_n^c$ résout donc (3.14). Notons que, puisque X est de plein rang colonne, $\|x\|_{X'X} = \sqrt{x'X'Xx}$ définit bien une norme. La caractérisation (3.14) est équivalente à

$$\hat{\theta}_n^c \in [0, \infty[^{q+1}, \quad \left\| \hat{\theta}_n - \hat{\theta}_n^c \right\|_{X'X} \leq \left\| \hat{\theta}_n - \theta \right\|_{X'X}, \quad \forall \theta \in [0, \infty[^{q+1}. \quad (3.15)$$

Comme $[0, \infty[^{q+1}$ est convexe, $\hat{\theta}_n^c$ existe, est unique et est la projection $X'X$ -orthogonale de $\hat{\theta}_n$ sur $[0, \infty[^{q+1}$. Cette projection se caractérise par

$$\hat{\theta}_n^c \in [0, \infty[^{q+1} \quad \text{et} \quad \left\langle \hat{\theta}_n - \hat{\theta}_n^c, \hat{\theta}_n^c - \theta \right\rangle_{X'X} \geq 0, \quad \forall \theta \in [0, \infty[^{q+1} \quad (3.16)$$

Cette caractérisation montre que, lorsque $\hat{\theta}_n \notin [0, \infty[^{q+1}$, l'estimation contrainte $\hat{\theta}_n^c$ doit être au bord de $[0, \infty[^{q+1}$. Sinon il suffit de prendre $\theta \in [0, \infty[^{q+1}$ entre $\hat{\theta}_n^c$ et $\hat{\theta}_n$ pour obtenir un produit scalaire égal à -1 .

□

La caractérisation (3.15) permet d'obtenir facilement la convergence forte de l'estimateur contraint.

Théorème 3.1.7 (Convergence de l'estimateur MCO contraint)

Sous les hypothèses de théorème (3.1.1), presque sûrement,

$$\hat{\theta}_n^c \longrightarrow \theta_0 \quad \text{pour} \quad n \longrightarrow \infty.$$

Preuve

Puisque $\theta_0 \in [0, \infty[^{q+1}$, on a d'après (3.15)

$$\left\| \hat{\theta}_n - \hat{\theta}_n^c \right\|_{X'X/n} \leq \left\| \hat{\theta}_n - \theta_0 \right\|_{X'X/n}.$$

Il s'ensuit que, en utilisant l'inégalité triangulaire,

$$\left\| \hat{\theta}_n^c - \theta_0 \right\|_{X'X/n} \leq \left\| \hat{\theta}_n^c - \hat{\theta}_n \right\|_{X'X/n} + \left\| \hat{\theta}_n - \theta_0 \right\|_{X'X/n} \leq 2 \left\| \hat{\theta}_n - \theta_0 \right\|_{X'X/n}.$$

Puisque, d'après le théorème (3.1.1), $\hat{\theta}_n \longrightarrow \theta_0$ p.s. et $X'X/n$ converge p.s. vers une matrice définie positive, il s'ensuit que $\left\| \hat{\theta}_n - \theta_0 \right\|_{X'X/n} \longrightarrow 0$ et donc que $\left\| \hat{\theta}_n - \hat{\theta}_n^c \right\|_{X'X/n} \longrightarrow 0$ p.s. \square

Calcul de l'estimateur MCO contraint

Nous donnons une manière explicite pour obtenir l'estimateur contraint. Nous avons déjà vu que si toutes les composantes de l'estimateur non contraint $\hat{\theta}_n$ sont positives, on a évidemment $\hat{\theta}_n^c = \hat{\theta}_n$. Supposons maintenant qu'une composante de $\hat{\theta}_n$ est négative, par exemple la dernière composante. On pose

$$X = (X^{(1)}, X^{(2)}), \quad X^{(1)} = \begin{pmatrix} 1 & \epsilon_0^2 & \cdots & \epsilon_{-q+2}^2 \\ 1 & \epsilon_1^2 & \cdots & \epsilon_{-q+3}^2 \\ \vdots & & & \vdots \\ 1 & \epsilon_{n-1}^2 & \cdots & \epsilon_{n-q+1}^2 \end{pmatrix},$$

et

$$\hat{\theta}_n = (X'X)^{-1}X'Y = \begin{pmatrix} \hat{\theta}_n^{c(1)} \\ \hat{\alpha}_q \end{pmatrix}, \quad \tilde{\theta}_n = \begin{pmatrix} \tilde{\theta}_n^{(1)} \\ 0 \end{pmatrix} = \begin{pmatrix} (X^{(1)'}X^{(1)})^{-1}X^{(1)'}Y \\ 0 \end{pmatrix}.$$

Notons que $\hat{\theta}_n^{c(1)} \neq \tilde{\theta}_n^{(1)}$ en général.

Théorème 3.1.8 (*Forme explicite de l'estimateur contraint*)

Supposons que X est de rang $q + 1$ et $\hat{\alpha}_q < 0$. Alors

$$\tilde{\theta}_n^{(1)} \in [0, +\infty[^q \iff \hat{\theta}_n^c = \tilde{\theta}_n.$$

Preuve

Soit $P^{(1)} = X^{(1)}(X^{(1)'}X^{(1)})^{-1}X^{(1)}$ le projecteur sur les colonnes de $X^{(1)}$ et soit $M^{(1)} = I - P^{(1)}$.

Nous avons

$$\begin{aligned}\tilde{\theta}'_n X' &= (Y'X^{(1)}(X^{(1)'}X^{(1)})^{-1}, 0) \begin{pmatrix} X^{(1)'} \\ X^{(2)'} \end{pmatrix} = Y'P^{(1)}, \\ \tilde{\theta}'_n X'X &= (Y'X^{(1)}, Y'P^{(1)}X^{(2)}), \\ \tilde{\theta}'_n X'X &= Y'X = (Y'X^{(1)}, Y'X^{(2)}), \\ (\hat{\theta}'_n - \tilde{\theta}'_n)X'X &= (0, Y'M^{(1)}X^{(2)}).\end{aligned}$$

Parce que $\hat{\theta}_n e_{q+1} < 0$, avec $e_{q+1} = (0, \dots, 0, 1)'$, on a $(\hat{\theta}'_n - \tilde{\theta}'_n)e_{q+1} < 0$. Cela peut être écrit comme

$$(\hat{\theta}'_n - \tilde{\theta}'_n)X'X(X'X)^{-1}e_{q+1} < 0,$$

Ou bien

$$Y'M^{(1)}X^{(2)}\{(X'X)^{-1}\}_{q+1, q+1} < 0.$$

Ainsi $Y'M^{(1)}X^{(2)} < 0$. Il s'ensuit que pour tout $\theta = (\theta^{(1)'}, \theta^{(2)'})$ tel que $\theta^{(2)} \in [0, \infty[$,

$$\begin{aligned}\left\langle \hat{\theta}_n - \tilde{\theta}_n, \tilde{\theta}_n - \theta \right\rangle_{X'X} &= (\hat{\theta}_n - \tilde{\theta}_n)'X'X(\tilde{\theta}_n - \theta) \\ &= \left(0, Y'M^{(1)}X^{(2)}\right) \begin{pmatrix} \tilde{\theta}_n^{(1)} - \theta^{(1)} \\ -\theta^{(2)} \end{pmatrix} \\ &= -\theta^{(2)'}Y'M^{(1)}X^{(2)} \geq 0.\end{aligned}$$

Au vu de (3.16), on a $\hat{\theta}_n^c = \tilde{\theta}_n$ car $\tilde{\theta}_n \in [0, +\infty[^{q+1}$.

□

3.2 Estimation par Quasi-Maximum de Vraisemblance des modèles GARCH

La méthode du quasi-maximum de vraisemblance (*QMV*) est particulièrement pertinente pour les modèles *GARCH* car elle fournit des estimateurs convergents et asymptotiquement normaux pour les processus *GARCH* strictement stationnaires sous des conditions de régularité simples, mais sans hypothèses de moment sur le processus observé.

En revanche, les méthodes des moindres carrés de la section précédente nécessitent au moins des moments d'ordre 4.

Dans cette section, en premier temps, nous étudions en détail la méthode *QMV* conditionnelle (conditionnelle aux valeurs initiales).

Nous considérons d'abord le cas où le processus observé est un *GARCH* pur. Nous présentons une procédure itérative du calcul de la *log - vraisemblance* gaussienne, conditionnellement à des valeurs initiales fixes ou aléatoires. La vraisemblance s'écrit comme si la loi des variables η_t était gaussienne $\mathcal{N}(0, 1)$ (on parle de pseudo ou quasi-vraisemblance), cette hypothèse n'est pas nécessaire pour la convergence forte de l'estimateur. Dans la deuxième partie, nous étudierons l'application de la méthode pour l'estimation des modèles *ARMA - GARCH*. Les propriétés asymptotiques de l'estimateur quasi-maximum de vraisemblance (*QMV*) sont établies à la fin de section.

3.2.1 Quasi-vraisemblance conditionnelle

Supposons que les observations $\epsilon_1, \dots, \epsilon_n$ constituent une réalisation (de longueur n) d'un processus *GARCH*(p, q), plus précisément une solution strictement stationnaire et non anticipative du modèle

$$\begin{cases} \epsilon_t = \sqrt{h_t} \eta_t \\ h_t = \omega_0 + \sum_{i=1}^q \alpha_{0i} \epsilon_{t-i}^2 + \sum_{j=1}^p \beta_{0j} h_{t-j}, \quad \forall t \in \mathbb{Z} \end{cases} \quad (3.16)$$

où (η_t) est une suite de variables *iid* de variance 1, $\omega_0 > 0$, $\alpha_{0i} \geq 0$, ($i = 1, \dots, q$) et $\beta_{0j} \geq 0$, ($j = 1, \dots, p$). Les ordres p et q sont supposés connus. Le vecteur des paramètres

$$\theta = (\theta_1, \dots, \theta_{p+q+1})' = (\omega, \alpha_1, \dots, \alpha_q, \beta_1, \dots, \beta_p)' \quad (3.17)$$

appartient à un espace de paramètres de la forme

$$\Theta \subset]0, +\infty[\times [0, \infty[^{p+q}. \tag{3.18}$$

La vraie valeur du paramètre est inconnue et est notée

$$\theta_0 = (\omega_0, \alpha_{01}, \dots, \alpha_{0q}, \beta_{01}, \dots, \beta_{0p})'.$$

Pour écrire la vraisemblance du modèle, il faut spécifier une distribution particulière pour les variables *iid* η_t .

Aucune hypothèse exigée sur la distribution de ces variables, mais en général on utilise une quasi-vraisemblance gaussienne, qui conditionnellement à certaines valeurs initiales, coïncide avec une vraisemblance quand les variables η_t sont distribuées par une loi gaussienne standard (centrée réduite).

Étant donné les valeurs initiales $\epsilon_0, \dots, \epsilon_{1-q}, \tilde{\sigma}_0^2, \dots, \tilde{\sigma}_{1-p}^2$ à préciser ci-dessous, la quasi-vraisemblance conditionnelle gaussienne est donnée comme suit

$$L_n(\theta) = L_n(\theta, \epsilon_1, \dots, \epsilon_n) = \prod_{t=1}^n \frac{1}{\sqrt{2\pi\tilde{\sigma}_t^2}} \exp\left(-\frac{\epsilon_t^2}{2\tilde{\sigma}_t^2}\right),$$

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

$$\tilde{\sigma}_t^2 = \tilde{\sigma}_t^2(\theta) = \omega + \sum_{i=1}^q \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^p \beta_j \tilde{\sigma}_{t-j}^2 \tag{3.19}$$

Pour une valeur donnée de θ , sous l'hypothèse de stationnarité du second ordre, la variance inconditionnelle (correspondant à cette valeur de θ) est un choix raisonnable pour les valeurs initiales inconnues :

$$\epsilon_0^2 = \dots = \epsilon_{1-q}^2 = \sigma_0^2 = \dots = \sigma_{1-p}^2 = \frac{\omega}{1 - \sum_{i=1}^q \alpha_i + \sum_{j=1}^p \beta_j} \tag{3.20}$$

De telle valeurs initiales ne sont cependant pas adaptées aux modèles IGARCH, en particulier, et plus généralement lorsque la stationnarité du second ordre n'est pas vérifiée, car la constante dans (3.20) peut prendre des valeurs négatives pour certaines valeurs de θ . Dans un tel cas, des valeurs initiales appropriées sont

$$\epsilon_0^2 = \dots = \epsilon_{1-q}^2 = \tilde{\sigma}_0^2 = \dots = \tilde{\sigma}_{1-p}^2 = \omega \tag{3.21}$$

ou encore

$$\epsilon_0^2 = \dots = \epsilon_{1-q}^2 = \tilde{\sigma}_0^2 = \dots = \tilde{\sigma}_{1-p}^2 = \epsilon_1^2 \tag{3.22}$$

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

$$\hat{\theta}_n = \arg \max_{\theta \in \Theta} L_n(\theta).$$

En prenant le logarithme, remarquant que maximiser la vraisemblance équivalent à minimiser par rapport à θ

$$\tilde{\mathbf{I}}_n(\theta) = n^{-1} \sum_{t=1}^n \tilde{\ell}_t, \quad \text{où} \quad \tilde{\ell}_t = \tilde{\ell}_t(\theta) = \frac{\epsilon_t^2}{\tilde{\sigma}_t^2} + \log \tilde{\sigma}_t^2 \tag{3.23}$$

et $\tilde{\sigma}_t^2$ est définie par (3.19). Un estimateur *QMV* est donc une solution mesurable de l'équation

$$\hat{\theta}_n = \arg \min_{\theta \in \Theta} \tilde{\mathbf{I}}_n(\theta). \tag{3.24}$$

On montrera que le choix des valeurs initiales n'a pas d'effet sur les propriétés asymptotiques de l'estimateur de *QMV*. Cependant, en pratique ce choix peut être important. Notons que d'autres méthodes sont possibles pour générer la suite $\tilde{\sigma}_t^2$, par exemple, en prenant $\tilde{\sigma}_t^2 = c_0(\theta) + \sum_{i=1}^{t-1} c_i(\theta)\epsilon_{t-i}^2$, où les $c_i(\theta)$ sont calculés récursivement. Remarquons que pour calculer $\tilde{\mathbf{I}}_n(\theta)$, cette procédure implique un certain nombre d'opérations d'ordre n^2 , alors que celui que nous proposons comporte un nombre d'ordre n . Il sera commode d'approximer la suite $\tilde{\ell}_t(\theta)$ par une suite stationnaire ergodique. En supposant que les racines de $\mathcal{B}_\theta(z) = 1 - \sum_{j=1}^p \beta_j z^j$, sont à l'extérieure du disque unité, soit la suite non anticipative et ergodique strictement stationnaire $(\sigma_t^2)_t = \{\sigma_t^2(\theta)\}_t$ comme solution de

$$\sigma_t^2 = \omega + \sum_{i=1}^q \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2, \quad \forall t. \tag{3.25}$$

Notons que $\sigma_t^2(\theta_0) = h_t$.

Equations de vraisemblance

Les équations de vraisemblance sont obtenues en annulant la dérivée par rapport à θ du critère $\tilde{\mathbf{I}}_n(\theta)$ ce qui donne

$$\frac{1}{n} \sum_{t=1}^n \{\epsilon_t^2 - \tilde{\sigma}_t^2\} \frac{1}{\tilde{\sigma}_t^4} \frac{\partial \tilde{\sigma}_t^2}{\partial \theta} = 0. \tag{3.26}$$

Ces équations peuvent être interprétées pour n grand comme des relations d'orthogonalité. En effet, comme on le verra dans la section suivante, le membre de gauche de l'équation (3.26) a le même comportement asymptotique que

$$\frac{1}{n} \sum_{t=1}^n \{\epsilon_t^2 - \sigma_t^2\} \frac{1}{\sigma_t^4} \frac{\partial \sigma_t^2}{\partial \theta}, \tag{3.27}$$

l'influence des valeurs initiales étant nulle lorsque $n \rightarrow \infty$. L'innovation de ϵ_t^2 est $v_t = \epsilon_t^2 - h_t^2$. Ainsi, sous l'hypothèse que l'espérance existe, on a,

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

parce que $\frac{1}{\sigma_t^4(\theta_0)} \frac{\partial \sigma_t^2(\theta_0)}{\partial \theta}$ est une fonction mesurable des $\epsilon_{t-i}^2, i > 0$. Ce résultat peut être considéré comme le version asymptotique de (3.26) en θ_0 , en utilisant le théorème ergodique.

I) Propriétés asymptotiques du QMV

Nous utiliserons la norme matricielle définie par $\|A\| = \sum |a_{ij}|$ pour toutes les matrices $A = (a_{ij})$.

Le rayon spectral d'une matrice carrée A est noté $\rho(A)$.

Convergence forte

Rappelons que le modèle (3.16) admet une solution strictement stationnaire si et seulement si la suite de matrices $A_0 = (A_{0t})$, où

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

admet un coefficient de Lyapunov strictement négatif, $\gamma(A_0) < 0$, où

$$\begin{aligned} \gamma(A_0) &= \inf_{t \in \mathbb{N}^*} \frac{1}{t} \mathbb{E}(\log \|A_{0t} A_{0t-1} \dots A_{01}\|) \\ &= \lim_{t \rightarrow \infty} \frac{1}{t} \log \|A_{0t} A_{0t-1} \dots A_{01}\|. \end{aligned} \tag{3.28}$$

Soit,

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

Par convention, $\mathcal{A}_\theta(z) = 0$ si $q = 0$ et $\mathcal{B}_\theta(z) = 1$ si $p = 0$. Pour montrer la convergence forte, les hypothèses suivantes seront nécessaires.

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

A2 : $\gamma(A_0) < 0$ et pour tout $\theta \in \Theta$, $\sum_{j=1}^p \beta_j < 1$.

A3 : η_t^2 a une distribution non dégénérée et $\mathbb{E}\eta_t^2 = 1$.

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

Notons que, la deuxième partie de l'hypothèse **A2** implique que les racines de $\mathcal{B}_\theta(z)$ sont en dehors du disque unité. Ainsi, (3.25) définit bien une suite (σ_t^2) strictement stationnaire ergodique et non anticipative. De même, on définit

$$\mathbf{l}_n(\theta) = \mathbf{l}_n(\theta; \epsilon_n, \epsilon_{n-1}, \dots) = n^{-1} \sum_{t=1}^n \ell_t, \quad \text{où} \quad \ell_t = \ell_t(\theta) = \frac{\epsilon_t^2}{\sigma_t^2} + \log \sigma_t^2.$$

Théorème 3.2.1 (*Convergence forte du QMV*)

Soit $(\hat{\theta}_n)$ une suite de QMV vérifiant (3.24), avec les conditions initiales (3.21) ou (3.22). Sous les hypothèses A1 – A4, presque sûrement

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

Remarque 3.2.1 .

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

2. Il est important de noter que la condition de stationnarité stricte n'est supposée qu'à θ_0 , pas sur tout l'ensemble Θ . Et la condition $\sum_{j=1}^p \beta_j < 1$ est plus faible que celle de stationnarité strict.

3. L'hypothèse **A4** disparaît dans le cas *ARCH*. Dans le cas général, cette hypothèse 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 *GARCH*($p, q - 1$)) si un modèle *GARCH*(p, q) est utilisé.

4. Lorsque $p \neq 0$, l'hypothèse **A4** exclut le cas où tous les α_{0i} sont nuls. Dans ce cas, la solution strictement stationnaire du modèle est un bruit blanc fort, qui s'écrit de multiples formes. Par exemple, un bruit blanc fort de variance 1 peut être écrit sous la forme d'un $GARCH(1, 1)$ avec $\sigma_t^2 = \sigma^2(1 - \beta) + 0 \times \epsilon_{t-1}^2 + \beta\sigma_{t-1}^2$.

5. L'hypothèse d'absence de racines communes, en **A4**, n'est restrictive que si $p > 1$ et $q > 1$. En effet, si $q = 1$, la racine unique de $\mathcal{A}_{\theta_0}(1)$ est 0 et on a $\mathcal{B}_{\theta_0}(0) \neq 0$. Si $p = 1$ et $\mathcal{B}_{01} \neq 0$, la racine unique de $\mathcal{B}_{\theta_0}(z)$ est $1/\beta_{01} > 0$ (si $\beta_{01} = 0$, le polynôme n'admet aucune racine). Comme les coefficients α_{0i} sont positifs, cette valeur ne peut pas annuler $\mathcal{A}_{\theta_0}(z)$.

6. L'hypothèse $\mathbb{E}\eta_t = 0$ n'est pas nécessaire pour la convergence (et la normalité asymptotique) de l'estimateur du QMV d'un $GARCH$. La variance conditionnelle de ϵ_t n'est donc, en général, que proportionnelle à $h_t : \text{Var}(\epsilon_t | \epsilon_u, u < t) = \{1 - (\mathbb{E}\eta_t)^2\}h_t$. L'hypothèse $\mathbb{E}\eta_t^2 = 1$ est fait pour des raisons d'identifiabilité et n'est pas restrictive dès que $\mathbb{E}\eta_t^2 < \infty$.

Normalité asymptotique

On considère les hypothèses supplémentaires suivantes.

A5 : $\theta_0 \in \hat{\Theta}$ où $\hat{\Theta}$ désigne l'intérieur de Θ .

A6 : $\kappa_\eta = \mathbb{E}\eta_t^4 < \infty$.

La loi limite de $\hat{\theta}_n$ est donnée par le résultat suivant.

Théorème 3.2.2 (*Normalité asymptotique du QMV*)

*Sous les hypothèses **A1-A6**,*

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{\mathcal{L}} \mathcal{N}(0, (\kappa_\eta - 1)J^{-1}),$$

où

$$J = \mathbb{E}_{\theta_0} \left(\frac{\partial^2 \ell_t(\theta_0)}{\partial \theta \partial \theta'} \right) = \mathbb{E}_{\theta_0} \left(\frac{1}{\sigma_t^4(\theta_0)} \frac{\partial \sigma_t^2(\theta_0)}{\partial \theta} \frac{\partial \sigma_t^2(\theta_0)}{\partial \theta'} \right) \tag{3.29}$$

est une matrice définie positive.

Remarque 3.2.2 .

1. L'hypothèse **A5** est standard et implique les conditions du premier ordre (au moins asymptotiquement).

En effet si $\hat{\theta}_n$ est convergent, il appartient également à l'intérieur de Θ , pour n grand. En tant que maximum, il doit annuler la dérivée de la fonction critère. Cependant, l'hypothèse **A5** est restrictive car elle exclut, par exemple, le cas $\alpha_{01} = 0$.

2. Lorsqu'une ou plusieurs composantes de θ_0 sont nulles, donc l'hypothèse **A5** ne se vérifie pas et le théorème ne peut pas être utilisé. Dans ce cas, la distribution asymptotique de $\sqrt{n}(\hat{\theta}_n - \theta_0)$ ne peut pas être normale car l'estimateur est contraint. Si, par exemple, $\alpha_{01} = 0$, la distribution de $\sqrt{n}(\hat{\alpha}_{01} - \alpha_{01})$ est, pour tout n , concentré dans $[0, \infty[$, et donc ne peut pas être asymptotiquement normale.

3. L'hypothèse **A6** ne concerne pas ϵ_t^2 , et n'exclut pas le cas IGARCH. Seule une hypothèse d'existence du moment de quatrième ordre sur η_t est nécessaire. Cette hypothèse est évidemment nécessaire pour l'existence de la variance du vecteur du score $\partial \ell_t(\theta_0)/\partial \theta$. Dans la démonstration de ce théorème, on montre que

$$\mathbb{E}_{\theta_0} \left\{ \frac{\partial \ell_t(\theta_0)}{\partial \theta} \right\} = 0, \quad \text{Var}_{\theta_0} \left\{ \frac{\partial \ell_t(\theta_0)}{\partial \theta} \right\} = (\kappa_\eta - 1)J.$$

4. Dans le cas ARCH($p = 0$), la variance asymptotique du QMV se réduit à celle de l'estimateur MCQG. En effet, dans ce cas on a $\partial \sigma_t^2(\theta)/\partial \theta = Z_{t-1}$.

La validité du théorème (3.1.3) reposait cependant sur l'existence d'un moment d'ordre 4 pour le processus observé, alors qu'il n'y a pas d'hypothèse de moment pour la normalité asymptotique du QMV. par ailleurs, le théorème (3.1.4) montre que le QMV d'un ARCH(q) est asymptotiquement plus précise que celui de l'estimateur MCO.

II) Cas de ARCH(1) : Évaluation numérique de la variance asymptotique

Considérons le modèle ARCH(1)

$$\epsilon_t = \{\omega_0 + \alpha_0 \epsilon_{t-1}^2\}^{1/2} \eta_t,$$

avec $\omega_0 > 0$ et $\alpha_0 > 0$, et supposons que les variables η_t vérifient l'hypothèse **A3**. Le paramètre est $\theta = (\omega, \alpha)$, et la contrainte de stationnarité stricte **A2** s'écrit

$$\alpha_0 < \exp\{-\mathbb{E}(\log \eta_t^2)\}.$$

L'hypothèse **A1** est vraie si, par exemple, l'espace des paramètres est de la forme $\Theta = [\delta, 1/\delta] \times [0, 1/\delta]$, où $\delta > 0$ est une constante, choisie suffisamment petite pour que $\theta_0 = (\omega_0, \alpha_0)'$ appartient à Θ .

D'après le théorème (3.2.1), l'estimateur de QMV de θ est alors fortement convergent.

Puisque $\partial \tilde{\sigma}_t^2 / \partial \theta = (1, \epsilon_{t-1}^2)'$, l'estimateur de QMV , $\hat{\theta}_n = (\hat{\omega}_n, \hat{\alpha}_n)'$ est caractérisé par les équations normales

$$\frac{1}{n} \sum_{t=1}^n \frac{\epsilon_t^2 - \hat{\omega}_n + \hat{\alpha}_n \epsilon_{t-1}^2}{(\hat{\omega}_n + \hat{\alpha}_n \epsilon_{t-1}^2)^2} \begin{pmatrix} 1 \\ \epsilon_{t-1}^2 \end{pmatrix} = 0$$

avec, par exemple, $\epsilon_0^2 = \epsilon_1^2$. Cet estimateur n'a pas de forme explicite et doit être obtenu numériquement. Le théorème (3.2.2), qui donne la distribution asymptotique de l'estimateur, ne demande que l'hypothèse supplémentaire que θ_0 appartient à $\overset{\circ}{\Theta} =]\delta, 1/\delta[\times]0, 1/\delta[$. Ainsi, si $\alpha_0 = 0$ (c'est à dire si le modèle est conditionnellement homoscédastique), l'estimateur reste convergent mais n'est plus asymptotiquement normale. La matrice J prend la forme

$$J = \mathbb{E}_{\theta_0} \begin{bmatrix} \frac{1}{(\omega_0 + \alpha_0 \epsilon_{t-1}^2)^2} & \frac{\epsilon_{t-1}^2}{(\omega_0 + \alpha_0 \epsilon_{t-1}^2)^2} \\ \frac{\epsilon_{t-1}^2}{(\omega_0 + \alpha_0 \epsilon_{t-1}^2)^2} & \frac{\epsilon_{t-1}^4}{(\omega_0 + \alpha_0 \epsilon_{t-1}^2)^2} \end{bmatrix},$$

et la variance asymptotique de $\sqrt{n}(\hat{\theta}_n - \theta_0)$ est

$$\text{Var}_{as} \left\{ \sqrt{n}(\hat{\theta}_n - \theta_0) \right\} = (\kappa_\eta - 1)J^{-1}.$$

III) ARCH(1) non stationnaire

Lorsque la contrainte de stationnarité stricte n'est pas satisfaite dans le cas $ARCH(1)$, c'est-à-dire lorsque

$$\alpha_0 \geq \exp \left\{ -\mathbb{E} \log \eta_t^2 \right\}, \tag{3.30}$$

on peut définir un processus $ARCH(1)$ commençant par des valeurs initiales. Pour une valeur ϵ_0 donnée, on définit

$$\epsilon_t = h_t^{1/2} \eta_t, \quad h_t = \omega_0 + \alpha_0 \epsilon_{t-1}^2, \quad t = 1, 2, \dots, \tag{3.31}$$

où $\omega_0 > 0$ et $\alpha_0 > 0$, avec les hypothèses usuelles sur la suite (η_t) . Comme déjà noté, σ_t^2 converge presque sûrement vers l'infini quand

$$\alpha_0 > \exp \left\{ -\mathbb{E} \log \eta_t^2 \right\}, \tag{3.32}$$

et seulement en probabilité lorsque l'inégalité (3.30) est une égalité. Est-il possible d'estimer les coefficients d'un tel modèle ? La réponse n'est qu'en partie positif : il est possible d'estimer de manière convergente le coefficient α_0 , mais le coefficient ω_0 ne peut pas être systématiquement estimé. L'impact pratique de ce résultat apparaît donc limité, mais parce que de son intérêt théorique, le problème de l'estimation des coefficients des modèles non stationnaires mérite l'attention. Considérons le QMV d'un ARCH(1), c'est-à-dire une solution mesurable de

$$(\hat{\omega}_n, \hat{\alpha}_n) = \arg \min_{\theta \in \Theta} \frac{1}{n} \sum_{t=1}^n \ell_t(\theta), \quad \ell_t(\theta) = \frac{\epsilon_t^2}{\sigma_t^2(\theta)} + \log \sigma_t^2(\theta), \quad (3.33)$$

où $\theta = (\omega, \alpha)$, est un ensemble compact de $]0, \infty[^2$, et $\sigma_t^2(\theta) = \omega + \alpha \epsilon_{t-1}^2$ pour $t = 1, \dots, n$ (pour une valeur initiale donnée de ϵ_0^2). La convergence presque sûre de ϵ_t^2 à l'infini sera utilisé pour montrer la convergence forte du QMV de α_0 .

Soit le modèle ARCH(1) défini par (3.31) avec une condition initiale quelconque $\epsilon_0^2 \geq 0$. Supposée vérifie la condition de non-stationnarité (3.32). Alors, presque sûrement, quand $n \rightarrow \infty$

$$\frac{1}{h_n} = o(\rho^n) \quad \text{et} \quad \frac{1}{\epsilon_n^2} = o(\rho^n)$$

pour toute constante ρ telle que

$$1 > \rho > \exp \{ -\mathbb{E} \log \eta_t^2 \} / \alpha_0. \quad (3.34)$$

Ce résultat permet d'obtenir la convergence forte et la normalité asymptotique du QMV de α_0 .

Théorème 3.2.3 .

Considérons les hypothèses du lemme (3.2.1) et l'estimateur du QMV défini par (3.32) où $\theta_0 = (\omega_0, \alpha_0) \in \Theta$. Alors

$$\hat{\alpha}_n \rightarrow \alpha_0 \quad p.s., \quad (3.35)$$

et quand θ_0 appartient à l'intérieur de Θ ,

$$\sqrt{n}(\hat{\alpha}_n - \alpha_0) \xrightarrow{\mathcal{L}} \mathcal{N} \{ 0, (\kappa_\eta - 1) \alpha_0^2 \} \quad (3.36)$$

pour $n \rightarrow \infty$.

Dans la preuve de ce théorème, on montre que le vecteur score satisfait

$$\frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{\partial}{\partial \theta} \ell_t(\theta) \xrightarrow{\mathcal{L}} \mathcal{N} \left\{ 0, J = (\kappa_\eta - 1) \begin{pmatrix} 0 & 0 \\ 0 & \alpha_0^{-2} \end{pmatrix} \right\}.$$

Dans le cadre classique d'inférence statistique ou l'échantillonnage, la variance J du vecteur de score est (proportionnelle à) l'information de Fisher. Selon l'interprétation usuelle, la forme de la matrice J montre que, asymptotiquement et pour presque toutes les observations, les variations de la log-vraisemblance $n^{-1/2} \sum_{t=1}^n \log \ell_t(\theta)$ sont non significatifs lorsque θ varie de (ω_0, α_0) à $(\omega_0 + h, \alpha_0)$ pour h petit. En d'autres termes, la log-vraisemblance limite est plate au point (ω_0, α_0) dans la direction de variation de ω_0 . Ainsi, la minimisation de cette fonction limite ne permet pas de trouver θ_0 . Cela nous amène à penser que le QMV de ω_0 est susceptible n'est pas convergent lorsque la condition de stationnarité stricte n'est pas satisfaite.

3.2.2 Estimation par Quasi-Maximum de Vraisemblance des modèles ARMA-GARCH

Dans cette section, les résultats précédents sont étendus au cas où le processus *GARCH* n'est pas directement observé, mais constitue l'innovation d'un processus *ARMA* observé. Ce cadre est pertinent car, même pour les séries financières, il est restrictif de supposer que la série observée est la réalisation d'un bruit. D'un point de vue théorique, l'extension au cadre *ARMA-GARCH* des techniques développées dans le cas *GARCH* pur n'est pas triviale. Supposons que les observations X_1, \dots, X_n sont générés par une solution strictement stationnaire et non anticipative du modèle *ARMA(P, Q) - GARCH(p, q)*

$$\begin{cases} X_t - c_0 = \sum_{i=1}^p a_{0i}(X_{t-i} - c_0) + e_t - \sum_{j=1}^Q b_{0j}e_{t-j}, \\ e_t = \sqrt{h_t}\eta_t \\ h_t = \omega_0 + \sum_{i=1}^q \alpha_{0i}e_{t-i}^2 + \sum_{j=1}^p \beta_{0j}h_{t-j}, \end{cases} \quad (3.37)$$

où (η_t) et les coefficients ω_0, α_{0i} et β_{0j} sont définis comme en (3.16). Les ordres P, Q, p, q sont supposée connus. Le vecteur des paramètres est noté

$$\varphi = (\vartheta', \theta')' = (c, a_1, \dots, a_P, b_1, \dots, b_Q, \theta')',$$

où θ est défini comme précédemment (voir (3.17)). L'espace des paramètres est

$$\Phi \subset \mathbb{R}^{P+Q+1} \times]0, +\infty[\times [0, \infty[^{p+q}.$$

La vraie valeur du paramètre est notée

$$\varphi_0 = (\vartheta'_0, \theta'_0)' = (c_0, a_{01}, \dots, a_{0P}, b_{01}, \dots, b_{0Q}, \theta'_0)'$$

Nous employons toujours la quasi-vraisemblance gaussienne conditionnelle aux valeurs initiales.

Si $q \geq Q$, les valeurs sont

$$X_0, \dots, X_{1-(q-Q)-P}, \tilde{\epsilon}_{-q+Q}, \dots, \tilde{\epsilon}_{1-q}, \tilde{\sigma}_0^2, \dots, \tilde{\sigma}_{1-p}^2.$$

Ces valeurs (dont les p dernières sont positives) peuvent dépendre des paramètres et des observations. Ceci permet, pour tout ϑ , de calculer les valeurs de $\tilde{\epsilon}(\vartheta)$, pour $t = -q + Q + 1, \dots, n$, puis, pour toute valeur de θ , la valeurs de $\tilde{\sigma}_t^2(\varphi)$, pour $t = 1, \dots, n$, peuvent donc être calculé à partir de

$$\begin{cases} \tilde{\epsilon}_t = \tilde{\epsilon}_t(\vartheta) = X_t - c \sum_{i=1}^P a_i (X_{t-i} - c) + \sum_{j=1}^Q b_j \tilde{\epsilon}_{t-j} \\ \tilde{\sigma}_t^2 = \tilde{\sigma}_t^2(\varphi) = \omega + \sum_{i=1}^q \alpha_i \tilde{\epsilon}_{t-i} + \sum_{j=1}^p \beta_j \tilde{\sigma}_{t-j}^2. \end{cases} \quad (3.38)$$

Lorsque $q < Q$, les valeurs initiales fixes sont

$$X_0, \dots, X_{1-(q-Q)-P}, \epsilon_0, \dots, \epsilon_{1-Q}, \tilde{\sigma}_0^2, \dots, \tilde{\sigma}_{1-p}^2.$$

Conditionnellement à ces valeurs initiales, la log-vraisemblance gaussienne est donnée par

$$\tilde{\mathbf{l}}_n(\varphi) = n^{-1} \sum_{t=1}^n \tilde{\ell}_t, \quad \tilde{\ell}_t = \tilde{\ell}_t(\varphi) = \frac{\tilde{\epsilon}_t^2(\vartheta)}{\tilde{\sigma}_t^2(\varphi)} + \log \tilde{\sigma}_t^2(\varphi).$$

Un estimateur QMV est défini comme une solution mesurable de l'équation

$$\hat{\varphi}_n = \arg \min_{\varphi \in \Theta} \tilde{\mathbf{l}}_n(\varphi). \quad (3.39)$$

Convergence forte

On considère les polynômes $a_\vartheta(z) = 1 - \sum_{i=1}^P a_i z^i$ et $b_\vartheta(z) = 1 - \sum_{j=1}^Q b_j z^j$. Nous procédons des hypothèses standard sur ces polynômes AR et MA , et adaptons l'hypothèse **A1** de la façon suivante :

A7 : $\varphi_0 \in \Phi$, où Φ est compact.

A8 : Pour tout $\varphi \in \Phi$, $a_\vartheta(z)b_\vartheta(z) = 0$ implique $|z| > 1$.

A9 : $a_{\vartheta_0}(z)$ et $b_{\vartheta_0}(z)$ n'ont pas de racines communes, $a_{0P} \neq 0$ ou $b_{0Q} \neq 0$.

Sous les hypothèses **A2** et **A8**, (X_t) est supposé comme l'unique solution strictement stationnaire et non anticipative de (3.37). Soit $\epsilon_t = \epsilon_t(\vartheta) = a_\vartheta(B)b_\vartheta^{-1}(B)(X_t - c)$ et $\ell_t = \ell_t(\varphi) = \epsilon_t^2/\sigma_t^2 + \log \sigma_t^2$, où $\sigma_t^2 = \sigma_t^2(\varphi)$ est la solution strictement stationnaire non anticipative et ergodique de (3.25). Notons que $e_t = \epsilon_t(\vartheta_0)$ et $h_t = \sigma_t^2(\varphi_0)$. Le résultat suivant étend le théorème 3.2.1.

Théorème 3.2.4 (*Convergence de l'estimateur QMV*)

Soit $(\hat{\varphi}_n)$ une suite d'estimateurs QMV et vérifiant (3.39). Supposons que $\mathbb{E}\eta_t = 0$. Alors, sous les hypothèses **A2-A4** et **A7-A9**, presque sûrement

$$\hat{\varphi}_n \longrightarrow \varphi_0, \quad \text{quand } n \longrightarrow \infty.$$

Remarque 3.2.3 .

1. Comme dans le cas *GARCH* pur, le théorème n'impose pas de variance finie pour e_t (et donc pour X_t). Dans le cas *ARMA* pur, où $e_t = \eta_t$ admet une variance finie, ce théorème réduit à un résultat classique concernant les modèles *ARMA* avec erreur *iid* (voir [11], p. 384).
2. En dehors de la condition $\mathbb{E}\eta_t = 0$, les conditions requises pour la convergence forte de l'estimateur QMV ne sont pas plus fortes que dans le cas *GARCH* pur.

Normalité asymptotique sous l'hypothèse d'existence du moment d'ordre 4

Pour tout ce qui précède, les résultats asymptotiques de l'estimateur de QMV (convergence et normalité asymptotique dans le cas *GARCH* pur, convergence dans le cas *ARMA – GARCH*) n'ont nécessité aucune hypothèse de moments sur le processus observé (un moment d'ordre 4 du processus *iid*, pas pour ϵ_t , est supposé pour la normalité asymptotique dans le cas *GARCH* pur). On pourrait penser qu'il en sera de même pour établir la normalité asymptotique dans le cas *ARMA – GARCH*. L'exemple suivant montre que ce n'est pas le cas.

Exemple 1 (*Non-existence de J sans hypothèse de moment*)

Considérons le modèle $AR(1) - ARCH(1)$

$$X_t = a_{01}X_{t-1} + e_t, \quad e_t = \sqrt{h_t}\eta_t, \quad h_t = \omega_0 + \alpha_0 e_{t-1}^2 \quad (3.40)$$

où $|a_{01}| < 1$, $\omega_0 > 0$, $\alpha_0 \geq 0$, et la distribution de la suite *iid* (η_t) est définie, pour $a > 1$, par

$$\mathbb{P}(\eta_t = a) = \mathbb{P}(\eta_t = -a) = \frac{1}{2a^2}, \quad \mathbb{P}(\eta_t = 0) = 1 - \frac{1}{a^2}.$$

Alors le processus (X_t) est toujours stationnaire, pour toute valeur de α_0 (car $\exp\{-\mathbb{E}(\log \eta_t^2)\} = +\infty > \alpha_0$: condition de contrainte de stationnarité; par contre, X_t n'admet pas de moment d'ordre 2 quand $\alpha_0 \geq 1$).

La première composante du vecteur des scores (normalisé) est

$$\begin{aligned} \frac{\partial \ell_t(\theta)}{\partial a_1} &= \left(1 - \frac{e_t^2}{\sigma_t^2}\right) \left(\frac{1}{h_t} \frac{\partial \sigma_t^2(\theta_0)}{\partial a_1}\right) + \frac{2e_t}{h_t} \frac{\partial \epsilon_t(\theta_0)}{\partial a_1} \\ &= -2\alpha_0 (1 - \eta_t^2) \left(\frac{e_{t-1}X_{t-2}}{h_t}\right) - 2\frac{\eta_t X_{t-1}}{\sqrt{h_t}}. \end{aligned}$$

Nous avons

$$\begin{aligned} &\mathbb{E} \left\{ \alpha_0 (1 - \eta_t^2) \left(\frac{e_{t-1}X_{t-2}}{h_t}\right) + \frac{\eta_t X_{t-1}}{\sqrt{h_t}} \right\}^2 \\ &\geq \mathbb{E} \left[\left\{ \alpha_0 (1 - \eta_t^2) \left(\frac{e_{t-1}X_{t-2}}{h_t}\right) + \frac{\eta_t X_{t-1}}{\sqrt{h_t}} \right\}^2 \middle| 1 - \eta_t = 0 \right] \mathbb{P}(\eta_{t-1} = 0) \\ &= \frac{\alpha_{01}^2}{\omega_0} \left(1 - \frac{1}{\alpha^2}\right) \mathbb{E}(X_{t-2}^2) \end{aligned}$$

puisque, $\eta_{t-1} = 0$ entraîne $\epsilon_{t-1} = 0$ et $X_{t-1} = a_{01}X_{t-2}$, et aussi, η_{t-1} et X_{t-2} sont indépendants.

Par conséquent, si $\mathbb{E}X_t^2 = \infty$ et $a_{01} \neq 0$, le vecteur des scores n'admet pas de variance.

Cet exemple montre qu'il n'est pas possible d'étendre le résultat de normalité asymptotique obtenu dans le cas *GARCH* aux modèles *ARMA - GARCH* sans hypothèses de moment supplémentaires. Ceci est logique car la normalité asymptotique de l'estimateur du *QMV* pour les modèles *ARMA* (peuvent être vus comme limites des modèles *ARMA - GARCH* lorsque les coefficients α_{0i} et β_{0j} tendent vers 0) est montrée avec des hypothèses de moment de second ordre. Pour un *ARMA* avec innovation ou bruit à variance infinie, les convergences des estimateurs peuvent être plus rapide que dans le cas standard et les lois asymptotique sont stable, mais non gaussienne en général.

Nous montrons la normalité asymptotique avec une hypothèse de moment d'ordre 4. Rappelons que, cette hypothèse équivaut à $\rho\{\mathbb{E}(A_{0t} \otimes A_{0t})\} < 1$ d'après le théorème 2.9 (voir [24]). Nous introduisons les hypothèses suivantes :

A10 : $\rho\{\mathbb{E}(A_{0t} \otimes A_{0t})\} < 1$ et, pour tout $\theta \in \Theta$, $\sum_{j=1}^p \beta_j < 1$.

A11 : $\varphi_0 \in \mathring{\Theta}$, où $\mathring{\Theta}$ désigne l'intérieur de Θ .

A12 : Il n'existe pas d'ensemble Λ de cardinal

2 tel que $\mathbb{P}(\eta_t \in \Lambda) = 1$.

L'hypothèse **A10** implique que $\kappa_\eta = \mathbb{E}(\eta_t^4) < \infty$ et rend l'hypothèse **A2** redondante. L'hypothèse d'identifiabilité **A12** est légèrement plus forte que la première partie de l'hypothèse **A3** lorsque la distribution de η_t n'est pas symétrique. Le théorème suivant énonce le résultat assurant la normalité asymptotique de l'estimateur du QMV d'un modèle *ARMA – GARCH*.

Théorème 3.2.5 (*Normalité asymptotique du QMV*)

*Supposons que $\mathbb{E}\eta_t = 0$ et que les hypothèses **A3–A4** et **A8–A12** sont vérifiées. Alors*

$$\sqrt{n}(\hat{\varphi}_n - \varphi_0) \xrightarrow{\mathcal{L}} \mathcal{N}\{0, \Sigma\},$$

où $\Sigma = \mathcal{J}^{-1}\mathcal{I}\mathcal{J}^{-1}$,

$$\mathcal{I} = \mathbb{E}_{\varphi_0} \left(\begin{array}{cc} \frac{\partial \ell_t(\varphi_0)}{\partial \varphi} & \frac{\partial \ell_t(\varphi_0)}{\partial \varphi'} \\ \frac{\partial \ell_t(\varphi_0)}{\partial \varphi'} & \frac{\partial \ell_t(\varphi_0)}{\partial \varphi'} \end{array} \right), \quad \mathcal{J} = \mathbb{E}_{\varphi_0} \left(\frac{\partial^2 \ell_t(\varphi_0)}{\partial \varphi \partial \varphi'} \right).$$

Si, de plus, la distribution de η_t est symétrique, on a

$$\mathcal{I} = \begin{pmatrix} I_1 & 0 \\ 0 & I_2 \end{pmatrix}, \quad \mathcal{J} = \begin{pmatrix} J_1 & 0 \\ 0 & J_2 \end{pmatrix},$$

avec

$$\begin{aligned}
 I_1 &= (\kappa_\eta - 1)\mathbb{E}_{\varphi_0} \left(\frac{1}{\sigma_t^4} \frac{\partial \sigma_t^2}{\partial \vartheta} \frac{\partial \sigma_t^2}{\partial \vartheta'}(\varphi_0) \right) + 4\mathbb{E}_{\varphi_0} \left(\frac{1}{\sigma_t^2} \frac{\partial \epsilon_t}{\partial \vartheta} \frac{\partial \epsilon_t}{\partial \vartheta'}(\varphi_0) \right) \\
 I_2 &= (\kappa_\eta - 1)\mathbb{E}_{\varphi_0} \left(\frac{1}{\sigma_t^4} \frac{\partial \sigma_t^2}{\partial \theta} \frac{\partial \sigma_t^2}{\partial \theta'}(\varphi_0) \right), \\
 J_1 &= \mathbb{E}_{\varphi_0} \left(\frac{1}{\sigma_t^4} \frac{\partial \sigma_t^2}{\partial \vartheta} \frac{\partial \sigma_t^2}{\partial \vartheta'}(\varphi_0) \right) + \mathbb{E}_{\varphi_0} \left(\frac{1}{\sigma_t^2} \frac{\partial \epsilon_t}{\partial \vartheta} \frac{\partial \epsilon_t}{\partial \vartheta'}(\varphi_0) \right), \\
 J_2 &= \mathbb{E}_{\varphi_0} \left(\frac{1}{\sigma_t^4} \frac{\partial \sigma_t^2}{\partial \theta} \frac{\partial \sigma_t^2}{\partial \theta'}(\varphi_0) \right).
 \end{aligned}$$

Remarque 3.2.4 .

1. Il est intéressant de noter que si η_t a une loi symétrique, alors la variance asymptotique Σ est bloc-diagonale, qui est interprétée comme une indépendance asymptotique entre les estimateurs des coefficients *ARMA* et celles des coefficients *GARCH*. La distribution asymptotique des estimateurs des coefficients *ARMA* dépend cependant des coefficients *GARCH* (vu la forme des matrices I_1 et J_1 faisant intervenir les dérivées de σ_t^2). En d'autre part, toujours lorsque la distribution de η_t est symétrique, la précision asymptotique de l'estimation des paramètres *GARCH* n'est pas affectée par la partie *ARMA* : le bloc en bas à gauche $J_2^{-1}I_2J_2^{-1}$ de Σ ne dépend que des coefficients *GARCH*. La forme bloc-diagonale peut également être intéressante pour tester des problèmes d'hypothèses jointes sur les paramètres *ARMA* et *GARCH*.

2. L'hypothèse **A11** impose la stricte positivité des coefficients *GARCH* et il est évident de voir que cette hypothèse ne contraint que les coefficients *GARCH*. Pour toute valeur de ϑ_0 , la restriction de Φ à ses $P + Q + 1$ premières coordonnées peut être choisie suffisamment grande pour que son intérieur contient ϑ_0 et l'hypothèse **A8** soit satisfaite.

3. Dans la preuve du théorème, la symétrie de la distribution du processus *iid* est utilisée pour montrer le résultat suivant, qui est d'un intérêt indépendant.

Si la distribution de η_t est symétrique alors,

$$\forall j, \quad \mathbb{E} \{ g(\epsilon_t^2, \epsilon_{t-1}^2, \dots) \epsilon_{t-j} f(\epsilon_{t-j-1}, \epsilon_{t-j-2}, \dots) \} = 0, \tag{3.41}$$

à condition que cette attente existe.

CHAPITRE

4

Measuring the impact of covid-19 pandemic on the oil prices volatility by GARCH modeling

ABSTRACT. The aim of this work is the study of crude oil price volatility, particularly during the Covid-19 pandemic, which changed the world scene and inflicted serious crises not only in the global health system, but also in the international financial markets and economy. This economic situation has many policy makers, financiers and portfolio managers worried about avoiding the risk of potential damage. Hedging strategies are based on the correct estimation of price volatility. For this aim, we use the *GARCH* model to measure the impact of volatility and shocks. The oil price series that will be examined here is *WTI* (West Texas Intermediate). In this type of series, such as financial or monetary variables which are characterized by high volatility, it is obvious that the analysis will be based on models which are qualified in such complex studies.

More precisely, the model used for predicting volatility associated with the price variable is the *GARCH*(1, 2) model, the best model among *GARCH*(1, 1), *GARCH*(1, 2) and *TGARCH*(1, 2), which are matched to *WTI* oil price data in through this analysis. Although the *EGARCH*

and *TGARCH* models are better because they take into account the asymmetry property of volatility, the so-called leverage effect to Black F. [1], the *GARCH*(1, 2) model was adopted as an appropriate model for the crude oil price data because it presents lower values of the forecasting criteria compared to the two other models. The forecast is made for the last three months of 2021. The result concludes that the predicted values and the current values are very close, and that the model *GARCH*(1, 2) is the best forecast model.

Key Words. Covid-19 pandemic, Oil price, Asymmetric Volatility, *GARCH* model, *TGARCH* model, *EGARCH* model, Forecasting.

Mathematical Subject Classification. 62M10, 60G25, 62P10, 62P20.

4.1 Introduction

The global financial crises and their impacts on international markets have attracted great attention from academic researchers, investors and policy makers. The first decade of this millennium saw at least two crises, such as the dotcom bubble of 2002 and the subprime financial crisis of 2008. This decade too was not distinct from the previous one, the COVID-19 pandemic had hit the world hard and caused serious crises, and which still surprises the whole world until now because of its significant effects, Asset price fluctuations, in world markets in particular, are one of them. To reminding, the outbreak of COVID–19, which was started in the Wuhan city of China during November 2019 has been declared as a global pandemic by the World Health Organisation (WHO) on March 11, 2020. The outbreak of COVID–19 has shaken the global financial markets, commodity markets, economic activities, employment and GDP of the countries, until they have their fully magnitudes as indicated by Maher B.K. and al. [16].

The reason for taking the crude oil because it's a vital commodity and has significance in both financial markets and economic growth.

The price of crude oil fell from 9 \$ per barrel in December 1998 to 145 \$ in July 2008. It then fell to 32 \$ in December 2008, before rising again in 2009 and reaching the end of the year a level of 80 dollars (79.36 \$ in 31 December 2009). To know other fluctuations at the end of 2014, falling to the threshold of 50 \$ then to 30 dollard in mid 2016,

And it continued to fall until reaching a collapse that the history of American oil prices has not known at all, on Monday (but not a black) April 20, 2020, where the price per barrel reached (-37.63), or 306%, for deliveries in May, i.e., the sellers are now offering to pay the buyers for this contract. Tuesday, April 21 in the morning, just a day later, the price of *WTI* for delivery in May rebounded on the Asian markets to trade around 1.1 dollar. For its part, the barrel for delivery in June rose above 21 dollars. Of course, this happened because of the Covid-19 pandemic which was at its peak during this time.

But after signs of recovery emerged from the Covid-19 pandemic, even signs of recovery in U.S. crude prices where it returned to a wild rise which exceeds 90 dollard in recent months, especially

with the recent strained geopolitical relations between the major powers. This conjunction of an upward trend and high volatility is likely to continue in the coming years. Of course, experience has shown how difficult it is to predict the evolution of the price of oil.

In fact, considerable interest in the literature has gone into examining the volatility of oil prices. Thus, in the same interest, we wanted to have a contribution in this study.

We will treat this study in its natural framework, which is the time series. The price data series is denoted *WTI*, It is the same universally recognized name (West Texas Intermediate : one of a type of crude oils used as a standard in pricing crude oil and as a commodity for oil futures contracts on the New York Mercantile Exchange (raw materials stock Exchange). The data sample frequency is daily oil closing prices and spans from January 01, 2015 to December 31, 2021 excluding holidays (1826 observations). The data can be downloaded from the website : investing.com.

GARCH models are the main tool in this analysis because they are a type of economic and monetary series that are characterized by irregular fluctuations over time.

The objective of this work is to seek, in the family of *GARCH* models, the adequate model and generator of *WTI* data in order to perform, mainly, the forecast, an important step in the methodology of Box-Jenkins. It should be noted that to study the behavior of the *WTI* series, we will also need the yield series, which represent the residuals $(\varepsilon_t)_{t \in \mathbb{Z}}$ which help us to check whether the generating process of *WTI* is an *ARIMA* or it is a model of the *GARCH* family. Generally, *WTI* yields or residuals are often expressed as

$$\varepsilon_t = \ln \frac{WTI_t}{WTI_{t-1}} = \ln WTI_t - \ln WTI_{t-1},$$

such that, \ln is the natural logarithm, WTI_t and WTI_{t-1} are the oil prices at date t and $t - 1$, respectively.

The evolution of the financial markets has certain empirical characteristics, which any model seeks to reproduce. The calculation of the risks taken by the players must, in order to be as predictive as possible, take account of these characteristics, it is for example the aggregation of volatility and its asymmetry. A popular class for the among practitioners in the finance area is the *GARCH* model proposed by Tim Bollerslev [2], which is a generalization of the *ARCH*

model proposed by Robert Engle [11] who quickly became one of the pioneers of financial econometrics. The model principle allows to consider an essential characteristic of the markets : volatility is not constant over time. In principle, taking this phenomenon into account increases the predictive potential of the models.

The remainder of the paper is organized as follows : Section 2 we will dedicate it to the review of the literature, section 3 is devoted to a brief theoretical presentation of *GARCH* models and its extensions, while section 4 presents data and the empirical results. Section 5 concludes this paper.

4.2 Literature review

Several GARCH extensions have been introduced in the finance and economic literature to improve some aspects of the GARCH model so that the models are more flexible and adequate in accommodating some characteristics and dynamics of a time series, for example, see Bollerslev T. and al. [5] and Higgins and Bera [28]. Ng and McAleer [38] used simple *GARCH*(1, 1) and *TARCH*(1, 1) models for testing estimation and forecasting the volatility of daily returns in S&P 500 Composite Index and the Nikkei 225 Index. They concluded that *TARCH*(1, 1) was the best performing model with S&P 500 data, whereas the *GARCH*(1, 1) model was better in some cases with Nikkei 225. Ramzan and al. [45] modeled exchange rate dynamics in Pakistan, using the *GARCH* family models, on the monthly data from July 1981 to May 2010. The study results showed that *GARCH*(1, 2) was better than *EGARCH*(1, 2) model. However, the *GARCH*(1, 2) model was used to remove the persistence in volatility while *EGARCH*(1, 2) successfully overcame the leverage effect in the exchange rate returns. Moreover they concluded that the *GARCH* family of models captures the volatility and leverage effect in the exchange rate returns, giving fairly good forecasting performance for the model.

Also, there are many research works have been done in the area of the commodities market, natural gas, crude oil and price volatility. Some of the important works are mentioned here. A paper investigates the behaviour of crude oil and natural gas price volatility in the United States since 1990 using *GARCH* model is by Pindyck [44]. Again, a study concerned with modelling

of price volatility of crude oil market in Nigeria employing both symmetric and asymmetric models of *GARCH* family appeared in Dum and Essi [18].

There was a study investigating the causal association between the stock market returns and crude oil price anomalies in the Indian stock market covering 10 companies of oil drilling and exploration sectors listed in the CNX NIFTY indexes and BSE Sensex from 2009 to 2018 (see Hawaldar and al [27]). Furthermore, the study focussed on comparing between realized *GARCH* model with some conventional *GARCH* models such as *GARCH*, *GJR – GARCH* and *EGARCH* by using the Gold 5 min intra-day data from April 2012 to April 2018. Nugroho and al. [39] used the $GARCH(1, 1)$ for modeling of the volatility of returns,. This study proposed two new classes of $GARCH(1, 1)$ model by applying the Tukey transformations to the returns and to the lagged variance and it recommends the use of Excel Solver for finance academics and practitioners working on volatility using $GARCH(1, 1)$ models. The empirical findings conclude that $GARCH(1, 1)$ models under Tukey transformations are more appropriate than standard for describing returns and volatility of financial time series and its stylized facts including fat tails and mean reverting.

In the paper of Merabet F. and al. [33], the behavior of the oil price series is examined. The study is based on a combination of the Box-Jenkins methodology with the *GARCH* processes. Of the models identified this analysis, the model $ARIMA(3, 1, 1) – EGARCH(1, 2)$ is retained and is the best forecast model.

In Mia M.S. and Rahman M.S. [34] authors have built a model to forecast the exchange rate of Bangladesh. A study on Monthly average exchange rates of Bangladesh for the period from August, 2004 to April, 2019. They have selected $ARIMA(1, 1, 1)$ as a main model for this study. Then they tried to model the volatility of exchange rate using *ARCH*, *GARCH*, *EGARCH*, *IGARCH* and *TARCH* models. $ARIMA(1, 1, 1) – GARCH(1, 1)$ is selected as a best model compared to others since it has the lower values of RMSE, MAE, MAPE and Theil than other models. However, to the authors' best knowledge, very few publications can be found on volatility of the oil price by symmetric and asymmetric models that capture most common stylized facts about oil price such as volatility clustering and leverage effect.

Not to mention the family of multivariate *GARCH* models which are in this regard very crucial

due to their efficiency and diversity, we will certainly discuss them in detail in the next articles, where we intend to use the multivariate model of $GARCH - BEKK$ as studied by El Ghini and Saidi [19], where they had studied the extent of the transmission of the 2008 financial crisis from America to other countries of the world. Olstad A. and al. [40] investigates the time-varying correlation between the volatilities of two oil benchmarks (Brent and WTI) and six currencies of the major oilimporters and oil-exporters, for the period from February 1, 1999 to May 30, 2016, using a multivariate $Diag - BEKK - GARCH$ model. The analysis reveals that oil and currency volatilities exhibit positive correlations during major global or region-specific economic events (such as the Global Financial Crisis of 2007–2009 and the EU debt crisis period). Yaya and al.[50] use the $CCC - MGARCH$ model to study the return transmission between oil and gold markets before and after the global financial crisis. They find a bidirectional return spillover before the crisis, and unidirectional spillover from gold to the oil market after the crisis. Guesmi and al.[26] apply the $VARMA - DCC - GJR - GARCH$ model and provide evidence of significant volatility spill-over between Bitcoin and other financial markets, including the oil market. Moreover, Bitcoin provides hedging and diversification benefits against oil, gold, and emerging stock markets.

The shocks affect, caused by the current Ukrainian crisis, in the world economy, in particular on the ambivalent fluctuations of the indices of the world stock exchanges and the prices of strategic raw materials, will mark many publications and scientific works using the technique of multivariate $GARCH$ models.

There are a lot of papers who made a studes of impact of Covid-19 on both oil prices volatility and auther stocks volatility prices using the $GARCH$ models family. Endri E. and al. [20] has examined the response of stock prices on the Indonesia Stock Exchange (IDX) to COVID-19 using an event study approach and the $GARCH$ model. While Meher B. K. and al. [32] have focused on measuring the impact and leverage of corona virus disease 2019 on price volatility of crude oil and natural gas quoted on MCX, India, using the $EGARCH$ model with general error distribution.

4.3 Research methodologie

4.3.1 Non-Linear Models and heteroskedasticity

Financial time series returns are notoriously unstable and characterized by heteroscedasticity and in particular asymmetric, the constraints that $ARMA(p, q)$ models suffer from taking them into account, and sometimes involve the use of nonlinear models that may make inadequate the $ARMA$ specification Eq. (3.1).

$$r_t = \mu_t + \varepsilon_t \quad (3.1)$$

where r_t is the series of price returns which is written as the sum of μ_t : the conditional expectation of returns (which is modeled by $ARMA$) and the residuals ε_t (which are assumed to be a white noise in the $ARMA$ specification).

When ε_t is not white noise in the (3.1), ε_t , an Autoregressive Conditionally Heteroscedastic ($ARCH$) models ([11]) are introduced by integrating volatility into the $ARMA$ specification (3.1).

Dans cette sous section on citera les deux types des modèles non linéaire à savoir les modèles coventionnels et les modèles asymetriques

I. Conventional models (Symetric)

i) ARCH Models

In 1982 Robert Engle developed the Autoregressive Conditional Heteroskedasticity ($ARCH$) models to model the time-varying volatility often observed in economical time series data that $ARMA$ linear models are no longer able to take them into account. For this contribution, he won the 2003 Nobel Prize in Economics with the co-integration of Cliv Granger.

$ARCH$ models assume the variance of the current error term or innovation to be a function of actual sizes of the previous time periods' error terms : often the variance is related to the squares of the previous innovations, i. e., the conditional varaince to the information available

at time $t - 1$ varies over time, and that is given as follows :

$$h_t = \mathbb{V}ar(\varepsilon_t | I_{t-1}) = w + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2,$$

where

$$\mathbb{E}(\varepsilon_t | I_{t-1}) = 0, \quad w > 0, \quad \alpha_i \geq 0 \text{ for } i = 1, \dots, q,$$

and

$$I_{t-1} = \{\varepsilon_{t-1}, \varepsilon_{t-2}, \dots, \varepsilon_{t-q}\}$$

is the information set available at time $t - 1$.

The condition of stationnarity of the *ARCH* (q) is

$$\sum_{i=1}^q \alpha_i < 1.$$

ii) Generalized ARCH models

The model is generalized by Robert Engle's doctoral student, Bollerslev Tim in 1986. The conditional variance not only depends on the previous innovations but also on its previous conditional variances, and that is given as follows :

$$h_t = \mathbb{V}ar(\varepsilon_t | I_{t-1}) = w + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j h_{t-j}, \quad (3.2)$$

such that the following conditions are necessary :

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

$$w > 0, \quad \alpha_i \geq 0 \text{ and } \beta_j \geq 0 \text{ for } i = 1, \dots, q \text{ and } j = 1, \dots, p,$$

where I_{t-1} is the information set available at time $t - 1$.

The stationarity condition of the model is as follows :

$$\sum_{i=1}^q \alpha_i + \sum_{j=1}^p \beta_j < 1. \quad (3.3)$$

The term $\sum_{i=1}^q \alpha_i + \sum_{j=1}^p \beta_j$, is a measure of the degree of volatility persistence.

The $GARCH(1, 1)$ model is presented in the form

$$h_t = \text{Var}(\varepsilon_t | I_{t-1}) = w + \alpha \varepsilon_{t-1}^2 + \beta h_{t-1},$$

with the same conditions of general case.

II. Asymmetrical models

The $ARCH$ and $GARCH$ models are based on a hypothesis of symmetry of the impact of shocks on volatility. They assign the same weight to positive shocks (good news) as to negative shocks (bad news), which is contrary to empirical reality. The asymmetry phenomena, can easily be observed through the graphs of the raw series and the volatility together : when the prices of a financial asset are in a bullish phase, low volatility is observed, on the other hand when they are in a bearish phase, high volatility is observed. Therefore, extensions of the $GARCH$ models that take into account the asymmetric behavior of volatility have been introduced :

- $EGARCH$ (Nelson [19]),
- $TGARCH$ or $GJR - GARCH$ (Glosten and al. [12])
- $TARCH$ (Zakoian [26])

1. EGARCH model

The standard $GARCH$ model is unable to capture the skewness or asymmetric nature caused by the negative correlation between returns and volatility which is referred to as the leverage effect. Therefore, if the conditions allow, the $EGARCH$ model , is the best framework we will rely on in our analysis to measure the leverage effect of COVID-19 on the price volatility of the crude oil. That is represented by

$$\ln(h_t) = w + \sum_{i=1}^q [\alpha_i |z_{t-i}| + \gamma_i z_{t-i}] + \sum_{j=1}^p \beta_j \ln(h_{t-j}), \quad (3.4)$$

while, h_t the actual volatility, $z_t = \frac{\varepsilon_t}{\sqrt{h_t}} \sim i.i.d. \mathcal{N}(0, 1)$ (standardized noise white) who ε_t the residuals of the estimation of a process $ARMA(p, q)$, $\ln(h_t) = \ln$ of variance or \ln returns, w a constant, α_i the $ARCH$ effects, γ_i is asymmetric effects and β_j is a $GARCH$ effects.

If $\gamma_i = 0$ the model is the standard *GARCH*.

If $\gamma_i < 0$ it implies that bad news (negative shocks) generate large volatility that good news (positive shocks).

The *EGARCH*(1, 1) is given as bellow,

$$\ln(h_t) = w + \alpha |z_{t-i}| + \gamma z_{t-i} + \beta \ln(h_{t-j}),$$

where $z_t = \frac{\varepsilon_t}{\sqrt{h_t}}$.

2. Threshold-GARCH model

Threshold – GARCH (TGARCH) model one of the ways taking into account the asymmetry of volatility which is a reality of financial series. To do that it simply adds into the variance equation a multiplicative dummy variable to check whether there is statistically significant difference when shocks are negative.

The model was introduced by Glosten, Jagannathan and Runkle **in 1993**.

The conditional variance for a TGARCH(1,1) is represented by :

$$h_t = \sigma_t^2 = w + \alpha \varepsilon_{t-1}^2 + \gamma I_{t-1} \varepsilon_{t-1}^2 + \beta h_{t-j}$$

where I_{t-1} takes the 1 (bad news) for $\varepsilon_t < 0$, and 0 otherwise. So "good news" and "bad news" have a different impact on the volatility. Good news (positive shock) has an impact of α , while bad news (negative shock) has an impact of $\alpha + \gamma$. Such as $\alpha \geq 0$, $\beta \geq 0$, $\gamma \geq 0$ and $w > 0$.

γ is known as the asymmetry or leverage term. $\gamma > 0$ is asymmetry, while $\gamma = 0$ is symmetry (model collapses to the standard GARCH). If γ is a significant and positive, negative shocks will have larger effects on the volatility h_t than positive shocks.

TGARCH models can be extended to higher order specifications by including more lagged terms and noted TGARCH(p,q), as follows :

$$h_t = \sigma_t^2 = w + \sum_{i=1}^q (\alpha_i + \gamma_i I_{t-1}) \varepsilon_{t-1}^2 + \sum_{j=1}^p \beta_j h_{t-j}, \quad (3.5)$$

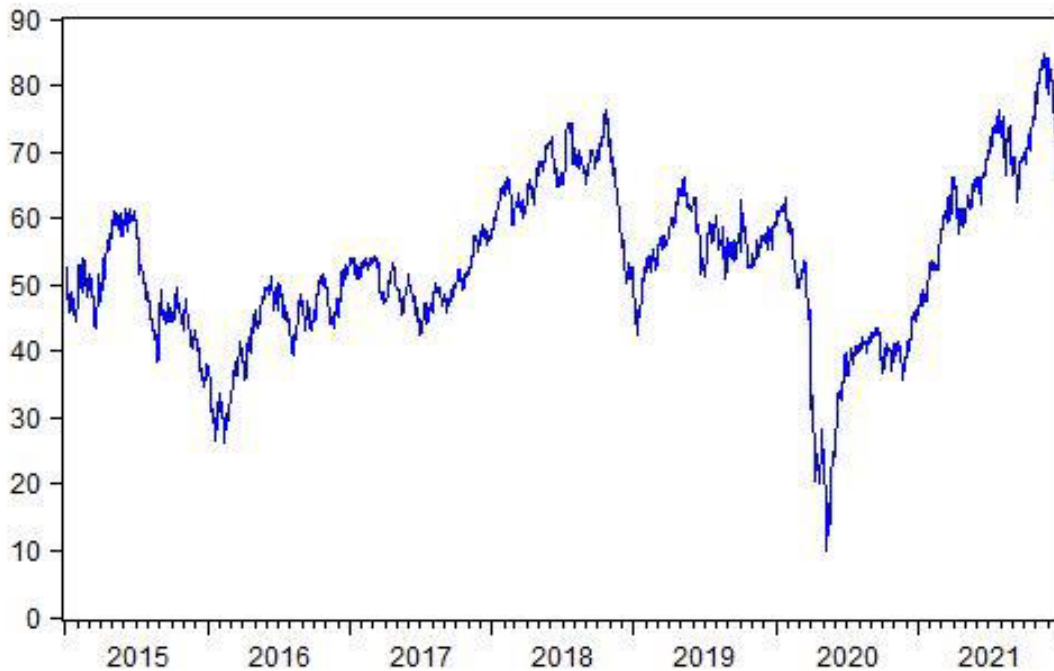
with specific conditions.

4.4 Analysis of results and discussion

The study is based on daily WTI crude oil prices covering the period from January 01, 2015 to December 31, 2021, the total of observations is 1827. The data was downloaded from the investing.com website. Our interest is to formulate models and forecasting price volatility of crude oil. E-Views 10 has been used.

4.4.1 Statistical analysis of WTI raw serie's behavior





























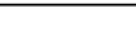
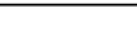
Figure 1 : Crude oil prices (WTI) "1/01/2015-12/31/2021"



Source : Athors' computation throught E-views 10

The *WTI* raw series in Figure 1 presents an unstable trend over time i. e, it is affected by an upward trend, which means that the series is not stationary in variance and in mean. the series shows a high variability throughout the period studied. This allows to say that a GARCH type process could be adapted to the modelling of the oil price series.

Table 1 : Correlogram of WTI((One significant peak

| Autocorrelation | Partial Correlation | AC | PAC | Q-Stat | Prob | |
|---|---|----|-------|--------|--------|-------|
|  |  | 1 | 0.994 | 0.994 | 1806.8 | 0.000 |
|  |  | 2 | 0.988 | 0.053 | 3594.2 | 0.000 |
|  |  | 3 | 0.983 | 0.025 | 5363.2 | 0.000 |
|  |  | 4 | 0.978 | 0.016 | 7114.8 | 0.000 |
|  |  | 5 | 0.972 | -0.017 | 8848.3 | 0.000 |
|  |  | 6 | 0.966 | -0.038 | 10562. | 0.000 |
|  |  | 7 | 0.961 | -0.009 | 12257. | 0.000 |
|  |  | 8 | 0.954 | -0.038 | 13930. | 0.000 |
|  |  | 9 | 0.947 | -0.096 | 15579. | 0.000 |
|  |  | 10 | 0.939 | -0.052 | 17202. | 0.000 |
|  |  | 11 | 0.932 | -0.008 | 18799. | 0.000 |
|  |  | 12 | 0.924 | 0.017 | 20372. | 0.000 |
|  |  | 13 | 0.917 | 0.018 | 21922. | 0.000 |
|  |  | 14 | 0.910 | -0.019 | 23447. | 0.000 |
|  |  | 15 | 0.902 | -0.023 | 24947. | 0.000 |

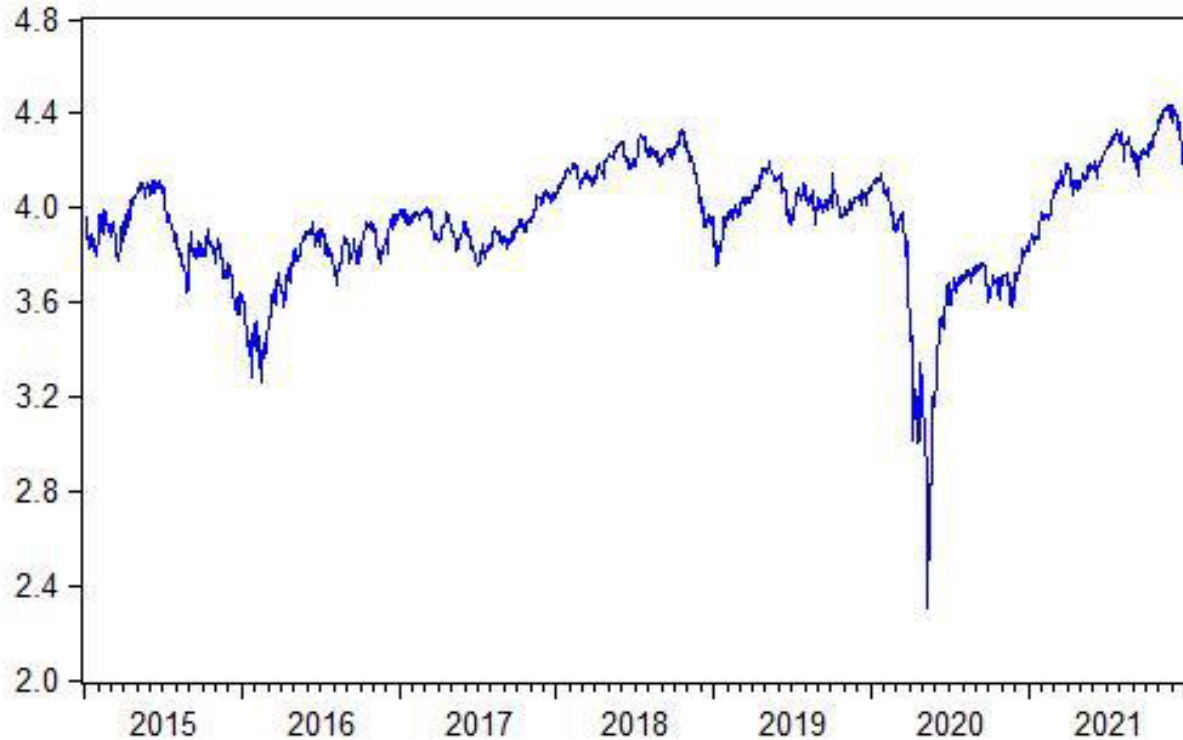
Source : Authors' computation through E-views 10

The preliminary analysis of stationnarity of the prices series, also give a sign of no-stationarity of the *WTI*, it would be much more in variance : the collerogram (Table 1) shows that, the all values of *Prob* of the all lags (36) are less than 0.05 and the slow decrease of the ACF terms. So then, the *WTI* raw series would follow an *AR*(1) but conventional process (*Cfr* correlogram).

4.4.2 Stationarity study

To formulate the appropriate model the logarithmic transformation of the prices oil data have been calculated. This has made the data of crude oil, less fluctuated but not stationary and keep the same shape and characteristic as presented in Figure 2 bellow.

Figure 2 : Logarithm of crude oil prices (LWTI)



Source : Athors' computation throught E-views 10

In first, we determine the type of non-stationarity by generating process of the raw time series, we perform the unit root test at level 5% using the Augmented-Dickey-Fuller test (*ADF*) (introduced by David A. Fuller and Wayne A. Fuller (see [16] and [17]) on the log oil prices (*WTI*) which based on hypotheses :

$$\begin{cases} H_0 : \text{has unit root} \\ H_1 : \text{hasn't unit root} \end{cases}$$

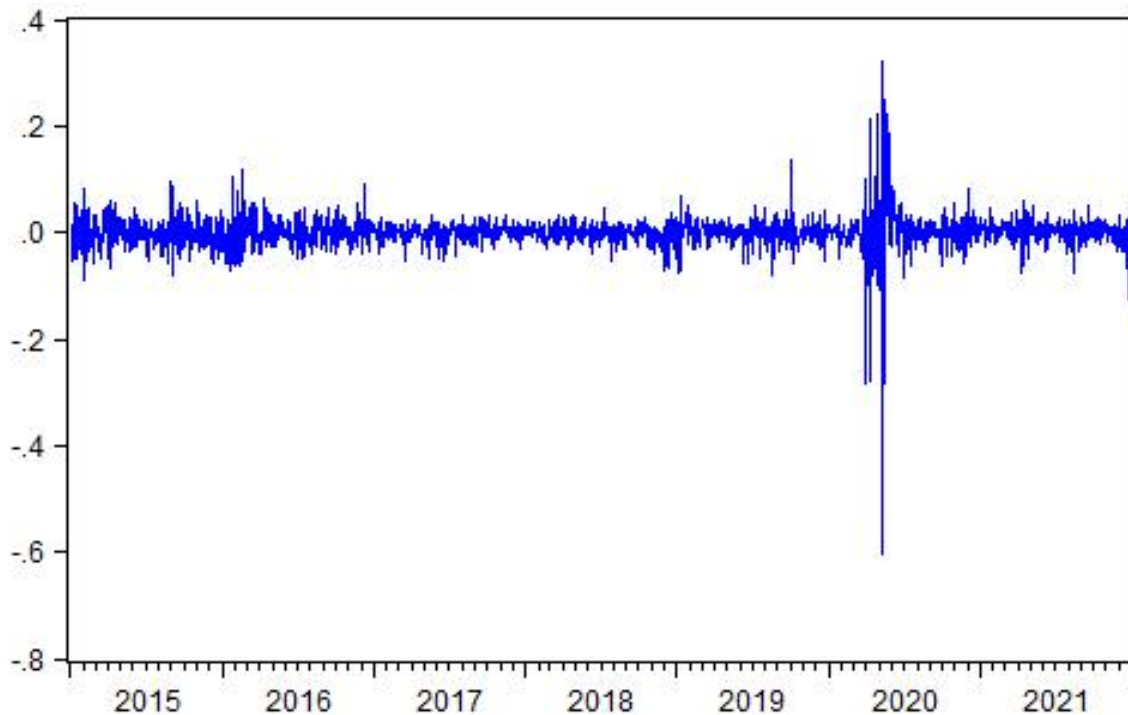
with the inclusion of the test equation as intercept, trend and intercept and none. This step is essential, wich allows to identify the nature of the trend, deterministic (*TS*) or stochastic (*DS*).

In effect, the analysis of unit root test, confirm that the *WTI* series has a unit root for all three models (Model 3 : $Prob^* = 0.4210 > 0.05$ and the trend and intercept coefficient is not significant ($Prob = 0.1843 > 0.05$) with $(1.3280_{t-stat} < 3.11_{ADF})$; Model 2 : $Prob^* = 0.2890 >$

0.05 and the intercept coefficient is significant ($Prob = 0.0430 < 0.05$) with ($2.0255_{t-stat} < 2.83_{ADF}$); Model 1 : $Prob^* = 0.6539 > 0.05$) then the serie of data is not stationnary and it is of the (DS) type, model two. The results are presented in Annex A1 section below.

Note that the method of stationarization of the data series will be done by differentiation ($d(LWTI)$) presented in Figure 3.

Figure 3 : Crude oil prices returns (DLWTI)



Source : Athors' computation throught E-views 10

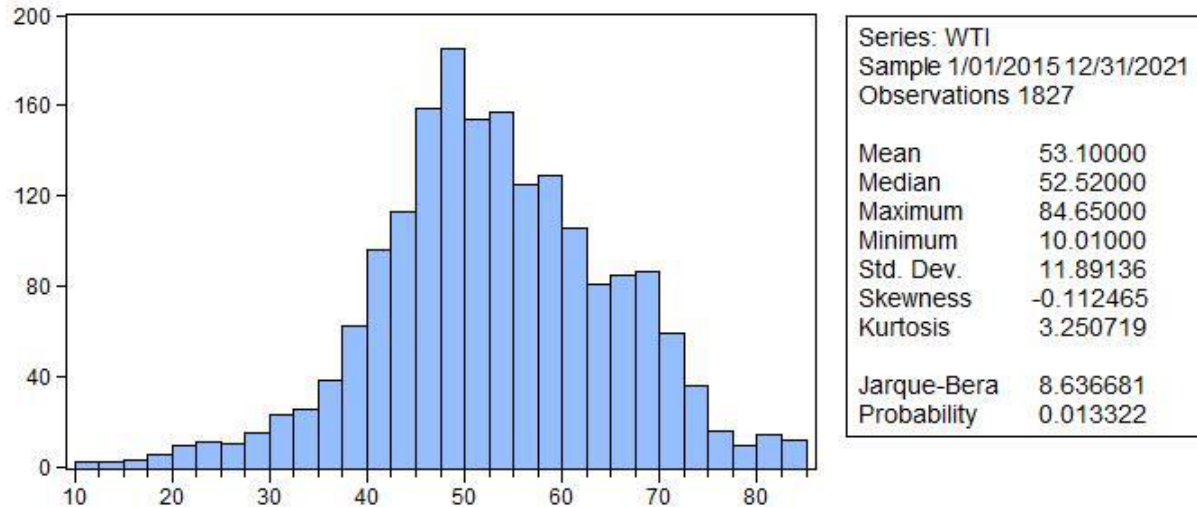
Figure 3 represents the volatility clustering of crude oil data from January 1, 2015 to December 31, 2022, i.m., small variations tracked by small variations and large variations tracked by large variations which implies that the volatility models can be formulated.

4.4.3 Statistical analysis of leverage effect and ARCH effect

To formulate the adequate *GARCH* model few more items need to be tested, i.e., asymmetry character, volatility clustering and presence of ARCH effects.

I. Leverage effect (asymmetry data)

Figure 4 : Asymmetry and peakedness character of crude oil data



Source : Authors' computation throught E-views 10

Statistical indicators of *WTI* series (Figure 4), in particular with the histogram and the test of Jarque-Bera, showed that there is a characteristic of asymmetry resulting from the effect of Covid-19 or leverage effect. Indeed, the Kurtosis coefficient (3.251) is lower than 3 and the skewness coefficient (-0.11) which is different from 0. So the data series is leptocurtic and Skewed, which means it is not normally distributed even *Jarque – Bera* Statistics (8.64) have also verified that the series data is leptokurtic and skewed, likewise, the associated probability lead to the same inference : $Probability = 0.0133 < 0.05$ so that allows us to reject the null hypothesis of the normality test. So then, the asymmetric or leverage characteristic is proven.

II. Heteroskedasticity and ARCH effect

i) Conditional Heteroskedasticity

Given the graph of the raw series (Figure 1) which reveals strong variability, we can assume the existence of conditional heteroskedasticity.

In effect, the conditional heteroskedasticity test confirm the existence of volatility, because the coefficient of the series ($DLWTI^2$) is statistically significant (0.291). That is to say, the

variation of the series of prices at time t is a function of its evolution at time $t - 1$. Hence the confirmation of the existence of volatility.

ii) ARCH effect

Table 2 : ARCH effect due to the Covid-19 pandemic

| Heteroskedasticity Test: ARCH | | | | |
|---|-------------|-----------------------|-------------|--------|
| F-statistic | 56.98643 | Prob. F(1,1823) | 0.0000 | |
| Obs*R-squared | 55.31967 | Prob. Chi-Square(1) | 0.0000 | |
| Test Equation: | | | | |
| Dependent Variable: RESID^2 | | | | |
| Method: Least Squares | | | | |
| Date: 04/04/22 Time: 16:57 | | | | |
| Sample (adjusted): 1/05/2015 12/31/2021 | | | | |
| Included observations: 1825 after adjustments | | | | |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
| C | 1.315696 | 0.114518 | 11.48902 | 0.0000 |
| RESID^2(-1) | 0.174100 | 0.023063 | 7.548936 | 0.0000 |
| R-squared | 0.030312 | Mean dependent var | 1.593064 | |
| Adjusted R-squared | 0.029780 | S.D. dependent var | 4.704125 | |
| S.E. of regression | 4.633551 | Akaike info criterion | 5.905619 | |
| Sum squared resid | 39139.43 | Schwarz criterion | 5.911657 | |
| Log likelihood | -5386.878 | Hannan-Quinn criter. | 5.907846 | |
| F-statistic | 56.98643 | Durbin-Watson stat | 2.017159 | |
| Prob(F-statistic) | 0.000000 | | | |

Source : Authors' computation through E-views 10

The *ARCH* effect can be judged from Lagrange Multiplier (*LM*) statistics which is represented as an observed R-squared. The observed R-square statistics is 55.3196 and it is considered significant as its probability value is < 0.05 . Moreover, the F-statistics (56.9864) is also significant as its probability value is < 0.05 . thus the coefficient associated with $RESID^2(-1) = 0.1741$ is statistically significant at the 5% level. This proves that there is an existence of *ARCH* effect in the data of the crude oil from January 1, 2015 to December 31, 2021, therefore the raw series generating process is an *ARCH* of order 1. which we will estimate in the next section.

4.4.4 Estimation of models and search for the optimal model

1. Estimation of eventual models

After having tested the existence of *ARCH* effects and the other essentials related to *GARCH* modeling, three models have been formulated, *GARCH*(1,1), *GARCH*(1,2) and *TGARCH*(1,2). The specific conditions for each model are established. The models : *GARCH* (1,1) and *GARCH*(1,2) vanish the positivity condition of coefficients of variance and stationarity equation. The *TGARCH*(1,2) model satisfy the condition of asymmetry.

Furthermore, the standard way to select a model is, the coefficients, *ARCH* and *GARCH* should be significant and there should not be the existence of heteroscedasticity or *ARCH* effect (with the help of *ARCH Lagrange Multiplier* test) and Autocorrelation (with the help of correlogram of residuals and squared residuals) after framing the model, which is well verified in any of the three models adopted for the series analyzed in this work (see Table 3, in the section "Autocorrelation residuals and heteroskedasticity").

Table 3 : Estimations of selected models

| | GARCH(1,1) | GARCH(1,2) | TGARCH(1,2) |
|---|----------------------|----------------------|-----------------------------|
| Main Equation | | | |
| C (Prob) | 0.292819 (0.0186) | 0.252381 (0.0416) | 0.062799 (0.5999) |
| WTI(-1) (Prob) | 0.995287 (0.0000) | 0.995940 (0.0000) | 0.999146 (0.0000) |
| Variance Equation | | | |
| w (Prob) | 0.041242 (0.0000) | 0.065049 (0.0000) | 0.065460 (0.0000) |
| α_1 (Prob) | 0.082836 (0.0000) | 0.125001 (0.0000) | 0.044011 (0.0023) |
| B₁ (Prob) | 0.894473 (0.0000) | 0.191560 (0.0062) | 0.199395 (0.0073) |
| B₂ (Prob) | * | 0.646135 (0.0000) | 0.652199 (0.0000) |
| γ (Prob) | * | * | 0.121670 (0.0000) |
| Autocorrelation Residuals and Heteroskedasticity | | | |
| Autoc. Residuals | Non | Non | Non |
| Autoc. Squared Residuals | Non | Non | Non |
| Heteroskedasticity (ARCH LM-test) | No Signifacant | No Signifacant | No Signifacant |

Source : Authors' computation through E-views 10

2. Research of the optimal model

A) Information criteria

The choice of the most appropriate model among the estimated models is made on the basis of minimization of the criterion *Akaike* (*AIC*), *Schwarz* (*SIC*), on the one hand, and maximization of *Log – likelihood*, on the other hand. All these specificities which helped in the preliminary selection of the appropriate models are represented in Table 4 bellow,

Table 4 : Informations criteria to select the suitable model

| | GARCH(1,1) | GARCH(1,2) | TGARCH(1,2) |
|-------------------------------|-------------------|-------------------|--------------------|
| AIC | 3.167363 | 3.163288 | 3.153744 |
| SIC | 3.182450 | 3.181393 | 3.174866 |
| LL | -2886.802 | -2882.082 | -2872.368 |
| Lj-Bx squar (Prob) | 34.248 (0.552) | 33.695 (0.579) | 33.661 (0.580) |

Source : Authors computation throught E-views 10

B) Forecasting criteria

We use the good forecast criteria, RMSE (Root Mean Square Error), MAE (Mean Absolut Error), MAPE (Mean Absolute Percentage Error) and Theil inequality coefficient (close to zero), shown in Talbe 5 below, to select the optimal model among the three selected models.

Table 5 : Forecast standard criteria to select the suitable model

| | GARCH(1,1) | GARCH(1,2) | TGARCH(1,2) |
|----------------|-------------------|-------------------|--------------------|
| R.M.S.E | 14.08002 | 13.95559 | 14.98503 |
| MAE | 11.29239 | 11.18234 | 11.80188 |
| MAPE | 27.08691 | 26.82942 | 28.70260 |
| THEIL | 0.121919 | 0.120988 | 0.127465 |

Source : Authors' computation through E-views 10

According to the criteria of Table 5, the model who having more minimal values is the $GARCH(1, 2)$, so it is the better then $GARCH(1, 1)$ and $TGARCH(1, 2)$. Moreover, as can be deduced from the Table 6, this adopted model satisfies the necessary conditions,

$$\alpha_1 \geq 0, \beta_j \geq 0, \forall_j = 1, 2,$$

and the sufficient stationnarity condition (volatility persistence)

$$\alpha_1 + \sum_{j=1}^2 \beta_j = 0.9627006 < 1.$$

Hence, the model retained for predicting crude oil price volatility is the $GARCH(1, 2)$. The its estimation results are presented in the Table 6 bellow :

Table 6 : Estimation results of GARCH(1,2) model

Dependent Variable: WTI
Method: ML ARCH - Normal distribution (BFGS / Marquardt steps)
Date: 04/24/22 Time: 20:15
Sample (adjusted): 1/02/2015 12/31/2021
Included observations: 1826 after adjustments
Convergence achieved after 38 iterations
Coefficient covariance computed using outer product of gradients
Presample variance: backcast (parameter = 0.7)
GARCH = C(3) + C(4)*RESID(-1)^2 + C(5)*GARCH(-1) + C(6)*GARCH(-2)

| Variable | Coefficient | Std. Error | z-Statistic | Prob. |
|--------------------|-------------|-----------------------|-------------|--------|
| C | 0.252381 | 0.123890 | 2.037134 | 0.0416 |
| WTI(-1) | 0.995940 | 0.002209 | 450.8203 | 0.0000 |
| Variance Equation | | | | |
| C | 0.065049 | 0.013273 | 4.900944 | 0.0000 |
| RESID(-1)^2 | 0.125001 | 0.010435 | 11.97901 | 0.0000 |
| GARCH(-1) | 0.191560 | 0.070014 | 2.736033 | 0.0062 |
| GARCH(-2) | 0.646135 | 0.068545 | 9.426414 | 0.0000 |
| R-squared | 0.988707 | Mean dependent var | 53.10022 | |
| Adjusted R-squared | 0.988701 | S.D. dependent var | 11.89461 | |
| S.E. of regression | 1.264370 | Akaike info criterion | 3.163288 | |
| Sum squared resid | 2915.903 | Schwarz criterion | 3.181393 | |
| Log likelihood | -2882.082 | Hannan-Quinn criter. | 3.169966 | |
| Durbin-Watson stat | 2.083557 | | | |

Source : Authors' computation through E-views 10

Table 6 shows the estimation results of the $GARCH(1, 2)$ model. The results contain two parts. The upper part shows the main equation and the lower part represents the variance equation.

In the main equation, both the constant (C) and the co-efficient of first lag [$WTI(-1)$] are significant as the probability value is < 0.05 . In case of variance equation, $C(3)$ is the constant, $C(4)$ is the $ARCH$ co-efficient, $C(5)$ and $C(6)$ are the $GARCH$ coefficients of first and second lag respectively. All the coefficients in the variance equation are significant as their probability values are < 0.05 .

Hence the variance equation can be shown as given below

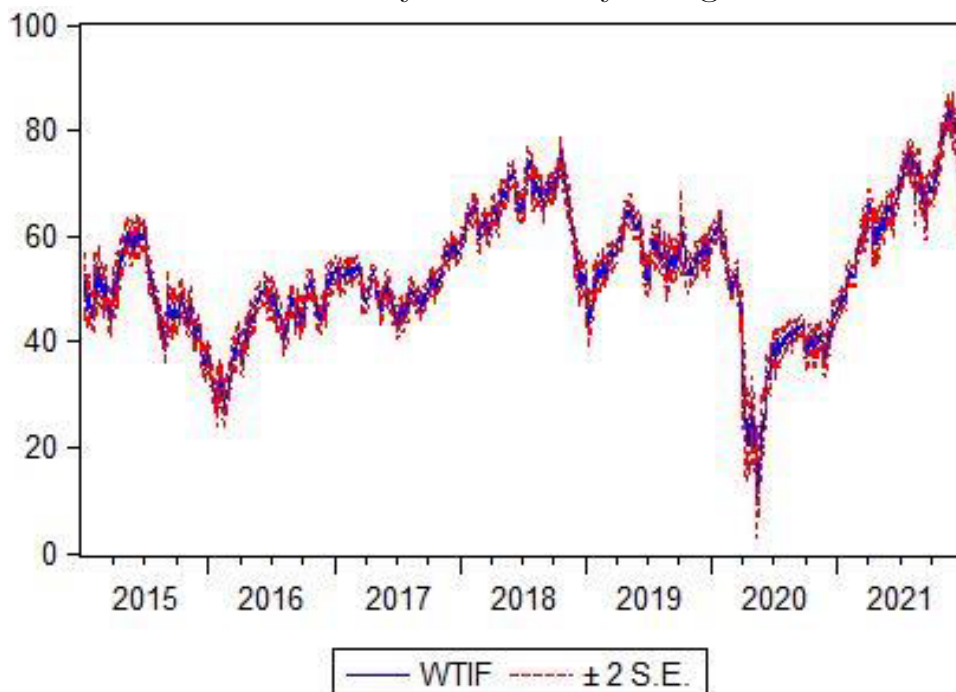
$$h_t = 0.065049 + 0.125001\varepsilon_{t-1}^2 + 0.191560h_{t-1} + 0.646135h_{t-2}. \quad (4.1)$$

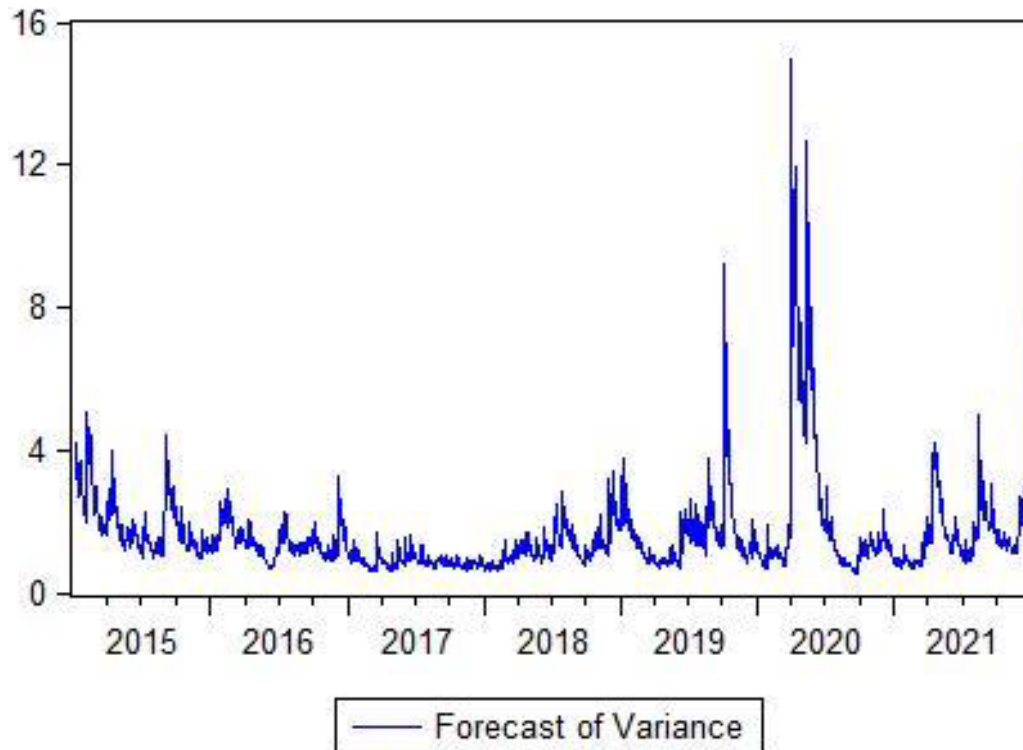
4.4.5 Forecasting using the formulated model

By using the above-formulated model, the crude oil prices and volatility of crude oil prices have been forecasted in the forecasting section.

The Figure 5 explains how the forecasting graph made by the adopted model (WTIF) coincides with that made by the actual WTI data (Figure 1), also for the graphs of the predictive volatility and the graph of the volatility of the actual data, which proves the validity of the adopted model.

Figure 5 : WTIF and Volatility Forecast by using the formulated model





Source : Authors' computation throught E-views 10

Table 7 : Actual and forecasted oil peices of last 3 months of 2021

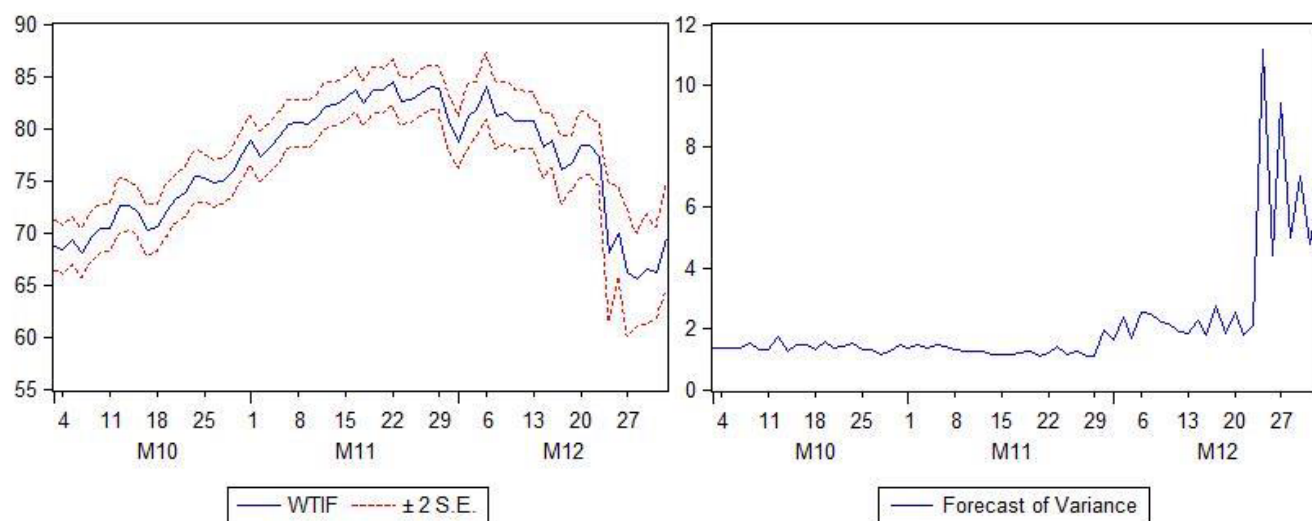
| Day | October Actual | October Forecast | Day | November Actual | November Forecast | Day | December Actual | December Forecast |
|-------|----------------|------------------|-------|-----------------|-------------------|-------|-----------------|-------------------|
| 01/10 | 75.88 | 74.98 | 01/11 | 84.05 | 83.48 | 01/12 | 65.57 | 66.16 |
| 04/10 | 77.62 | 75.82 | 02/11 | 83.91 | 83.96 | 02/12 | 66.5 | 65.56 |
| 05/10 | 78.93 | 77.56 | 03/11 | 80.86 | 83.82 | 03/12 | 66.26 | 66.48 |
| 06/10 | 77.43 | 78.86 | 04/11 | 78.81 | 80.78 | 06/12 | 69.49 | 66.24 |
| 07/10 | 78.3 | 77.37 | 05/11 | 81.27 | 78.74 | 07/12 | 72.05 | 69.46 |
| 08/10 | 79.35 | 78.23 | 08/11 | 81.93 | 81.19 | 08/12 | 72.36 | 70.23 |
| 11/10 | 80.52 | 79.28 | 09/11 | 84.15 | 81.85 | 09/12 | 70.94 | 68.89 |
| 12/10 | 80.64 | 80.45 | 10/11 | 81.34 | 84.06 | 10/12 | 71.67 | 70.02 |
| 13/10 | 80.44 | 80.57 | 11/11 | 81.59 | 81.26 | 13/12 | 71.29 | 69.88 |
| 14/10 | 81.31 | 80.37 | 12/11 | 80.79 | 81.51 | 14/12 | 70.73 | 69.12 |
| 15/10 | 82.28 | 81.23 | 15/11 | 80.88 | 80.71 | 15/12 | 70.87 | 71.11 |
| 18/10 | 82.44 | 82.20 | 16/11 | 80.76 | 80.80 | 16/12 | 72.38 | 70.95 |
| 19/10 | 82.96 | 82.36 | 17/11 | 78.36 | 80.68 | 17/12 | 70.86 | 68.75 |
| 20/10 | 83.87 | 82.88 | 18/11 | 79.01 | 78.29 | 20/12 | 68.23 | 69.65 |
| 21/10 | 82.5 | 83.788 | 19/11 | 76.1 | 78.94 | 21/12 | 71.12 | 70.44 |
| 22/10 | 83.76 | 82.42 | 22/11 | 76.75 | 76.04 | 22/12 | 72.76 | 73.01 |
| 25/10 | 83.76 | 83.672 | 23/11 | 78.5 | 76.69 | 23/12 | 73.79 | 72.78 |
| 26/10 | 84.65 | 83.672 | 24/11 | 78.39 | 78.43 | 27/12 | 75.57 | 74.02 |
| 27/10 | 82.66 | 84.56 | 25/11 | 77.41 | 78.32 | 28/12 | 75.98 | 76.21 |
| 28/10 | 82.81 | 82.58 | 26/11 | 68.15 | 77.35 | 29/12 | 76.56 | 77.05 |
| 29/10 | 83.57 | 82.73 | 29/11 | 69.95 | 68.13 | 30/12 | 76.99 | 75.36 |
| | | | 30/11 | 66.18 | 69.92 | 31/12 | 75.21 | 76.08 |

Source : Authors' computation throught E-views10

The Table 7, can be confirmed these results by comparing the values of the real data (actual)

and the values provided by the equation of the $GARCH(1,2)$ model adopted (forecast).

Figure 6 : Forecasting crude oil volatility based on modified 3 months data (from 1st October 2021 to 31th December 2021 by using the formulated model



Source : Authors' computation through E-views10

Figure 6 shows the forecasted prices of crude oil and the forecasted variance for crude oil for the modeling period from October 01, 2021 to December 31, 2021. The first graph of forecasted prices of crude oil depicts that its general trend has seen two essential phases, the first is characterized by growth, crude oil prices have increased since the beginning of October 2021 until reaching a high peak on November 22, while the second phase is characterized by a decrease, in a single month the prices fell in a continuous and remarkable way with fluctuations, from November 22 until the last of December of the year, when they started to rise again another time.

While the predictive variance graph in turn experienced the same two phases, the first witnessing a quiet period of variance from October 1 to November 22, which corresponds exactly to the upward phase of crude oil prices (good news : The expansion of the Corona virus epidemic recovery and the removal of the global containment measure), the second witness to a wildly fluctuating period of variance (volatility) from November 22 until the end of December, which corresponds to the bearish phase of crude oil prices.

Conclusion

Plusieurs résultats importants peuvent être conclus à partir de notre analyse des données, c'est que l'effet de levier ou l'asymétrie de la volatilité est observé en raison de la propagation de la pandémie qui a un impact sur la volatilité des prix du pétrole brut, ce qui est mis en évidence par le nombre de modèles asymétriques découverts à travers l'étude, autres que ceux mentionnés dans cet article, tels que $TGARCH(1, 1)$, $TGARCH(2, 1)$, $TGARCH(3, 1)$, $EGARCH(1, 2)$, $EGARCH(2, 1)$ and $EGARCH(3, 1)$, qui font part aux modèles asymétriques, tels que le coefficient d'asymétrie λ soit négatif. Malgré qu'en a certains peuvent s'adapter aux données étudiées, mais ils sont exclus de la sélection à cause de l'absence des critères de minimisation ou de maximisation.

Les graphes de prévision ci-dessus des trois mois des prix du pétrole brut indique, notamment au dernier de cette période (décembre 2021), qu'il est difficile de prédire la volatilité des prix car le graphe de volatilité est extrêmement fluctuant. Les investisseurs dans les marchés des matières premières qui concentrent leurs activités sur le pétrole brut peuvent utiliser les modèles formulés pour prendre des décisions d'investissement.

En général, les modèles GARCH, intéressés par l'étude des séries chronologiques et qui intervient dans la résolution des problèmes influant dans l'économie mondiale, dignent de l'attention par rapport aux autres méthodologies, en raison de la réputation qu'il a acquise notamment après que c'est la raison pour laquelle l'économètre et son fondateur Robert Engel a remporté le prix Nobel en 2003. Ces modèles ont réellement provoqué une révolution scientifique dans cet aspect et surtout que leurs prédécesseurs les modèles ARMA ont montré une limite par rapport aux plusieurs domaines, notamment économiques et financière. Cet étude qui a été portée dans le cas univarié peut étendra au cas multivarié en introduisant l'interaction de plusieurs phénomènes ou actifs dans le cadre de la méthodologie de Box-Jenkins, par les modèles GARCH qui sont capables de gérer des situations multivariées telles que le modèle BEKK-GARCH, modèle CCC ou aussi modèle DCC.

Annexe

A1. Test de stationnarité ADF (Augmented Dickey-Fuller) :

A1.1 Table ADF :

| N | Model 0 - no constant, no trend | | | | Model 1 - constant, no trend | | | | Model 2 - constant, trend | | | |
|------|---------------------------------|--------|--------|--------|------------------------------|--------|--------|--------|---------------------------|--------|--------|--------|
| | 0.01 | 0.025 | 0.05 | 0.10 | 0.01 | 0.025 | 0.05 | 0.10 | 0.01 | 0.025 | 0.05 | 0.10 |
| 25 | -2.661 | -2.273 | -1.955 | -1.609 | -3.724 | -3.318 | -2.986 | -2.633 | -4.375 | -3.943 | -3.589 | -3.238 |
| 50 | -2.612 | -2.246 | -1.947 | -1.612 | -3.568 | -3.213 | -2.921 | -2.599 | -4.152 | -3.791 | -3.495 | -3.181 |
| 100 | -2.588 | -2.234 | -1.944 | -1.614 | -3.498 | -3.164 | -2.891 | -2.582 | -4.052 | -3.722 | -3.452 | -3.153 |
| 250 | -2.575 | -2.227 | -1.942 | -1.616 | -3.457 | -3.136 | -2.873 | -2.573 | -3.995 | -3.683 | -3.427 | -3.137 |
| 500 | -2.570 | -2.224 | -1.942 | -1.616 | -3.443 | -3.127 | -2.867 | -2.570 | -3.977 | -3.670 | -3.419 | -3.132 |
| >500 | -2.567 | -2.223 | -1.941 | -1.616 | -3.434 | -3.120 | -2.863 | -2.568 | -3.963 | -3.660 | -3.413 | -3.128 |

Tale A1.1 : Augmented Dickey-Fuller Table

A1.2 Test de trois modèles de racine unitaire de la série brute WTI :

Null Hypothesis: WTI has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 0 (Automatic - based on SIC, maxlag=24)

| | t-Statistic | Prob.* |
|---|------------------|---------------|
| Augmented Dickey-Fuller test statistic | -2.322360 | 0.4210 |
| Test critical values: | | |
| 1% level | -3.963076 | |
| 5% level | -3.412271 | |
| 10% level | -3.128067 | |

*Mackinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(WTI)
Method: Least Squares
Date: 07/18/23 Time: 17:43
Sample (adjusted): 1/02/2015 12/31/2021
Included observations: 1826 after adjustments

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|---------------------|-------------|-----------------------|-------------|--------|
| WTI(-1) | -0.006123 | 0.002637 | -2.322360 | 0.0203 |
| C | 0.263563 | 0.135642 | 1.943074 | 0.0522 |
| @TREND("1/01/2015") | 7.89E-05 | 5.94E-05 | 1.328046 | 0.1843 |
| R-squared | 0.003143 | Mean dependent var | 0.010602 | |
| Adjusted R-squared | 0.002049 | S.D. dependent var | 1.265083 | |
| S.E. of regression | 1.263786 | Akaike info criterion | 3.307743 | |
| Sum squared resid | 2911.615 | Schwarz criterion | 3.316796 | |
| Log likelihood | -3016.969 | Hannan-Quinn criter. | 3.311082 | |
| F-statistic | 2.873663 | Durbin-Watson stat | 2.082323 | |
| Prob(F-statistic) | 0.056748 | | | |

Null Hypothesis: WTI has a unit root
Exogenous: Constant
Lag Length: 0 (Automatic - based on SIC, maxlag=24)

| | t-Statistic | Prob.* |
|---|------------------|---------------|
| Augmented Dickey-Fuller test statistic | -1.995483 | 0.2890 |
| Test critical values: | | |
| 1% level | -3.433730 | |
| 5% level | -2.862919 | |
| 10% level | -2.567551 | |

*Mackinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(WTI)
Method: Least Squares
Date: 07/18/23 Time: 17:45
Sample (adjusted): 1/02/2015 12/31/2021
Included observations: 1826 after adjustments

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|--------------------|-------------|-----------------------|-------------|--------|
| WTI(-1) | -0.004967 | 0.002489 | -1.995483 | 0.0461 |
| C | 0.274322 | 0.135428 | 2.025587 | 0.0430 |
| R-squared | 0.002178 | Mean dependent var | 0.010602 | |
| Adjusted R-squared | 0.001631 | S.D. dependent var | 1.265083 | |
| S.E. of regression | 1.264051 | Akaike info criterion | 3.307615 | |
| Sum squared resid | 2914.432 | Schwarz criterion | 3.313650 | |
| Log likelihood | -3017.852 | Hannan-Quinn criter. | 3.309841 | |
| F-statistic | 3.981952 | Durbin-Watson stat | 2.082716 | |
| Prob(F-statistic) | 0.046139 | | | |

Table A1.2 : Test de racine unitaire pour WTI (Modle3).

Table A1.2 : Test de racine unitaire pour WTI (Modle 2)..

Null Hypothesis: WTI has a unit root
Exogenous: None
Lag Length: 0 (Automatic - based on SIC, maxlag=24)

| | t-Statistic | Prob.* |
|---|------------------|---------------|
| Augmented Dickey-Fuller test statistic | -0.086028 | 0.6539 |
| Test critical values: | | |
| 1% level | -2.566229 | |
| 5% level | -1.940997 | |
| 10% level | -1.616583 | |

*Mackinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(WTI)
Method: Least Squares
Date: 07/18/23 Time: 17:46
Sample (adjusted): 1/02/2015 12/31/2021
Included observations: 1826 after adjustments

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|--------------------|-------------|-----------------------|-------------|--------|
| WTI(-1) | -4.68E-05 | 0.000544 | -0.086028 | 0.9315 |
| R-squared | -0.000066 | Mean dependent var | 0.010602 | |
| Adjusted R-squared | -0.000066 | S.D. dependent var | 1.265083 | |
| S.E. of regression | 1.265125 | Akaike info criterion | 3.308766 | |
| Sum squared resid | 2920.988 | Schwarz criterion | 3.311784 | |
| Log likelihood | -3019.904 | Hannan-Quinn criter. | 3.309880 | |
| Durbin-Watson stat | 2.088294 | | | |

Table A1.2 : Test de racine unitaire pour WTI (Modle 1)..

A1.3 Test de trois modèles de racine unitaire de la série DLWTI :

Null Hypothesis: DLWTI has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 3 (Automatic - based on SIC, maxlag=24)

| | t-Statistic | Prob.* |
|---|------------------|---------------|
| Augmented Dickey-Fuller test statistic | -20.80389 | 0.0000 |
| Test critical values: | | |
| 1% level | -3.963087 | |
| 5% level | -3.412277 | |
| 10% level | -3.128071 | |

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(DLWTI)
Method: Least Squares
Date: 07/18/23 Time: 17:31
Sample (adjusted): 1/08/2015 12/31/2021
Included observations: 1822 after adjustments

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|---------------------|-------------|-----------------------|-------------|--------|
| DLWTI(-1) | -1.079964 | 0.051912 | -20.80389 | 0.0000 |
| D(DLWTI(-1)) | 0.015722 | 0.043405 | 0.362215 | 0.7172 |
| D(DLWTI(-2)) | -0.062235 | 0.034096 | -1.825284 | 0.0681 |
| D(DLWTI(-3)) | -0.134601 | 0.023249 | -5.789684 | 0.0000 |
| C | -0.000240 | 0.001556 | -0.154399 | 0.8773 |
| @TREND("1/01/2015") | 5.20E-07 | 1.47E-06 | 0.352594 | 0.7244 |
| R-squared | 0.546244 | Mean dependent var | 1.83E-05 | |
| Adjusted R-squared | 0.544995 | S.D. dependent var | 0.049052 | |
| S.E. of regression | 0.033088 | Akaike info criterion | -3.976021 | |
| Sum squared resid | 1.988153 | Schwarz criterion | -3.957884 | |
| Log likelihood | 3628.155 | Hannan-Quinn criter. | -3.969330 | |
| F-statistic | 437.2305 | Durbin-Watson stat | 2.013715 | |
| Prob(F-statistic) | 0.000000 | | | |

Table A1.3 : Test de racine unitaire pour DLWTI (Modèle 3)..

Null Hypothesis: DLWTI has a unit root
Exogenous: Constant
Lag Length: 3 (Automatic - based on SIC, maxlag=24)

| | t-Statistic | Prob.* |
|---|------------------|---------------|
| Augmented Dickey-Fuller test statistic | -20.80599 | 0.0000 |
| Test critical values: | | |
| 1% level | -3.433737 | |
| 5% level | -2.862923 | |
| 10% level | -2.567553 | |

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(DLWTI)
Method: Least Squares
Date: 07/18/23 Time: 17:11
Sample (adjusted): 1/08/2015 12/31/2021
Included observations: 1822 after adjustments

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|--------------------|-------------|-----------------------|-------------|--------|
| DLWTI(-1) | -1.079702 | 0.051894 | -20.80599 | 0.0000 |
| D(DLWTI(-1)) | 0.015529 | 0.043391 | 0.357870 | 0.7205 |
| D(DLWTI(-2)) | -0.062373 | 0.034085 | -1.829907 | 0.0674 |
| D(DLWTI(-3)) | -0.134675 | 0.023242 | -5.794471 | 0.0000 |
| C | 0.000235 | 0.000775 | 0.303841 | 0.7613 |
| R-squared | 0.546213 | Mean dependent var | 1.83E-05 | |
| Adjusted R-squared | 0.545214 | S.D. dependent var | 0.049052 | |
| S.E. of regression | 0.033080 | Akaike info criterion | -3.977050 | |
| Sum squared resid | 1.988289 | Schwarz criterion | -3.961936 | |
| Log likelihood | 3628.093 | Hannan-Quinn criter. | -3.971474 | |
| F-statistic | 546.7706 | Durbin-Watson stat | 2.013722 | |
| Prob(F-statistic) | 0.000000 | | | |

Table A1.3 : Test de racine unitaire pour DLWTI (Modèle 2)..

Null Hypothesis: DLWTI has a unit root
Exogenous: None
Lag Length: 3 (Automatic - based on SIC, maxlag=24)

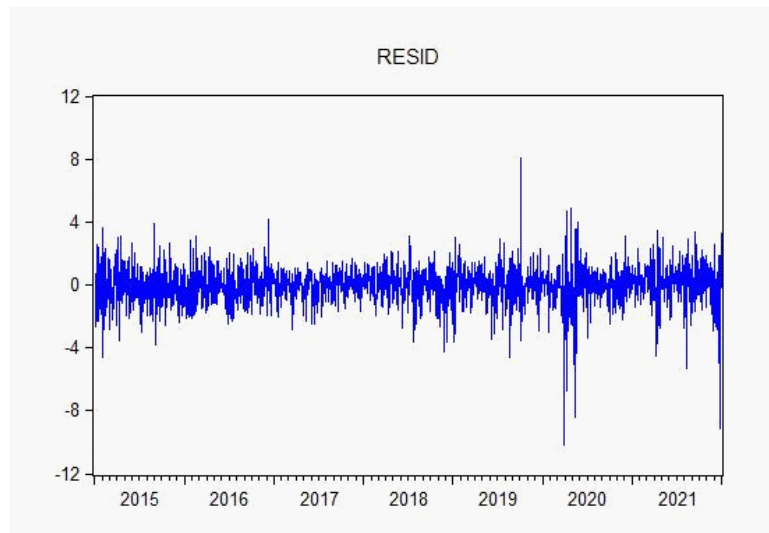
| | t-Statistic | Prob.* |
|---|------------------|---------------|
| Augmented Dickey-Fuller test statistic | -20.80911 | 0.0000 |
| Test critical values: | | |
| 1% level | -2.566232 | |
| 5% level | -1.940997 | |
| 10% level | -1.616582 | |

*MacKinnon (1996) one-sided p-values.

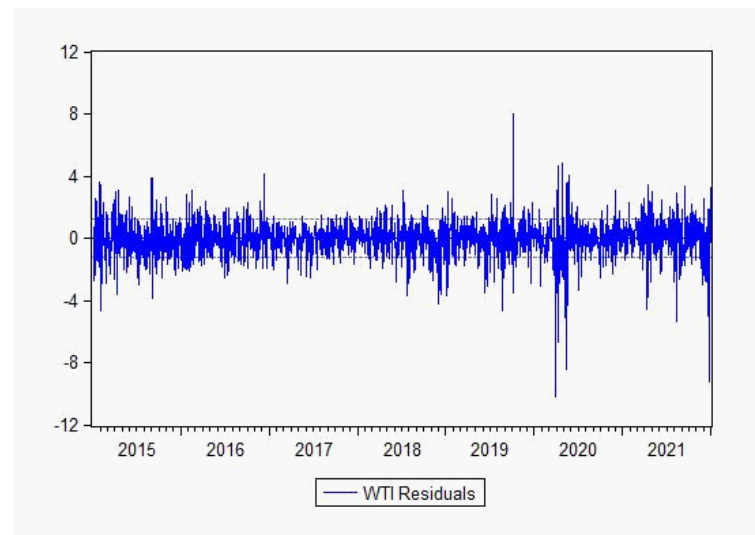
Augmented Dickey-Fuller Test Equation
Dependent Variable: D(DLWTI)
Method: Least Squares
Date: 07/18/23 Time: 17:21
Sample (adjusted): 1/08/2015 12/31/2021
Included observations: 1822 after adjustments

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|--------------------|-------------|-----------------------|-------------|--------|
| DLWTI(-1) | -1.079531 | 0.051878 | -20.80911 | 0.0000 |
| D(DLWTI(-1)) | 0.015407 | 0.043379 | 0.355162 | 0.7225 |
| D(DLWTI(-2)) | -0.062449 | 0.034076 | -1.832654 | 0.0670 |
| D(DLWTI(-3)) | -0.134710 | 0.023236 | -5.797485 | 0.0000 |
| R-squared | 0.546190 | Mean dependent var | 1.83E-05 | |
| Adjusted R-squared | 0.545441 | S.D. dependent var | 0.049052 | |
| S.E. of regression | 0.033071 | Akaike info criterion | -3.978097 | |
| Sum squared resid | 1.988390 | Schwarz criterion | -3.966005 | |
| Log likelihood | 3628.046 | Hannan-Quinn criter. | -3.973636 | |
| Durbin-Watson stat | 2.013723 | | | |

Table A1.3 : Test de racine unitaire pour DLWTI (Modèle 1)..

A2. Graphes des Résidus :**A2.1 Graphe des Résidus de la série WTI :**

Graphe A 2.1 : Rsidus de WTI

A2.2. Graphe des Résidus par modèle estimé :

Graphe A 2.2 : Rsidus de WTI par modle estim

BIBLIOGRAPHIE

- [1] Amiri O., Remita M.R., (2022), Measuring the impact of covid-19 pandemic on the oil prices volatility by GARCH modeling, *advances in mathematics : Scientific Journal*, 11(6), 515-537.
- [2] Bartlett M.S., (1937), "Properties of sufficiency and statistical tests". *Proceedings of the Royal Statistical Society, Series A* 160, 268-282.
- [3] Black F. (1976), *Studies of Stock Price Volatility Changes*. *Proceedings of the Business and Economics Section of the American Statistical Association*, ,177–181.
- [4] Bollerslev T., (1986), Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* 31, 307–327
- [5] Bollerslev T., Chou R.Y., Kroner K.F., (1992), ARCH modeling in finance : a review of the theory and empirical evidence. *Journal of Econometrics* 52, 5–59
- [6] Bougerol P., Picard N., (1992a), Strict stationarity of generalized autoregressive processes. *Annals of Probability* 20, 1714–1729.
- [7] Bougerol P., Picard N., (1992b), Stationarity of GARCH processes and of some non negative time series. *Journal of Econometrics* 52, 115–127.

-
- [8] Box G.E.P., JENKINS G.M., (1970), Time Series Analysis : Forecasting and Control, Holden-Day, San Fransisco : Holden-Day, First edition.
- [9] Box G.E.P., JENKINS G.M., (1994), Time Series Analysis : Forecasting and Control, Holden-Day, San Fransisco : Holden-Day, (1976). Revised Edition.
- [10] Box G.E.P., JENKINS G.M., Reinsel G.C., Time Series Analysis : Forecasting and Control, Prentice Hall.
- [11] Brockwell P.J., Davis R.A., (1991), Time Series : Theory and Methods, 2nd edition, Springer Science, New York.
- [12] Brockwell P.J., Davis R.A., (2002), Inroduction to Time Series and Forecasting, 2nd edition, Springer-Verlag, New York, Inc.
- [13] Charpentier A., Introduction à la théorie des processus en tempts discret : Modèles ARIMA et méthode Box & Jenkins, Cours de séries temporelles théorie et applications, Université Paris Dauphine, 2000-2005.
- [14] Cooley J.W., Tukey J.W., (1965), An Algorithm for the Machine Calculation of Complex Fourier Series,. American Mathematical Society-Mathematics of Computation, 19 (90) : 297-301.
- [15] Dickey D.A., Fuller W.A., (1979), "Distribution of the Estimators for Autoregressive Time Series with Unit Root". Journal of the American (366) : 74. Statistical Association : doi 427-431.
- [16] Dickey D.A., Fuller W.A., (1981), Likelihood Ratio Statistics for Autoregressive Time Series With a Unit Root, *Economica*, 49(4), 1057-1072.
- [17] Drost F.C., Nijman T.E., (1993), Temporal aggregation of GARCH processes. *Econometrica* 61, 909-927.
- [18] Dum D.Z., Essi I.D., (2017), Modeling Price Volatility of Nigerian Crude Oil Markets Using Garch Model : 1987-2017. *International Journal of Applied Science and Mathematical Theory*, 3(4), 23-49.

- [19] EL-GHINI A., SAIDI Y., (2016), Return and volatility spillovers in the moroccan stock market during the financial crisis, *Empirical Economics*, ISSN 0377-7332, DOI 10.1007/s00181-016-1110-8.
- [20] Endri E., Aipama W., Razak A., Sari L., Septiano R., (2021), Stock price volatility during the COVID-19 pandemic : The GARCH model, *Investment Management and Financial Innovations Journal*, 18(4), 12-20.
- [21] Engle R.F., (1982), Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation. *Econometrica* 50, 987–1008
- [22] Engle, R.F., Bollerslev T., (1986), Modelling the persistence of conditional variances (with comments and a reply by the authors). *Econometric Reviews* 5, 1–87.
- [23] Francq C., Zakoian J.-M., (2007), Quasi-maximum likelihood estimation in GARCH processes when some coefficients are equal to zero. *Stochastic Processes and their Applications* 117, 1265–1284.
- [24] Francq C., Zakoian J.-M., (2009), Modèles GARCH : Structure, Inférence statistique et applications financières, *Economica*.
- [25] Glosten L.R., Gannathan R., Runkle D.E., (1993), On The Relation between The Expected Value and The Volatility of Nominal Excess Return on stocks, *Journal of Finance*, 48, 1779-1801.
- [26] Guesmi K., Saadi S., Abid I., Ftiti Z., (2019), Portfolio diversification with virtual currency : Evidence from bitcoin, *International Review of Financial Analysis*, 63, 431-437.
- [27] Hawaldar I.T., Rajesha T.M., Sarea A.M., (2020), Causal nexus between the anomalies in the crude oil price and stock market, *International Journal of Energy Economics and Policy*, 10(3), 233-238.
- [28] Higgins M.L., Bera A.K., (1992), A class of nonlinear ARCH models. *International Economic Review*, 33(1), 137–158.
- [29] Kim, S., Shephard, N. and Chib, S. (1998), Stochastic volatility : likelihood inference and comparison with ARCH models. *Review of Economic Studies* 65, 361– 396

- [30] Lee J, Strazicich MC (2013), Minimum LM unit root test with one structural break. *Econ Bull* 33(4) :2483–2492
- [31] Ljung, G.M. and Box, G.E.P. (1978), On the measure of lack of fit in time series models. *Biometrika* 65, 297–303.
- [32] Meher B.K., Hawaldar I.T., Mohapatra L., Sarea A.M., (2020), The Impact of COVID-19 on Price Volatility of Crude Oil and Natural Gas Listed on Multi Commodity Exchange of India, *International Journal of Energy Economics and Policy*, 10(5), 422-431.
- [33] Merabet F., Zeghdoudi H., Ahmed Y.R., Saba I., (2021), Modelling of oil price volatility using ARIMA-GARCH models, *Advances in Mathematics : Scientific Journal* 10(5), 2361–2380.
- [34] Mia M.S., Rahman M.S., Evaluating the Forecast Accuracy of Exchange Rate Volatility in Bangladesh Using ARCH Family of Models., *Am. J. Math. Stat.*, 9(5), (2019), 183–190.
- [35] Nelson, D.B. (1990a), Stationarity and persistence in the GARCH(1, 1) model. *Econometric Theory* 6, 318– 334
- [36] Nelson, D.B. (1991), Conditional heteroscedasticity in asset returns : a new approach, *Econometrica*, 59(2), 347-370.
- [37] Ng A (2000), Volatility spillover effects from Japan and the US to the Pacific Basin. *J Int Money Finan* 19 :207–233.
- [38] Ng H.G., M. McAleer, (2004), Recursive Modelling of Symmetric and Asymmetric Volatility in the Presence of Extreme Observations, *Int. J. Forecast.*, 20, 115–129.
- [39] Nugroho D.B., Susanto B., Prasetia K.N.P., Rorimpandey R., (2019), Modeling of Returns Volatility using GARCH (1,1) Model under Tukey Transformations, *Journal Akuntansi dan Keuangan*, 21(1), 12-20.
- [40] Olstad A. , Filis G., Degiannakis S., (2021), Oil and currency volatilities : Co-movements and hedging opportunities, *International Journal of Finance & Economics*, 26(2), 2351-2374.
- [41] Pantula, S.G. (1989) Estimation of autoregressive models with ARCH errors. *Sankhya B* 50, 119–138.

-
- [42] Peng L. and Yao, Q. (2003), Least absolute deviations estimation for ARCH and GARCH models. *Biometrika* 90, 967–975.
- [43] Phillips P. C. B., Perron P., (1988), Testing for a Unit Root in a Time Series Regression. *Biometrika*, 75, p. 335-346
- [44] Pindyck R.S. (2004), Volatility in natural gas and oil markets, *The Journal of Energy and Development*, 30(1), 1-20.
- [45] Ramzan S., Ramzan S.Z., Zahid F.M., (2012), Modelling and forecasting exchange rate dynamics in Pakistan using ARCH family of models, *Electronic Journal of Applied Statistical Analysis*, 5(1), 15-29.
- [46] Sentana, E. (1995), Quadratic ARCH models. *Review of Economic Studies* 62, 639–661.
- [47] Slutsky E., (1937), the summation of random causes as the source of cyclical processes, *Econometrica-The Econometric Society*, 5(2), p. 105-146.
- [48] Straumann, D. (2005), Estimation in Conditionally Heteroscedastic Time Series Models. *Lecture Notes in Statistics* 181. Berlin : Springer.
- [49] Straumann, D. and Mikosch, T. (2006), Quasi-maximum-likelihood estimation in heteroscedastic time series : a stochastic recurrence equation approach. *Annals of Statistics* 34, 2449–2495
- [50] Yaya O.S., Tumala M.M., Udomboso C.G (2016), Volatility persistence and returns spillovers between oil and gold prices : Analysis before and after the global financial crisis. *Resources Policy*, 49, 273–281.
- [51] Zakoïan J.M., (1994), Threshold heteroskedastic models, *Journal of Economic Dynamics and Control*, 18(5), 931-955.
- [52] Yule G.U., (1927), On a method of investigating periodicities disturbed series, with special reference to Wolfer’s sunspot numbers, *Royal Society*, vol. 226 Issue 636-646, pp. 267-298.