

Remerciements

En tout premier lieu, je remercie le bon Dieu, tout puissant, de m'avoir donné la force pour survivre, ainsi que l'audace pour dépasser toutes les difficultés.

Ma plus grande gratitude va à mon encadreur M^m **Atil Lynda**, pour sa disponibilité, l'autonomie et la confiance qu'elle m'a accordée. J'aimerais aussi la remercier pour ses précieux conseils qui m'ont permis de mener à bien ce travail.

L'enseignement de qualité dispensé par l'équipe de formation du master Probabilités et Statistiques de l'université Mouloud Mammeri de Tizi-Ouzou a su nourrir mes réflexions et a représenté une profonde satisfaction intellectuelle, merci donc aux enseignants-chercheurs.

Mes sincères remerciements vont aux membres de jury pour l'intérêt qu'ils ont porté à mon travail en acceptant de l'examiner et de l'enrichir par leurs propositions et remarques.

Je remercie mes très chers parents, qui ont toujours été là pour moi, «Vous avez tout sacrifié pour mon frère et moi n'épargnant ni santé ni efforts. Vous m'avez donné un magnifique modèle de labeur et de persévérance. Je suis redevable d'une éducation dont je suis fière».

Un grand merci à mon petit frère chéri, mon ami, mon confident **Idir** pour ton amour, ton soutien sans limites malgré la distance qui nous sépare que dieu te garde pour moi trésor.

Enfin, je remercie tou(te)s mes ami(e)s que j'aime tant, Massiva, Dyhia, Nedjma, Soraya, Hassane, Mehdi, Hakim, ... Pour leur sincère amitié et confiance, et à qui je dois ma reconnaissance et mon attachement.

À vous tous, je présente mes remerciements, mon respect et ma gratitude.

Table des matières

Introduction générale

La très grande fréquence des données financières a favorisé l'étude des impacts de court et de long terme d'une brusque fluctuation sur la dynamique des séries. Elle met en évidence le concept de mémoire longue dans les séries financières et microéconomiques. Pendant de nombreuses années, le modèle de marche aléatoire a été considéré le meilleur prédicteur de la dynamique de la volatilité sur les marchés de change. Ce modèle présume qu'on ne peut pas battre systématiquement le marché : les marchés sont efficients. L'étude de séries financières révèle une dépendance temporelle du risque qui disparaît souvent progressivement. Les recherches se sont orientées vers l'étude des dynamiques conditionnelles. Toutefois, ces travaux ne différencient pas la dépendance conditionnelle révélant de la moyenne de celle de la variance. Deux volets sont alors issus. Le premier étudie l'équation de la moyenne conditionnelle avec les modèles ARMA, ARIMA, ARFIMA... Le second s'est focalisé sur la variance conditionnelle. Deux classes de modèles non linéaires se sont développées pour caractériser cette variance.

- La première classe se base sur les modèles autoregressifs conditionnellement hétéroscédastiques (ARCH) de Engle (1982) et ses extensions, le modèle d'hétéroscédasticité conditionnelle autoregressif généralisé (GARCH) de Bollerslev (1986) et le modèle GARCH intégré de Engle et de Bollerslev (1986).

- La deuxième classe fait référence aux modèles de volatilité stochastique (SV) introduit par Taylor (1986).

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

que plus l'innovation est lointaine dans le passé plus faible sera son impact sur la variance conditionnelle : la valeur de la volatilité est une fonction décroissante de l'ordre de retard. La spécification linéaire de la variance conditionnelle semble saisir l'effet des erreurs passées ou de mémoire longue trouvé dans les travaux empiriques. Cette spécification ne pose pas de problèmes d'estimation particuliers. Toutefois, pour que la variance conditionnelle soit positive, des contraintes supplémentaires sur les paramètres du modèle sont nécessaires. Pour contourner cette dernière difficulté, Engle (1982, 1983) et Engle et Kraft (1983) proposent de fixer le nombre de retards à prendre en compte au lieu de travailler avec des ARCH infinis. En se basant sur les travaux faits sur l'équation de la moyenne conditionnelle, une extension généralisée des modèles ARCH a été développée : le modèle GARCH. Ce dernier tient compte non seulement de la volatilité courante exprimée par les carrés des résidus passés mais aussi de la volatilité passée, il présente ainsi une spécification plus flexible de la variance conditionnelle. Toutefois, le modèle GARCH ne permet de modéliser que la dépendance temporelle de court terme : l'effet des chocs décroît exponentiellement dans le temps. Pour étendre la mémoire du modèle GARCH, Engle et Bollerslev (1986) introduisent une extension du modèle GARCH dite intégrée (IGARCH) qui présente une mémoire explosive : les effets des chocs sont persistants à l'infini. Compte tenu des propriétés opposées des modèles traditionnels : décroissance exponentielle de l'impact des chocs sur la volatilité des modèles GARCH par opposition leur persistance à l'infinie dans les modèles IGARCH. Un troisième modèle plus flexible a été introduit sur le marché de change, Baillie, Bollerslev et Mikkelsen (1996) ont proposé une version ajustée des deux modèles : c'est le processus GARCH fractionnement intégré (FIGARCH). Ces modèles sont spécifiés par analogie avec les modèles ARFIMA dans l'équation de la moyenne. Le but de notre travail est de faire l'inférence statistique des modèles autoregressifs conditionnellement hétéroscédastiques.

Chapitre 1

Généralités

1.1 Introduction

Dans ce chapitre nous allons décrire les principales formulations des modèles avec hétéroscédasticité conditionnelle qui ont été proposés dans la littérature (Engle(1982)). En tout premier lieu nous allons présenter le modèle ARCH de manière générale et brièvement nous citerons les différentes extensions qui ont été faites sur ce modèle. Puis, nous étudierons les propriétés de ce modèle.

1.2 Définitions et Préliminaires

Définition 1.1. *Stationnarité stricte*

Un processus temporel aléatoire $(X_t)_{t \in \mathbb{Z}}$ est dit fortement stationnaire si $\forall t_1 < t_2 < \dots < t_n$ tel que $t_i \in \mathbb{Z}$ et pour tout $h \in \mathbb{Z}$ $t_i + h \in \mathbb{Z}$, $\forall i = 1, \dots, n$, la suite $(X_{t_1+h}, \dots, X_{t_n+h})$ à la même loi de probabilité que la suite $(X_{t_1}, \dots, X_{t_n})$.

Définition 1.2. *Stationnarité faible*

Un processus $(X_t)_{t \in \mathbb{Z}}$ est stationnaire au second ordre si :

- i) $E(X_t^2) < \infty \quad \forall t \in \mathbb{Z}$.
- ii) $E(X_t) = m$, indépendant de t , $\forall t \in \mathbb{Z}$.
- iii) $cov(X_t, X_{t+h}) = E[(X_{t+h} - m)(X_t - m)] = \gamma(h)$ indépendante de $t \quad \forall (t, h) \in \mathbb{Z}^2$.

Définition 1.3. Bruit Blanc

Le processus $(\varepsilon_t)_{t \in \mathbb{Z}}$ est appelé bruit blanc faible (processus hasard pure) s'il vérifie pour une constante positive σ^2 :

i) $E(\varepsilon_t) = 0, \forall t \in \mathbb{Z}$.

ii) $E(\varepsilon_t^2) = \sigma^2, \forall t \in \mathbb{Z}$

iii) $cov(\varepsilon_t, \varepsilon_s) = E(\varepsilon_t \varepsilon_s) = 0, \forall t, s \in \mathbb{Z} \text{ et } t \neq s$.

Si de plus les variables ε_t et ε_s sont indépendantes on parle alors de bruit blanc fort.

Définition 1.4. L'innovation

L'innovation d'un processus $(X_t)_{t \in \mathbb{Z}}$ est définie comme suit :

1-Les innovations représentent la différence entre la valeur prévue et celle réalisée.

$$\varepsilon_t = X_t - E(X_t | X_{t-1}, X_{t-2}, \dots)$$

2-Le carré des innovations représente la différence entre la valeur réalisée et la variance conditionnelle.

$$\varepsilon_t^2 = X_t^2 - V(X_t | I_{t-1})$$

3-Le carré des innovations normalisées est défini par :

$$\varepsilon_t^2 = \frac{X_t - E(X_t | I_{t-1})}{\sqrt{V(X_t | I_{t-1})}}$$

Définition 1.5. Kurtosis

Le kurtosis ou le coefficient d'aplatissement pour un échantillon X_1, \dots, X_T de taille T s'écrit :

$$K = \frac{\mu_4}{\mu_2^2}$$

avec $\mu_k = E(X_t - E[X_t])^k = \frac{1}{T} \sum_{t=1}^T (X_t - \bar{X})^k$.

Remarque 1.1. Le Kurtosis mesure le caractère pointu ou plat de la distribution de la série. Le Kurtosis de la distribution normale est 3. Si le Kurtosis est supérieur à 3 (queues

épaisses), la distribution est plutôt pointue (distribution leptokurtique); si le Kurtosis est inférieur à 3, la distribution est plutôt plate (distribution dite platikurtique).

Définition 1.6. Causalité

Un processus $(X_t)_{t \in \mathbb{Z}}$ est dit causal, s'il existe une suite de constantes ψ_j telle que :

$$X_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j} \quad \text{avec} \quad \sum_{j=0}^{\infty} |\psi_j| < \infty$$

avec $\{\varepsilon_t\}_t$ une suite de variables aléatoires de moyenne nulle et de variance constante.

Définition 1.7. Différence de martingale

Un processus $(X_t)_{t \in \mathbb{Z}}$ est une différence de martingale homoscédastique si et seulement si :

$$E(X_t | X_{t-1}, \dots, X_0) = 0$$

$$V(X_t) = \sigma^2, \quad \forall t \in \mathbb{Z}.$$

avec σ^2 est une constante.

Théorème 1.1. Théorème érgodique

Soit (X_t) un processus stationnaire ergodique.

Pour toute fonction f mesurable telle que $E[|f(\dots, X_{t-1}, X_t, X_{t+1}, \dots)|] < \infty$ alors

$$\frac{1}{n} \sum_{t=1}^n f(\dots, X_{t-1}, X_t, X_{t+1}, \dots) \xrightarrow{P.S} E[f(\dots, X_{t-1}, X_t, X_{t+1}, \dots)]$$

Définition 1.8. Produit de Kronecker

Soit $A = (a_{ij})$ et $B = (b_{ij})$ deux matrices de dimension respectivement $n \times p$ et $m \times q$, la matrice de dimension $mn \times pq$ défini par :

$$A \otimes B = \begin{pmatrix} a_{11}B & a_{12}B & \dots & a_{1p}B \\ a_{21}B & a_{22}B & \dots & a_{2p}B \\ \vdots & \vdots & & \vdots \\ a_{n1}B & a_{n2}B & \dots & a_{np}B \end{pmatrix}$$

est appelé produit de Kronecker des matrices A et B .

Propriété de Cramer-Wold

Pour une suite (Z_n) de vecteurs aléatoires de dimension d , $Z_n \xrightarrow{L} Z$ si et seulement si

$$\lambda' Z_n \xrightarrow{L} \lambda' Z \quad \forall \lambda \in \mathbb{R}^d$$

avec λ' est la transposée de λ

Définition 1.9. Coefficient de Lyapounov

Soit $\{A_t, t \in \mathbb{Z}\}$ une suite de matrices aléatoires de même ordre, strictement stationnaire et ergodique, telle que $E(\log^+ \|A_t\|)$ est fini. On a

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

γ sera donc appelé coefficient de Lyapounov de la suite de matrices $\{A_t, t \in \mathbb{Z}\}$. De plus

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

Propriété 1.1. Propriété des espérances itérées

Soit X une variable aléatoire, on a :

$$E(X/I_1) = E[E(X/I_2)/I_1]$$

avec : $I_{t-s} = \sigma\{\varepsilon_{t-s}, \varepsilon_{t-s-1}, \dots\} \quad \forall s < t \text{ et } I_1 \subset I_2$

Définition 1.10. Modèle AR(p)

$(X_t)_{t \in \mathbb{Z}}$ est un processus autorégressif d'ordre p , noté $AR(p)$ s'il s'écrit sous la forme suivante :

$$X_t = \sum_{j=1}^p \phi_j X_{t-j} + \varepsilon_t$$

où les ϕ_j sont des nombres réels et ε_t est un bruit blanc de moyenne nulle et de variance σ_ε^2 .

Définition 1.11. Modèle ARMA(p,q)

Un processus (X_t) est appelé processus $ARMA(p,q)$ (Auto Regressive Moving Average) où p et q sont des entiers, s'il existe des réels $a_1, \dots, a_p, b_1, \dots, b_q$ tel que :

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

où (ε_t) est le processus des innovations linéaires de (X_t) .

1.3 La famille des modèles ARCH

Les modèles classiques de prévision fondés sur les modèles ARMA supposent des séries temporelles à variance constante (hypothèse d'homoscédasticité). Cette modélisation néglige donc, éventuellement l'information contenue dans le facteur résiduel de la chronique. Les modèles de type ARCH (Autorégressive Conditional Hétéroscedasticity) permettent de modéliser des chroniques [la plupart du temps financières, car en effet, les séries financières sont particulièrement concernées par les modèles ARCH, car on constate des périodes de forte spéculation (variabilité élevée) suivi de périodes d'accalmie (variabilité faible)] qui ont une volatilité (variance) instantanée qui dépend du passé.

Il est ainsi possible d'élaborer une prévision dynamique de la chronique en terme de moyenne et de variance.

Présentés initialement par Engle (1982), ces modèles ont connu des développements et des applications très importants, par la suite.

Dans ce chapitre nous allons aborder les différentes classes de modèles ARCH, ainsi que leurs propriétés.

1.3.1 Représentation ARCH

Commençons par présenter le modèle ARCH(1) introduit par Engle (1982).

Définition 1.12. *Un processus $\{\varepsilon_t\}$ satisfait une représentation ARCH(1) si*

$$\varepsilon_t = \eta_t h_t$$

avec $h_t = \sqrt{\alpha_0 + \alpha_1 \varepsilon_{t-1}^2}$ et η_t désigne un bruit blanc faible tel que

$$E(\eta_t) = 0 \text{ et } E(\eta_t^2) = \sigma_\eta^2, \alpha_0 > 0, \alpha_1 \geq 0.$$

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

Théorème 1.2. *Si un processus $\{\varepsilon_t\}$ satisfait une représentation ARCH(1) , alors ε_t^2 satisfait une représentation AR (1) telle que :*

$$\varepsilon_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + v_t \quad (1.1)$$

où $v_t = \varepsilon_t^2 - h_t^2$ vérifiant $E[v_t/I_{t-1}] = 0$ est un processus d'innovation pour ε_t^2 et $\alpha_0 > 0$, $\alpha_1 \geq 0$

1.3.2 Propriétés des processus ARCH(p)

1.3.2.1 Les moments des modèles ARCH(p)

Considérons le cas d'un modèle ARCH(1) défini comme suit :

$$\varepsilon_t = \eta_t h_t$$

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

Théorème 1.3. *(Engle 1982)*

Pour tout entier r , les moments d'ordre $2r$ du processus ARCH(1) défini précédemment avec $\alpha_0 > 1$ et $\alpha_1 \geq 0$ existent pour tout r , si

$$\alpha_1^r \prod_{j=1}^r (2j - 1) < 1$$

Remarque

Les modèles ARCH(p) définis par $\varepsilon_t = \eta_t h_t$ avec $h_t = \sqrt{\alpha_0 + \sum_{i=1}^p \alpha_i \varepsilon_{t-i}^2}$, en tant que processus non linéaires ne sont pas complètement définis par leur structure d'ordre deux, donc une connaissance des moments d'ordre supérieur apparaît nécessaire. La distribution de ε_t étant conditionnellement symétrique autour de zéro, tous les moments d'ordre impairs sont nuls. Tous les produits $\varepsilon_{t_1}^{n_1} \dots \varepsilon_{t_m}^{n_m}$, $t_i \neq t_j$ ont une moyenne nulle dès que l'un des facteurs a un exposant nul, donc tous les cumulants d'ordre impair sont nuls. Il est alors possible d'obtenir une généralisation du théorème (1.3) pour l'existence des moments d'ordre $2r$ pour un processus ARCH(p). La construction est très technique, pour une approche précise on renvoie à Milhoj (1985).

Propriété 1.2. Soit $\varepsilon_t \sim ARCH(p)$, on peut noter que pour tout $s > 1$: $E(\varepsilon_t/I_{t-s}) = 0$, tel que $I_{t-1} = \sigma\{\varepsilon_{t-1}, \varepsilon_{t-2}, \dots, \varepsilon_{t-j}, \dots\}$

Preuve 1.1. Pour démontrer cela, utilisons la propriété des espérances itérées.

On a :

$$\begin{aligned} E[\varepsilon_t/I_{t-s}] &= E[E[\varepsilon_t/I_{t-1}]/I_{t-s}] \\ &= E[0/I_{t-s}] \\ &= 0 \quad \forall s > 1. \end{aligned} \tag{1.2}$$

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

Propriété 1.3. La variance conditionnelle du processus $\{\varepsilon_t\}$ satisfaisant une représentation ARCH (1), définie par l'équation

$$\begin{aligned} \varepsilon_t &= \eta_t h_t \\ h_t &= \sqrt{\alpha_0 + \alpha_1 \varepsilon_{t-1}^2} \quad \alpha_0 > 0 \text{ et } 0 < \alpha_1 < 1 \end{aligned}$$

est non constante dans le temps et vérifie :

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

Ceci est la caractéristique principale des processus ARCH(p), le processus (ε_t) a une variance conditionnelle qui dépend du temps. **Preuve**

On sait que $E[\varepsilon_t/I_{t-s}] = 0$ dès lors, $V[\varepsilon_t/I_{t-s}] = E[\varepsilon_t^2/I_{t-s}]$. Considérons le processus ε_t^2 défini par la relation (1.1) où v_t est un bruit blanc faible. Par itération successive, on a :

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

En considérant l'espérance conditionnelle de chaque côté, on obtient :

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

Par définition du bruit blanc, on a :

$E[v_{t-j}/I_{t-s}] = 0, \forall j = 1, 2, \dots, s-1$ et par définition $E[\varepsilon_{t-s}^2/I_{t-s}] = \varepsilon_{t-s}^2$, on obtient ainsi la formule de la variance.

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

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

$$\begin{aligned} V[\varepsilon_t] &= \lim_{s \rightarrow \infty} V[\varepsilon_t / I_{t-s}] \\ &= \lim_{s \rightarrow \infty} \left[\alpha_0 \left[\frac{1 - \alpha_1^s}{1 - \alpha_1} \right] + \alpha_1^s \varepsilon_{t-s}^2 \right] \\ &= \frac{\alpha_0}{1 - \alpha_1} \quad \text{car } \alpha_0 > 0 \text{ et } 0 < \alpha_1 < 1 \end{aligned}$$

d'où les contraintes sur les paramètres de la représentation ARCH à savoir $\alpha_0 > 0, 0 < \alpha_1 < 1$

Propriété 1.4. Les autocovariances conditionnelles du processus (ε_t) , suivant un modèle ARCH (1), définies par l'équation

$$\begin{aligned} \varepsilon_t &= \eta_t h_t \\ h_t &= \sqrt{\alpha_0 + \alpha_1 \varepsilon_{t-1}^2} \end{aligned}$$

sont nulles

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

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

Preuve 1.2. Cette propriété s'obtient de la façon suivante :

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

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

Propriété 1.5. .

1. Le moment conditionnel centré d'ordre quatre du processus (ε_t) suivant un modèle ARCH(1) vérifie

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

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

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

3. Le kurtosis non conditionnelle associée au processus ARCH(1) est

$$K = \frac{E[\varepsilon_t^4]}{E^2[\varepsilon_t^2]} = 3 \left[\frac{1-\alpha_1^2}{1-3\alpha_1^2} \right] > 3$$

Preuve 1.3. 1. On rappelle que si une variable centrée X suit une loi normale centrée, alors

$$E(X^4) = 3(V(X))^2 = 3(E(X^2))^2 \quad (\star\star)$$

donc

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

2. Sous l'hypothèse $3\alpha_1^2 < 1$ on a :

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

3. Le kurtosis non conditionnel associé au processus ARCH(1) est définie par :

$$K = \frac{E[\varepsilon_t^4]}{E^2[\varepsilon_t^2]}$$

et d'après les résultats obtenus précédemment on obtient :

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

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

1.3.3 Modèle avec erreur ARCH

1.3.3.1 Modèle ARCH(p)

Définition 1.13. Un processus ε_t satisfait une représentation ARCH(p) si $\varepsilon_t = \eta_t h_t$ avec

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

où η_t est un bruit blanc faible, d'espérance nulle et de variance σ_η^2 .

Pour ce type de processus, on retrouve les deux propriétés essentielles vues précédemment, à savoir la propriété de différence de martingale (ou bruit blanc faible) $E(\varepsilon_t/I_{t-s}) = 0$ et la propriété de variance conditionnelle variable dans le temps puisque :

$$V[\varepsilon_t/I_{t-1}] = h_t^2 = \alpha_0 + \sum_{i=1}^p \alpha_i \varepsilon_{t-i}^2$$

de plus,

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

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

Dans ce qui suit, on ne considère plus un processus ARCH pour modéliser directement la série financière, mais le résidu d'un modèle linéaire. Prenons l'exemple d'un modèle linéaire auto-régressif avec résidus de type ARCH(p).

1.3.3.2 Modèle Autoregressif avec erreur ARCH(p)

$(X_t)_{t \in \mathbb{Z}}$ est un processus autoregressif avec erreur ARCH(P), s'il s'écrit sous la forme suivante

Définition 1.14. On considère un modèle linéaire auto-régressif de la forme

$$X_t = \delta + \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + \varepsilon_t$$

ε_t est un bruit blanc faible, tel que $E(\varepsilon_t) = 0$ et $E(\varepsilon_t \varepsilon_s) = 0$ si $s \neq t$, satisfaisant la condition de différence de martingale $E(\varepsilon_t / \varepsilon_{t-1}) = 0$. On suppose que ce résidu admet une représentation de type ARCH(p) :

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

où η_t est un bruit blanc faible

et $\alpha_0 > 0$, $\alpha_i \geq 0$

On a donc un modèle qui décrit à la fois l'évolution de l'espérance conditionnelle et de la variance conditionnelle du processus (X_t) dans le temps. Envisageons le cas le plus simple d'un processus de type AR(1) avec erreur ARCH(1) :

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

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

Les résidus satisferont alors les principales propriétés étudiées précédemment.

1. Le processus (ε_t) est orthogonal aux valeurs passées, quelque soit le retard

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

et suit un processus ARCH(1)

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

dont la variance conditionnelle est

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

et la variance non conditionnelle est

$$V[\varepsilon_t] = \frac{\alpha_0}{1 - \alpha_1}$$

2. La propriété d'orthogonalité implique que les corrélations conditionnelles sont nulles :
 $cov[\varepsilon_t \varepsilon_{t+k} / I_{t-s}] = 0$,

Il y a donc absence de corrélation entre les valeurs présentes et futures du processus, quels que soient les retards s et k , mais si la variance conditionnelle de (ε_t) n'est pas constante, la variance non conditionnelle est constante.

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

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

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

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

En effet

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

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

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

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

Propriété 1.6. *La variance conditionnelle du processus AR (1) avec erreur ARCH (1), X_t , s'écrit*

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

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

$$V[X_t/I_{t-1}] = \delta + \alpha_1 \varepsilon_{t-1}^2$$

Preuve 1.4.

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

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

1.3.4 Les modèles ARCH généralisés : GARCH

Pour palier aux inconvénients des modèles ARCH, Bollerslev (1986) introduit le modèle GARCH (Generalized Autoregressive Conditional Heteroskedasticity), une généralisation du modèle ARCH dans laquelle sont ajoutées p retards de la variance conditionnelle h_t . Ce modèle permet une plus grande flexibilité dans la structure des retards. Il s'agit d'une extension du processus ARCH, et constitue un modèle ARMA de la variance conditionnelle alors que le processus ARCH est un modèle AR de cette même variance. Toutefois, Engle (1995) observe que le modèle ARCH est une moyenne mobile puisque la variance conditionnelle est moyenne mobile des résidus au carré.

1.3.4.1 Modèle GARCH(p,q)

Définition 1.15. Un processus (ε_t) satisfait une représentation GARCH(p,q) si

$$\varepsilon_t = \eta_t h_t \quad (1.3)$$

$$h_t = \sqrt{\alpha_0 + \sum_{i=1}^p \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^q \beta_j h_{t-j}^2} \quad (1.4)$$

où η_t est un bruit blanc faible, et où $\alpha_0 > 0, \alpha_i > 0$, pour $i = 1, 2, \dots, p$ et $\beta_j > 0$ pour $j = 1, 2, \dots, q$.

1.3.4.2 Propriétés des processus GARCH (p,q)

Les propriétés des processus GARCH se déduisent de la même façon que celle que nous avons traité pour les processus ARCH.

Propriété 1.7. Une condition nécessaire de l'existence de la variance d'un processus GARCH (p,q) défini par les équations (1.3) et (1.4) est

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

Remarque 1.2. Si la condition \star est vérifiée en plus des contraintes de non négativité données précédemment, elle sera également suffisante pour dire que le processus GARCH est faiblement stationnaire ou stationnaire au second ordre.

Dans le cas où $\sum_{i=1}^p \alpha_i + \sum_{j=1}^q \beta_j = 1$, on dira alors que le processus GARCH(p,q) est intégré, et on parlera de processus **IGARCH**.

Propriété 1.8. Le processus (ε_t^2) satisfaisant une représentation GARCH(p,q) peut être représenté sous la forme d'un processus ARMA (max(p,q),q) défini $v_t = \varepsilon_t^2 - h_t^2$ tel que :

$$\varepsilon_t^2 = \sum_{i=1}^n (\alpha_i + \beta_i) \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j v_{t-j} + v_t \quad n = \max(p, q)$$

avec la convention $\alpha_i = 0$ si $i > p$ et $\beta_j = 0$ si $j > q$.

1.3.4.3 Modèle avec Erreur GARCH(1,1)

Pour modéliser les données empiriques sur le marché financier, un modèle de régression avec erreur GARCH (1,1) est souvent suffisant. Il est donné par l'équation

$$Y_t = cX_t + \varepsilon_t$$

avec ε_t un GARCH(1,1) qui vérifie $\varepsilon_t = \eta_t h_t$ et $h_t = \sqrt{\alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1}^2}$ avec $\alpha_0 > 0$, $\alpha_1 > 0$ et $\beta_1 > 0$.

Dans ce modèle, les carrés des résidus suivent un processus ARMA (1,1)

$$\varepsilon_t^2 = \alpha_0 + (\alpha_1 + \beta_1)\varepsilon_{t-1}^2 + \beta_1 v_{t-1} + v_t$$

ε_t^2 est stationnaire pour $0 < \alpha_1 + \beta_1 < 1$, où $v_t = \varepsilon_t^2 - h_t^2$ est un processus d'innovation pour ε_t . Sous la condition de stationnarité de second ordre, la variance inconditionnelle du processus ε_t existe et est constante. Sachant que $V(\varepsilon_t) = E(\varepsilon_t^2)$ il suffit à partir de la forme ARMA (1,1) sur ε_t de définir la variance du processus : $V(\varepsilon_t) = \frac{\alpha_0}{1 - (\alpha_1 + \beta_1)}$.

Selon Bollerslev (1986), le kurtosis existe si

$$3\alpha_1^2 + 2\alpha_1\beta_1 + \beta_1^2 < 1$$

et est donnée par :

$$\begin{aligned} K &= \frac{E[\varepsilon_t^4]}{E^2[\varepsilon_t^2]} \\ &= 3 \frac{1 - (\alpha_1 + \beta_1)^2}{1 - (\alpha_1 + \beta_1)^2 - 2\alpha_1^2} \end{aligned}$$

Il est toujours supérieur à trois. Ainsi, si α_1 tend vers zéro, l'hétéroscédasticité disparaît et la valeur du kurtosis tend vers trois. Enfin, on peut montrer que pour un processus GARCH, le kurtosis est directement lié à l'hétéroscédasticité conditionnelle.

Considérons le cas du kurtosis associé à la loi non conditionnelle dans un processus ε_t suivant un GARCH(p,q) conditionnellement gaussien tel que $\eta_t \sim N(0, 1)$. Dans ce cas, les moments conditionnels d'ordre 2 et 4 du processus ε_t sont liés :

$$E[\varepsilon_t^4 / I_{t-1}] = 3[E(\varepsilon_t^2 / I_{t-1})]^2$$

Si on applique l'espérance sur les deux côtés de l'équation ($\star\star$), il devient

$$\begin{aligned} E(\varepsilon_t^4) &= E(E[\varepsilon_t^4/I_{t-1}]) = 3E([E(\varepsilon_t^2)]^2) \\ &\geq 3E([E(\varepsilon_t^2)]^2) \\ &= 3E^2(\varepsilon_t^2) \end{aligned}$$

On peut calculer le kurtosis comme suit :

$$\begin{aligned} K &= \frac{E[\varepsilon_t^4]}{E^2[\varepsilon_t^2]} \\ &= \frac{3E([E(\varepsilon_t^2/I_{t-1})]^2)}{E^2[\varepsilon_t^2]} \\ &= 3\frac{E^2[\varepsilon_t^2]}{E^2[\varepsilon_t^2]} + \frac{3}{E^2[\varepsilon_t^2]}(E([E(\varepsilon_t^2/I_{t-1})]^2) - E^2[\varepsilon_t^2]) \\ &= 3 + \frac{3}{E^2[\varepsilon_t^2]}(E([E(\varepsilon_t^2/I_{t-1})]^2) - E^2[E(\varepsilon_t^2/I_{t-1})]) \\ &= 3 + 3\frac{V[E(\varepsilon_t^2/I_{t-1})]}{E^2[\varepsilon_t^2]} > 3 \end{aligned}$$

Le kurtosis est donc liée à une mesure de l'hétéroscédasticité conditionnelle.

1.4 Extension des modèles ARCH

1.4.1 Modèle ARCH-M

Le modèle ARCH-M est un modèle dérivé du modèle ARCH, il permet à la variance d'affecter la moyenne. Ce modèle a été introduit par Engle et al (1987).

Définition 1.16. *On dit qu'un processus $\{X_t, t \in \mathbb{Z}\}$ suit un modèle de régression ARCH-M s'il est défini par les équations suivantes*

$$X_t = \beta'Y_t + \delta h_t + \varepsilon_t$$

où

$$Var[\varepsilon_t|I_{t-1}] = h_t^2$$

avec

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

1.4.2 Modèle β -ARCH

Définition 1.17. On appelle processus β -ARCH, un processus $\{X_t, t \in Z\}$, défini par les relations

$$X_t = T(X_{t-1}) + h(X_{t-1})\varepsilon_t$$

où

$$T(x) = \alpha x$$

et

$$h(x) = [\alpha_0 + (\alpha_1^+ x^+ + \alpha_1^- x^-)^{2\beta}]^{1/2}$$

Où $\alpha^+ = \max(\alpha, 0)$, $\alpha^- = \max(-\alpha, 0)$ où $\beta_0, \alpha_0, \alpha_1^+, \alpha_1^-$ sont des constantes positives, et où ε_t est un bruit blanc non supposé gaussien.

Remarque 1.3. 1. Si $\alpha = 0$ on a un pur modèle β -ARCH. Dans le cas où $\beta = 1$, on retrouve le modèle AR(1) avec bruit ARCH(1).

2. Ce modèle est important car il va permettre de prendre en compte une grande variabilité des données, de modéliser des distributions marginales plus ou moins leptokurtiques en fonction de β .

1.4.3 Modèle TARARCH

Engle et Bollerslev (1986) ont étudié les modèles ARCH à seuil (Threshold ARCH)(TARCH) où la variance est une fonction linéaire définie par morceaux qui permet différentes fonctions de volatilité selon le signe et la valeur des chocs.

Définition 1.18. On appelle modèle TARARCH, le processus (X_t) défini par

$$X_t = \alpha X_{t-1} + (\alpha_0 + \alpha_1^+(X_{t-1})^+ + \alpha_1^-(X_{t-1})^-)\varepsilon_t$$

Où $\alpha^+ = \max(\alpha, 0)$ et $\alpha^- = \max(-\alpha, 0)$ et où ε_t est un bruit blanc gaussien.

Remarque 1.4. Une étude plus détaillée se trouve dans Zakoïan (1990)

1.4.4 Modèle GARCH exponentiel EGARCH

Nelson(1990) s'est intéressé aux évolutions asymétrique de la variance à l'aide des modèles EGARCH.

Définition 1.19. Soit $(z_t)_{t \in \mathbb{Z}}$ une suite de variables iid telles que $E(z_t) = 0$ et $Var(z_t) = 1$. On dit que (ε_t) est un processus GARCH exponentiel s'il vérifie une équation de la forme

$$\begin{cases} \varepsilon_t = z_t h_t, \\ \log(h_t^2) = \omega + \sum_{i=1}^q \alpha_i g(z_{t-i}) + \sum_{j=1}^p \beta_j \log(h_{t-j}^2), \end{cases}$$

Où

$$g(z_{t-i}) = \theta z_{t-i} + \gamma[|z_{t-i}| - E|z_{t-i}|]$$

$\omega, \alpha_i, \beta_j, \theta$ et γ sont des réels.

1.4.5 Modèle IGARCH

Introduits par Engle et Bollerslev (1987), les modèles IGARCH correspondent au cas d'une racine unitaire dans le processus de variance conditionnelle et sont caractérisés par un effet de persistance dans la variance.

Définition 1.20. Un processus $\{\varepsilon_t\}_{t \in \mathbb{Z}}$ satisfait une représentation IGARCH(p, q) si et seulement si :

$$V(\varepsilon_t | I_{t-1}) = h_t = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^p \beta_i h_{t-i}$$

avec : $\alpha_0 > 0, \alpha_i \geq 0 \quad \forall i = 1, \dots, q$ et $\beta_j \geq 0 \quad \forall j = 1, \dots, p$ et $\sum_{i=1}^p \alpha_i + \sum_{j=1}^q \beta_j = 1$

Remarque 1.5. Persistance sur la variance signifie qu'un choc sur la variance conditionnelle actuelle se répercute sur toutes les valeurs futures prévues, ie : choc infiniment persistant sur h_t .

Chapitre 2

Estimation et tests dans les modèles hétéroscédastiques

2.1 Introduction

Dans ce chapitre, nous allons traiter l'estimation des paramètres d'un modèle (G)ARCH, et plus particulièrement, d'un modèle de régression avec erreur (G)ARCH. Les modèles introduits reposent sur des formulations des moyennes et variances conditionnelles. En pratique, celles-ci sont souvent paramétrées de telle façon que la moyenne conditionnelle $m_t(\theta)$ et la variance conditionnelle $h_t^2(\theta)$ apparaissent comme des fonctions de paramètres inconnus et de valeurs passées du processus. La connaissance de ces moments ne suffit cependant pas sans hypothèse supplémentaire à caractériser la loi conditionnelle du processus. La vraisemblance est écrite en considérant que η_t suit une loi normale centrée réduite (on parle de pseudo ou quasi-vraisemblance), mais cette hypothèse n'est pas nécessaire pour la convergence forte de l'estimateur. Elle a évidemment un effet sur la variance de la loi asymptotiquement normale de l'estimateur. Les modèles avec erreurs hétéroscédastiques peuvent être estimés généralement par ces méthodes :

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

Nous présenterons ici ces deux méthodes d'estimation. Commençons par étudier la méthode des moindres carrés.

Nous présenterons aussi deux statistiques de tests, qui permettent de détecter les modèles ARCH en présence de la corrélation en série.

2.2 Méthode d'estimation des moindres carrés ordinaire

Soit un processus $\{\varepsilon_t\}_{t \in \mathbb{Z}}$ suivant un modèle ARCH(q) défini par les équations suivantes :

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

(η_t) est une suite de variables aléatoires *i.i.d*, $E(\eta_t) = 0$, $\text{Var}(\eta_t) = 1$.

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

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

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

où $\mu_t = \varepsilon_t^2 - h_t^2 = (\eta_t^2 - 1)h_t^2$. La suite $(\mu_t; I_{t-1})$, constitue une différence de martingale.

avec $I_{t-1} = \sigma \varepsilon_{t-1}, \varepsilon_{t-2}, \dots, \varepsilon_{t-j}, \dots$

On suppose que l'on dispose d'observations $\varepsilon_0, \dots, \varepsilon_n$ avec $n \in \mathbb{N}$, réalisations partielles du processus (ε_t) , et de valeurs initiales $\varepsilon_0, \dots, \varepsilon_{t-q}$.

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

Soit le vecteur

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

On déduit le système

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

Soit $Y = X\theta_0 + U$

en définissant la matrice X d'ordre $n \times (q+1)$ et les vecteurs Y et U d'ordre $n \times 1$

$$X = \begin{bmatrix} Z'_{n-1} \\ \vdots \\ Z'_0 \end{bmatrix}, Y = \begin{bmatrix} \varepsilon_n^2 \\ \vdots \\ \varepsilon_1^2 \end{bmatrix}, U = \begin{bmatrix} \mu_n \\ \vdots \\ \mu_1 \end{bmatrix}$$

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

$$\hat{\theta}_n = (X'X)^{-1} X'Y = \arg \min_{\theta} \|Y - X\theta\|^2 \quad (2.4)$$

L'estimateur des MCO de $h_0^2 = \text{Var}(\mu_t)$ est :

$$\hat{h}^2 = \frac{1}{n-q-1} \|Y - X\hat{\theta}_n\|^2 = \frac{1}{n-q-1} \sum_{t=1}^n \left\{ \varepsilon_t^2 - \hat{\alpha}_0 - \sum_{i=1}^q \hat{\alpha}_i \varepsilon_{t-i}^2 \right\}^2$$

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

H1 : $\{\varepsilon_t\}$ est une solution non anticipative strictement stationnaire (causal) du modèle (2.1).

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

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

Théorème 2.1. (convergence des estimateurs MCO pour un ARCH(p))

Soit $(\hat{\theta}_n)_{n \in \mathbb{N}}$ une suite d'estimateurs satisfaisant (2.4). Sous les hypothèses H1-H3, presque sûrement, on a

$$\hat{\theta}_n \xrightarrow{P.S} \theta_0, \hat{h}_t^2 \xrightarrow{P.S} h_t^2, \text{ quand } n \rightarrow \infty.$$

Preuve

- **étape 1** : Nous avons vu que l'unique solution stationnaire non anticipative ε_t est ergodique. Le processus $(Z_t)_{t \in \mathbb{Z}}$ est également ergodique car Z_t s'écrit comme fonction mesurable des ε_{t-i} .

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

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

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

- **étape 2** : Montrons par l'absurde l'inversibilité de la matrice $E(Z_{t-1} Z'_{t-1}) = E(Z_t Z'_t)$ supposons qu'il existe un vecteur non nul $c \in \mathbb{R}^{q+1}$ tel que

$$c' E(Z_t Z'_t) = 0$$

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

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

Donc (η_t) s'exprime p.s comme fonction mesurable des variables $\varepsilon_{t-1}, \varepsilon_{t-2}, \dots, \varepsilon_{t-q}$, et comme (η_t) est indépendante de ces variables, donc (η_t^2) est égale à une constante p.s.

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

- **étape 3** : Il découle de ce qui précède que $\frac{1}{n} X'X$ est p.s inversible, pour n assez grand et que p.s quand $n \rightarrow \infty$,

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

- **étape 4** : Le processus (μ_t) est l'innovation de ε_t^2 . On a donc, en particulier, les relations d'orthogonalité

$$E(\mu_t) = E(\mu_t \varepsilon_{t-1}^2) = \dots = E(\mu_t \varepsilon_{t-q}^2) = 0$$

c'est-à-dire

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

d'où

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

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

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

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

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

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

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

L'inversibilité de la matrice A se démontre de la même manière que précédemment (convergence des estimateurs MCO pour un ARCH). Celle de I sera montrée dans la preuve du résultat suivant, qui établit la normalité asymptotique de l'estimateur des MCO.

Théorème 2.2. *Sous les hypothèses H1-H4*

$$\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{D} N(0, (\mu_4 - 1)A^{-1}IA^{-1}) \quad \text{avec } \mu_4 = E(Z_t^4)$$

Preuve : On a :

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

Donc

$$\sqrt{n}(\hat{\theta}_n - \theta_0) = \left(\frac{1}{n} \sum_{t=1}^n Z_{t-1} Z'_{t-1} \right)^{-1} \left\{ \frac{1}{\sqrt{n}} \sum_{t=1}^n Z_{t-1} \mu_t \right\} \quad (2.6)$$

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

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

On déduit que, pour tout $\lambda \neq 0$ (par application du théorème central limite)

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

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

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

On montre que cette loi limite est non dégénérée, c'est-à-dire que I est inversible, par le même raisonnement que celui utilisé pour établir l'inversibilité de A dans la preuve du Théorème (2.1). Par conséquent, on déduit de (2.5), (2.6) et (2.7), par un raisonnement classique, que $\sqrt{n}(\hat{\theta}_n - \theta_0)$ est asymptotiquement normal, de moyenne le vecteur nul, et de variance la matrice du théorème(2.2).

2.3 Méthode d'estimation du maximum de vraisemblance sur un GARCH(p,q)

Dans cette partie, nous étudions la méthode du maximum de vraisemblance conditionnelle (à des valeurs initiales). Nous présentons de façon parallèle la méthode d'estimation du maximum de vraisemblance (MV) sous l'hypothèse de normalité de la distribution conditionnelle des résidus et la méthode d'estimation du quasi maximum de vraisemblance

(QMV). En effet, l'idée générale des estimateurs du pseudo maximum de vraisemblance (PMV) consiste à démontrer que si l'on commet une erreur sur la distribution conditionnelle des résidus en utilisant à tort une log-vraisemblance fondée sur une loi normale, l'estimateur du MV ainsi obtenu peut tout de même être convergent si la loi des résidus est une loi normale (Gouriéroux, Monfort, 1989). L'estimateur sera non seulement asymptotiquement convergent mais aussi asymptotiquement normal. Par conséquent, la fonction de vraisemblance définissant l'estimateur du MV sous l'hypothèse de normalité et la fonction de quasi-vraisemblance de l'estimateur du QMV sont les mêmes.

2.3.1 Quasi-vraisemblance conditionnelle

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

où encore

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

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

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

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

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

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

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

2.3.1.1 Equations de vraisemblance

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

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

Pour un n assez grand, ces équations s'interprètent comme des relations d'orthogonalité. Le terme de gauche de l'égalité précédente se comporte asymptotiquement comme

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

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

$$E_{\theta_0} \left(\mu_t \frac{1}{h_t^4(\theta_0)} \frac{\partial h_t^2(\theta_0)}{\partial \theta} \right) = 0 \quad (2.20)$$

Car $\frac{1}{h_t^4(\theta_0)} \frac{\partial h_t^2(\theta_0)}{\partial \theta}$ est une fonction mesurable des ε_{t-i} , $\forall i > 0$. Ce résultat n'est autre que la version asymptotique de (2.19) en θ_0 , en utilisant le théorème ergodique.

2.3.1.2 Propriétés asymptotiques de l'estimateur du QMV

Dans ce qui suit, nous utilisons comme norme d'une matrice $A = (a_{ij})$ avec i, j quelconque la norme $\|A\| = \sum_{i=1}^n \sum_{j=1}^n |a_{ij}|$. Le rayon spectral d'une matrice A carrée sera noté $\rho(A)$. Le produit de Kronecker sera noté \otimes .

Convergence forte

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

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

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

Notons

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

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

Pour la convergence, on pose les hypothèses suivantes :

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

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

A3 : η_t^2 a une loi non dégénérée.

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

Il est pratique d'approximer la suite $(\tilde{l}_t(\theta))$ par rapport à t par une suite stationnaire ergodique. La condition de stricte stationnarité (**A2**) implique que les racines de $\mathcal{B}_\theta(\eta)$ sont extérieures au disque unité. Notons donc $(h_t^2)_t = \{h_t^2(\theta)\}_t$ la solution strictement stationnaire ergodique et non anticipative de

$$h_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i} + \sum_{j=1}^p \beta_j h_{t-j}^2, \quad \forall t \in \mathbb{Z} \quad (2.21)$$

et soit

$$I_n(\theta) = I_n(\theta; \varepsilon_n, \varepsilon_{n-1}, \dots) = n^{-1} \sum_{t=1}^n \tilde{l}_t, \text{ où } \tilde{l}_t(\theta) = \frac{\varepsilon_t^2}{h_t^2} + \ln(h_t^2)$$

Théorème 2.3. (Convergence forte de l'estimateur du QMV)(Zakoïan 1991)

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

$$\hat{\theta}_n \xrightarrow{P.S} \theta_0, \text{ quand } n \rightarrow \infty$$

Remarques

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

2- L'hypothèse (**A4**) disparaît dans le cas d'un modèle ARCH. Elle permet de sur-estimer l'un des ordres, p ou q , mais pas les deux.

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

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

Normalité asymptotique

Pour démontrer la normalité asymptotique, les hypothèses supplémentaires suivantes sont nécessaires.

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

A6 : $k_\eta = E(\eta_t^4) < \infty$.

Théorème 2.4. (Normalité asymptotique des estimateurs du QMV)

Sous les hypothèses **(A1)**-**(A6)**,

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{D} N(0, (k_\eta - 1)J^{-1})$$

où

$$J = E_{\theta_0} \left(\frac{\partial^2 l_t(\theta_0)}{\partial \theta \partial \theta'} \right) = E_{\theta_0} \left(\frac{1}{h_t^4(\theta_0)} \frac{\partial h_t^2(\theta_0)}{\partial \theta} \frac{\partial h_t^2(\theta_0)}{\partial \theta'} \right) \quad (2.22)$$

Remarques

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

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

Exemple appliqué sur un modèle ARCH(1)

Soit (ε_t) un modèle ARCH(1) défini par les équations suivantes

$$\varepsilon_t = \eta_t h_t, \quad t \in \mathbb{Z}$$

$$h_t^2 = w + \alpha \varepsilon_{t-1}^2 \quad \text{avec } w > 0, \quad \alpha \geq 0$$

Lorsque $q = 1$, les conditions de moments **H2** et **H4** prennent la forme $\mu_4 \alpha^2 < 1$ et $\mu_8 \alpha^4 < 1$. On a

$$A = \begin{pmatrix} 1 & E_{\theta_0} \varepsilon_{t-1}^2 \\ E_{\theta_0} \varepsilon_{t-1}^2 & E_{\theta_0} \varepsilon_{t-1}^4 \end{pmatrix}, \quad I = \begin{pmatrix} E_{\theta_0} h_t^4 & E_{\theta_0} h_t^4 \varepsilon_{t-1}^2 \\ E_{\theta_0} h_t^4 \varepsilon_{t-1}^2 & E_{\theta_0} h_t^4 \varepsilon_{t-1}^4 \end{pmatrix}$$

On a

$$E_{\theta_0} \varepsilon_t^2 = \frac{\omega}{1 - \alpha}, \quad E_{\theta_0} \varepsilon_t^4 = \mu_4 E_{\theta_0} h_t^4 = \frac{\omega^2(1 + \alpha)}{(1 - \mu_4 \alpha^2)(1 - \alpha)} \mu_4.$$

Les autres termes de la matrice I s'obtiennent en utilisant $h_t^4 = (\omega + \alpha \varepsilon_{t-1}^2)^2$ et en calculant les moments d'ordre 6 et 8 de ε_t^2 .

A titre d'exemple, le tableau suivant présente quelques valeurs numériques de la variance asymptotique, pour différentes valeurs des paramètres, lorsque η_t suit la loi $N(0, 1)$. On a $\mu_4 = 3$, $\mu_8 = 105$ et l'hypothèse **H4** prend la forme $\alpha < 105^{-\frac{1}{4}} = 0.312$.

Paramètres	$\omega = 1, \alpha = 0.1$	$\omega = 1, \alpha = 0.2$	$\omega = 1, \alpha = 0.3$
$Var_{as}\{\sqrt{n}(\hat{\theta}_n - \theta_0)\}$	$\begin{pmatrix} 3.98 & -1.85 \\ -1.85 & 2.15 \end{pmatrix}$	$\begin{pmatrix} 8.03 & -5.26 \\ -5.26 & 5.46 \end{pmatrix}$	$\begin{pmatrix} 151.0 & -106.5 \\ -106.5 & 77.6 \end{pmatrix}$

La précision asymptotique du paramètre ω , et globalement celle du couple θ , devient très faible au voisinage de la frontière d'existence du moment d'ordre 8. La méthode peut cependant être utilisée pour estimer α , même pour des valeurs plus élevées de ce paramètre (l'estimateur est convergent pour $\alpha < 3^{-1/2} = 0.577$), et fournir ainsi une valeur initiale dans l'algorithme de maximisation de la vraisemblance.

2.4 Test sur les modèles GARCH(p)

Comme nous l'avons souligné précédemment, la maximisation de la vraisemblance s'avère délicate, compte tenu de l'augmentation du nombre de dérivées à calculer. L'estimation des modèles ARCH-GARCH, requiert dans un premier temps la présence d'hétéroscédasticité conditionnelle dans les aléas du modèle spécifié. Pour ce faire, il existe plusieurs tests prévus à cet effet, nous en citerons deux des plus utilisés.

Test contre l'hétéroscédasticité de type-ARCH

Engle(1982) a proposé le test du Multiplicateur de Lagrange basé sur le modèle suivant :

$$y_t = x_t' \beta + u_t, \quad t = 1, \dots, n$$

$$u_t = \sigma_t \varepsilon_t$$

$$u_t / \psi_{t-1} \sim N(0, \sigma_t^2)$$

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

Où ψ_{t-1} est l'ensemble d'informations disponibles jusqu'à l'instant $t - 1$.

L'hypothèse d'homoscédasticité peut être formulée de la manière suivante :

$$H_0 : \alpha_1 = \alpha_2 = \dots = \alpha_q = 0$$

La statistique de test d'Engle est donnée par nR^2 , où n est la taille de l'échantillon, R^2 est le coefficient de détermination dans la régression des carrés des résidus \hat{u}_t des moindres carrés ordinaires sur une constante et \hat{u}_{t-i}^2 ($i=1, \dots, q$). Sous l'hypothèse nulle, la distribution asymptotique de cette statistique est la loi du chi-deux à q degrés de liberté.

La statistique de test de Lee et King (1993)

Lee (1991) a montré que le même test précédent (test contre l'hétéroscédasticité de type-ARCH) est approprié, mais cette fois ci contre les alternatives GARCH(p, q) c'est à dire

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^p \theta_i \sigma_{t-i}^2 + \sum_{i=1}^q \alpha_i u_{t-i}^2$$

Et l'hypothèse nulle serait alors

$$H_0 : \alpha_1 = \alpha_2 = \dots = \alpha_q = \theta_1 = \dots = \theta_p = 0$$

Lee et King (1993) ont proposé un test alternatif (G)ARCH qui exploite la nature unilatérale de $H_A : \sigma_t^2 = \sigma_s^2$.

La statistique de test est

$$LK = \frac{\left\{ (n - q) \sum_{t=q+1}^n [(\hat{u}_t^2 / \hat{\sigma}^2 - 1)] \sum_{i=1}^q \hat{u}_{t-i}^2 \right\}}{\left\{ \sum_{t=q+1}^n (\hat{u}_t^2 / \hat{\sigma}^2 - 1)^2 \right\}^{1/2}} \frac{1}{\left\{ (n - q) \sum_{t=q+1}^n \left(\sum_{i=1}^q \hat{u}_{t-i}^2 \right)^2 - \left(\sum_{t=q+1}^n \left(\sum_{i=1}^q \hat{u}_{t-i}^2 \right) \right)^2 \right\}^{1/2}}$$

et sa distribution asymptotique sous l'hypothèse nulle est normale.

Chapitre 3

Estimation par la méthode en deux étapes et application

3.1 Introduction

Nous avons vu dans le chapitre précédent l'estimation des paramètres d'un modèle ARCH en utilisant l'estimation par MCO et par maximum de vraisemblance ou quasi-maximum de vraisemblance. Cette dernière est la plus utilisée du fait que l'estimateur du maximum de vraisemblance obtenu après maximisation de la fonction log-vraisemblance est consistant et asymptotiquement normal, ce résultat a été établi par Weiss(1986). Il est également connu pour être efficace lorsque la distribution des erreurs est gaussienne.

Cependant, l'estimateur du quasi-maximum de vraisemblance n'admet pas de formule analytique il n'est donc pas facile à obtenir, la discussion de ce problème peut être trouvée dans Shepard (1996).

A cet effet, Bose et Mukherjee (2003), nous proposent un estimateur pour les paramètres du modèle ARCH issu de la méthode en deux étapes, dont la forme analytique est facile à calculer et qui en même temps contient les mêmes propriétés que l'estimateur quasi-maximum de vraisemblance.

3.2 Description de la méthode en deux étapes

Soit le modèle ARCH(p) suivant proposé par Bose et Mukherjee (2003) :

$$X_t = \varepsilon_t h_{t-1}(\beta) \quad 1 \leq t \leq n \quad (3.1)$$

tel que :

$\beta' = (\beta_0, \beta_1, \dots, \beta_p)$ sont les paramètres inconnus avec $\beta_0 > 0$ et $\beta_j \geq 0$, ($1 \leq j \leq p$).

$$h_{t-1}(\beta) = \sqrt{\beta_0 + \beta_1 X_{t-1}^2 + \dots + \beta_p X_{t-p}^2}.$$

Les variables ε_t ($1 \leq t \leq n$) sont *iid* de moyenne zéro, de variance 1 avec les moments d'ordre inférieur ou égal à 4 finis et indépendant de X_i .

Il est supposé que le processus est stationnaire et ergodique. Nelson (1990) montre que pour $p = 1$, la condition suffisante de stationnarité est

$$E \{ \log(\beta_1 \varepsilon_1^2) \} < 0$$

Pour p quelconque, la condition suffisante de stationnarité et d'ergodicité est donnée par Bougerol et Picard (1992, Théorème 1.3).

3.2.1 Construction de l'estimateur

Posons

$$\begin{aligned} Y_t &= X_t^2 \\ Z_{t-1} &= (1, Y_{t-1}, \dots, Y_{t-p})' = (1, X_{t-1}^2, \dots, X_{t-p}^2)' \\ \eta_t &= \varepsilon_t^2 - 1 \quad 1 - p \leq t \leq n \end{aligned}$$

Le modèle peut alors s'écrire

$$h_{t-1}^2(\beta) = Z_{t-1}' \beta$$

Après avoir élevé l'équation (3.1) au carré, on utilise la forme de $h_{t-1}(\beta)$ donné plus haut et on obtient :

$$Y_t = Z_{t-1}' \beta + h_{t-1}^2(\beta) \eta_t \quad 1 \leq t \leq n \quad (3.2)$$

avec

$$E(h_{t-1}^2 \eta_t) = E(h_{t-1}(\beta)^2) \cdot E(\eta_t) = 0 \quad 1 \leq t \leq n$$

L'équation (3.2) est similaire à un modèle linéaire du type

$$Y_t = Z'_{t-1} \beta + \mu_t \quad 1 \leq t \leq n$$

avec $E(\mu(t)) = 0$.

La procédure comporte deux étapes :

Première étape

On ignore le caractère aléatoire de $h_{t-1}(\beta)^2$ et la présence de β dans son expression, on utilise la méthode des moindres carrés pour trouver l'estimateur préliminaire qui est

$$\hat{\beta}_{pr} = (Z'Z)^{-1}Z'Y \quad (3.3)$$

avec Z la matrice $n \times (1 + p)$ dont la i ème ligne est Z'_{i-1} et $Y = (Y_1, Y_2, \dots, Y_n)'$.

Deuxième étape

On divise chaque terme de l'équation (3.2) par $h_{t-1}^2(\beta)$ et on aura :

$$\frac{Y_t}{h_{t-1}^2(\beta)} = \frac{Z'_{t-1}}{h_{t-1}^2(\beta)} \beta + \eta_t \quad 1 \leq t \leq n \quad (3.4)$$

Puis, on remplace $h_{t-1}^2(\beta)$ par $h_{t-1}^2(\hat{\beta}_{pr})$. On obtient l'expression

$$\frac{Y_t}{h_{t-1}^2(\hat{\beta}_{pr})} = \frac{Z'_{t-1}}{h_{t-1}^2(\hat{\beta}_{pr})} \beta + \eta_t \quad 1 \leq t \leq n \quad (3.5)$$

Toujours, en ignorant le caractère aléatoire de $h_{t-1}^2(\hat{\beta}_{pr})$, la méthode des moindres carrés appliquée au modèle de régression (3.5) donne l'estimateur final

$$\hat{\beta} = \frac{\sum_{t=1}^n Z'_{t-1} Y_t / h_{t-1}^4(\hat{\beta}_{pr})}{\sum_{t=1}^n Z'_{t-1} Z_{t-1} / h_{t-1}^4(\hat{\beta}_{pr})}$$

3.2.2 Convergence asymptotique de l'estimateur β

Rappelons que nous avons supposé que le processus est stationnaire et ergodique. Nous commençons par énoncer un lemme concernant la distribution de l'estimateur préliminaire $\hat{\beta}_{pr}$ donné par l'équation (3.3). Nous supposons dans ce qui suit que $\forall 1 \leq j, k, l, m \leq p$ la condition suivante est vérifiée :

$$E\{Y_j Y_k Y_l Y_m\} < \infty \quad (3.6)$$

La condition (3.6) nous assure que $E(Z_0 Z_0')$ et $E\{(\beta' Z_0)^2 Z_0 Z_0'\}$ sont finis. Quand les erreurs sont normales la condition suffisante pour l'existence de moments d'ordre supérieur à 1 de Y en fonction du paramètres β est donnée par Engle (1982, Théorème 1 et 2).

Lemme 3.1. *En plus des hypothèses formulées en (3.1), on suppose que l'hypothèse (3.6) est retenue. Alors*

$$\sqrt{n}(\hat{\beta}_{pr} - \beta) \xrightarrow{D} N \left[0, \text{Var}(\varepsilon_t^2) \{E(Z_0 Z_0')\}^{-1} E\{(\beta' Z_0)^2 Z_0 Z_0'\} \{E(Z_0 Z_0')\}^{-1} \right]. \quad (3.7)$$

Preuve

De (3.2) et (3.3), on a

$$\sqrt{n}(\hat{\beta} - \beta) = \left[n^{-1} \sum_{t=1}^n Z_{t-1} Z_{t-1}' \right]^{-1} \left[n^{-1/2} \sum_{t=1}^n Z_{t-1} h_{t-1}^2(\beta) \eta_t \right] \quad (3.8)$$

En considérant les conditions de stationnarité et d'ergodicité, nous obtenons la convergence

$$n^{-1} \sum_{t=1}^n Z_{t-1} Z_{t-1}' \xrightarrow{P.S} E(Z_0 Z_0').$$

qui est fini par (3.6).

Pour conclure que la somme $\sum_{t=1}^n n^{-1/2} Z_{t-1} h_{t-1}^2(\beta) \eta_t$ converge vers un vecteur normal de moyenne nulle et de variance $\text{Var}(\eta_1) E\{h_0^4(\beta) Z_0 Z_0'\}$, il est suffisant de montrer que $\forall l \in \mathbb{R}^p - \{0\}$,

$$n^{-1/2} \sum_{t=1}^n l' Z_{t-1} \beta' Z_{t-1} \eta_t \rightarrow N \left[0, \text{Var}(\eta_1) E\{(l' Z_0)^2 (\beta' Z_0)^2\} \right]. \quad (3.9)$$

Pour démontrer cela, on vérifie que les conditions du théorème central limite de martingale (Hall et al, 1980, Cor.3.1) sont satisfaites.

Soit la suite croissante des σ -algèbre $F_t = \{\sigma < Z_0, \dots, Z_t >; t \geq 0\}$.

Alors $\{Z_{t-1}h_{t-1}^2(\beta)\eta_t; F_t; t \geq 1\}$ est une différence de martingale.

Maintenant on vérifie uniquement la condition de Lindeberg conditionnelle car les autres sont faciles à vérifier.

Soit E_{t-1} l'espérance conditionnelle avec

$$F_{t-1} = \sigma(Z_0, \dots, Z_{t-1}) \quad t \geq 1.$$

On montre que $\forall \varepsilon > 0$,

$$\begin{aligned} L_n(\varepsilon) &= n^{-1} \sum_{t=1}^n \{l'Z_{t-1}\beta'Z_{t-1}\}^2 E_{t-1}\eta_t^2 I(|\eta_t l'Z_{t-1}\beta'Z_{t-1}| > n^{1/2}\varepsilon^2) \\ &= o_p(1) \end{aligned} \tag{3.10}$$

Notons que

$$\begin{aligned} L_n(\varepsilon) &\leq n^{-1} \sum_{t=1}^n \{l'Z_{t-1}\beta'Z_{t-1}\}^2 E_{t-1} [\eta_t^2 \{I(|\eta_t| > n^{1/4}\varepsilon) + I(|l'Z_{t-1}\beta'Z_{t-1}| > n^{1/4}\varepsilon)\}] \\ &= n^{-1} \sum_{t=1}^n \{l'Z_{t-1}\beta'Z_{t-1}\}^2 E \{ \eta_1^2 I(|\eta_1| > n^{1/4}\varepsilon) \} \\ &\quad + E(\eta_1^2) n^{-1} \sum_{t=1}^n \{l'Z_{t-1}\beta'Z_{t-1}\}^2 I(|l'Z_{t-1}\beta'Z_{t-1}| > n^{1/4}\varepsilon) \\ &= S_1 + E(\eta_1^2)S_2 \end{aligned}$$

Comme

$$n^{-1} \sum_{t=1}^n \{l'Z_{t-1}\beta'Z_{t-1}\}^2 = E \{l'Z_0\beta'Z_0\}^2 + o_p(1)$$

et $E(\eta_1^2) < \infty$, on a $S_1 = o_p(1)$. On remarque que $S_2 \geq 0$ avec

$$E(S_2) = E \left\{ \{l'Z_0\beta'Z_0\}^2 I(|l'Z_0\beta'Z_0| > n^{1/4}\varepsilon) \right\} = o(1)$$

Grâce à la condition (3.6). D'où la condition (3.10) est satisfaite et (3.9) sera donc établie.

Le théorème suivant donne la distribution asymptotique de $\hat{\beta}$.

Dans ce qui suit, on suppose que les propriétés suivantes de $\hat{\beta}_{pr}$ sont vérifiées.

$$\sqrt{n}(\hat{\beta}_{pr} - \beta) = o_p(1) \quad (3.11)$$

et les propriétés des moments de $\{X_t; 1 - p \leq t \leq 0\}, \forall 1 \leq j, k, l \leq p,$

$$E \left\{ \frac{Y_{-j}Y_{-k}Y_{-l}}{(\beta'Z_0)^3} \right\} < \infty \quad (3.12)$$

Notons que les conditions du lemme ci dessus impliquent (3.11).

On note également que dans l'étude asymptotique de l'EQMV, Weiss(1986) suppose que $1 \leq j, k \leq p, E \{Y_{-j}Y_{-k}/(\beta'Z_0)^2\} < \infty$.

Quand $\beta_j > 0 \forall 1 \leq j \leq p,$ (3.12) est automatiquement vérifiée du fait que la fonction $y \rightarrow y/(\beta_0 + \beta_j y)$ est bornée.

Théorème 3.1. *En plus des hypothèses du modèle (3.1). On suppose que (3.11) et (3.12) sont vérifiées. Alors*

$$\sqrt{n}(\hat{\beta} - \beta) \xrightarrow{D} N \left[0, Var(\varepsilon_1^2) \{E \{Z_0 Z_0' (\beta' Z_0)^{-2}\} \}^{-1} \right]. \quad (3.13)$$

Preuve

Comme dans (3.8),

$$\sqrt{n}(\hat{\beta} - \beta) = \left[n^{-1} \sum_{t=1}^n \frac{Z_{t-1} Z'_{t-1}}{h_{t-1}^4(\hat{\beta}_{pr})} \right]^{-1} \left[n^{-1/2} \sum_{t=1}^n \frac{Z_{t-1} h_{t-1}^2(\beta) \eta_t}{h_{t-1}^4(\hat{\beta}_{pr})} \right].$$

puis on montre que

$$n^{-1} \sum_{t=1}^n \left\{ \frac{1}{h_{t-1}^4(\hat{\beta}_{pr})} - \frac{1}{h_{t-1}^4(\beta)} \right\} Z_{t-1} Z'_{t-1} = o_p(1) \quad (3.14)$$

et

$$n^{-1/2} \sum_{t=1}^n h_{t-1}^2(\beta) \left\{ \frac{1}{h_{t-1}^4(\hat{\beta}_{pr})} - \frac{1}{h_{t-1}^4(\beta)} \right\} Z_{t-1} \eta_t = o_p(1). \quad (3.15)$$

Alors

$$n^{1/2}(\hat{\beta}_{pr} - \beta) = \left[n^{-1} \sum_{t=1}^n \frac{Z_{t-1} Z'_{t-1}}{h_{t-1}^4(\beta)} \right]^{-1} \left[n^{-1/2} \sum_{t=1}^n \frac{Z_{t-1} h_{t-1}^2(\beta) \eta_t}{h_{t-1}^4(\beta)} \right] + o_p(1) \quad (3.16)$$

et donc, en appliquant le théorème central limite de martingal, (3.13) sera similaire à (3.7).

Remarque

Pour plus de détails voir Bose et Mukherjee (2003).

3.3 Comparaison de la performance avec des données simulées

Dans le tableau suivant, nous présentons quelques résultats de simulation pour le modèle ARCH(1) défini par $\varepsilon_t = h_t \eta_t$ avec $h_t = \sqrt{\alpha_0 + \alpha_1 \varepsilon_{t-1}^2}$ avec erreur standard normale. Ici pour chacun des i -ème essais ($1 \leq i \leq 10$), une paire particulière de (β_0, β_1) est sélectionnée de manière aléatoire suivant une distribution \mathcal{X}^2 . Puis on calcule les estimateurs de (β_0, β_1) à la k -ième itération où $1 \leq k \leq 1000$, en considérant l'échantillon $\{X_t; 0 \leq t \leq n = 30\}$, avec la méthode en deux étapes et la méthode du QMV. Soit

$$S_0 = \sum_{k=1}^{1000} \frac{(\beta_0 - \hat{\beta}_{0k})^2}{1000}$$

et

$$S_1 = \sum_{k=1}^{1000} \frac{(\beta_1 - \hat{\beta}_{1k})^2}{1000}$$

où les estimateurs $\hat{\beta}_{0k}$ et $\hat{\beta}_{1k}$ sont obtenus par la méthode de résolution des équations linéaires à la k -ème itération. De même que S_{0Q} et S_{1Q} sont définis lorsque les estimateurs sont obtenus par la méthode du QMV.

Essai	β_0	β_1	S_0	S_{0Q}	S_1	S_{1Q}
1	0.35463	0.11346	0.0123	0.01649	0.0522	0.13894
2	0.69867	0.38751	0.0800	0.08463	0.0685	0.07058
3	1.82138	0.04873	0.2758	0.36670	0.0384	0.07021
4	0.72256	0.25672	0.0543	0.06757	0.0472	0.07285
5	1.10568	0.27014	0.1682	0.17849	0.0455	0.08187
6	0.89437	0.41837	0.0988	0.11150	0.0567	0.11589
7	1.18623	0.40427	0.3102	0.31742	0.06482	0.10209
8	0.61056	0.25444	0.0526	0.06219	0.0333	0.04743
9	0.84292	0.02405	0.07132	0.07347	0.0434	0.12475
10	1.83434	0.29898	0.4594	0.60428	0.0532	0.09289

TAB. 3.1 – Valeurs de S_0 , S_1 , S_{0Q} et S_{1Q} avec des erreurs normal.

On constate du tableau ci dessus, que même avec un échantillon de petite taille ($n=30$) et des erreurs gaussiennes, $S_0 < S_{0Q}$ et $S_1 < S_{1Q}$ pour chaque essai.

De ce fait, on déduit que la méthode a deux étapes est plus performante que la méthode du QMV, même dans le cas des échantillons de petites tailles.

La colonne 4 et 6 ont été obtenues par le programme (3) cité en annexe.

3.4 Calcul du biais de l'estimateur de la méthode a deux étapes

Dans le tableau suivant, nous présentons quelques résultats obtenu avec le logiciel MATLAB du calcul du biais des estimateurs de la methode a deux étapes. Le modèle utilisé ici est le même utilisé précédemment.

Soit

$$S_{00} = \sum_{k=1}^{10000} \frac{(\beta_0 - \hat{\beta}_{0k})}{10000}$$

et

$$S_{11} = \sum_{k=1}^{10000} \frac{(\beta_1 - \hat{\beta}_{1k})}{10000}$$

Essai	β_0	β_1	S_{00}	S_{11}
1	0.35463	0.11346	-0.0792	-0.0047
2	0.69867	0.38751	-0.3730	-0.0039
3	1.82138	0.04873	-1.6514	0.2734
4	0.72256	0.25672	-0.3946	0.0627
5	1.10568	0.27014	-0.8287	0.1483
6	0.89473	0.41837	-0.6020	0.0430
7	1.18623	0.40427	-0.9380	0.1009
8	0.61056	0.25444	-0.2746	0.0220
9	0.84292	0.02405	-0.5741	0.2027
10	1.83434	0.29898	-1.6430	0.1993

TAB. 3.2 – Valeurs du biais des estimateurs de la méthode a deux étapes.

On constate dans le tableau ci dessus, que les valeurs des biais obtenus sont très petites, ce qui fait que les estimateurs de la méthode a deux étapes sont efficaces.

Les biais calculés ont été obtenu par le programme (3) cité en annexe.

3.5 Simulation du comportement asymptotique de l'estimateur de la méthode a deux étapes

Nous avons fait une étude de simulation pour le comportement asymptotique de l'estimateur obtenu par la méthode a deux étapes en utilisant le logiciel MATLAB.

Cette étude empirique est faite pour illustrer le résultat du lemme (3.1) et du théorème (3.1).

3.5.1 Etude du comportement asymptotique de l'estimateur préliminaire $\hat{\beta}_{pr}$

Pour étudier le comportement asymptotique de cet estimateur, nous allons faire varier la taille de l'échantillon du modèle ARCH(1) défini par

$$X_t = \varepsilon_t h_{t-1}(\beta)$$

avec

$$h_{t-1}(\beta) = \sqrt{\beta_0 + \beta_1 X_{t-1}^2}$$

Où ε_t sont iid de moyenne nulle et de variance unitaire.

$$\hat{\beta}_{pr} = (Z'Z)^{-1}Z'Y$$

Etude du comportement de l'estimateur $\hat{\beta}_{0pr}$

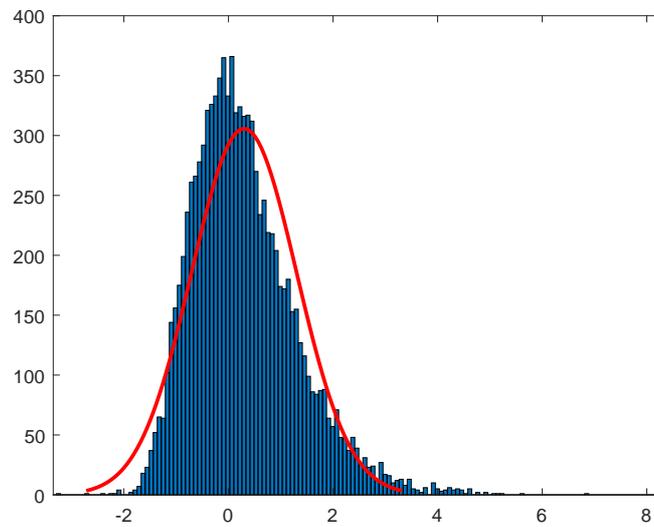


FIG. 3.1 – Comportement de l'estimateur $\hat{\beta}_{0pr}$ pour $n=20$.

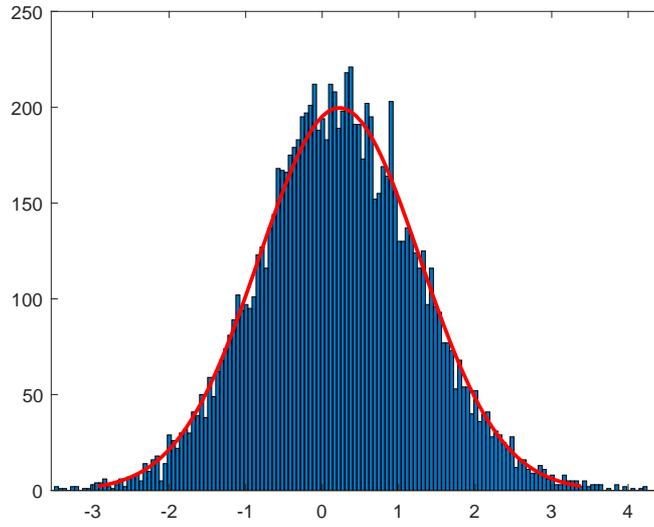


FIG. 3.2 – Comportement de l'estimateur $\hat{\beta}_{0pr}$ pour $n=200$.

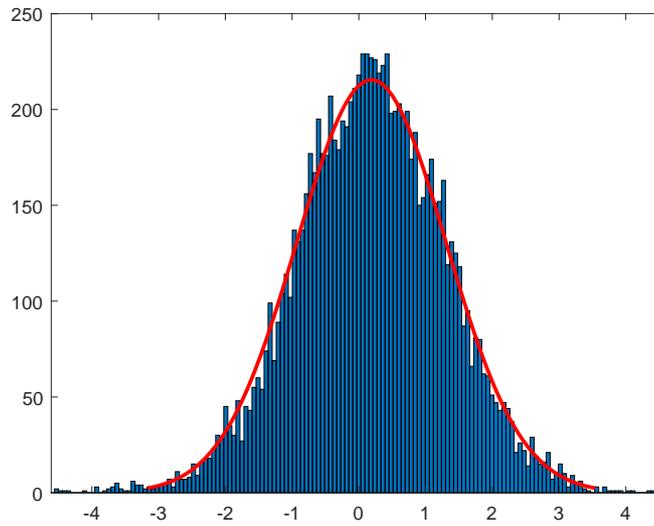


FIG. 3.3 – Comportement de l'estimateur $\hat{\beta}_{0pr}$ pour $n=500$.

Etude du comportement de l'estimateur $\hat{\beta}_{1pr}$

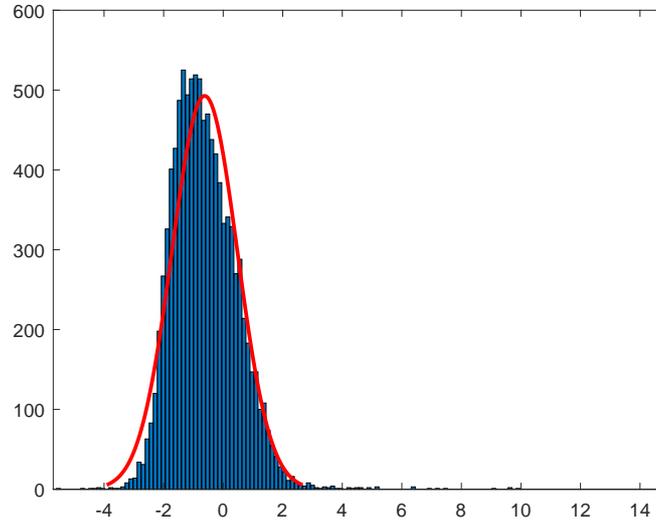


FIG. 3.4 – Comportement de l'estimateur $\hat{\beta}_{1pr}$ pour $n=20$.

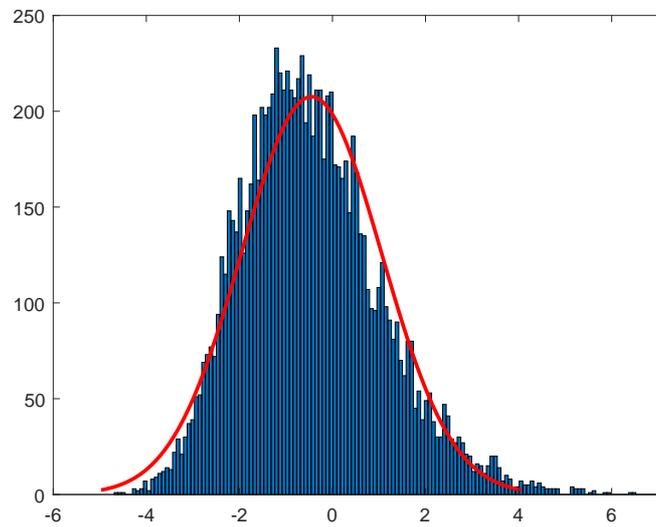


FIG. 3.5 – Comportement de l'estimateur $\hat{\beta}_{1pr}$ pour $n=200$.

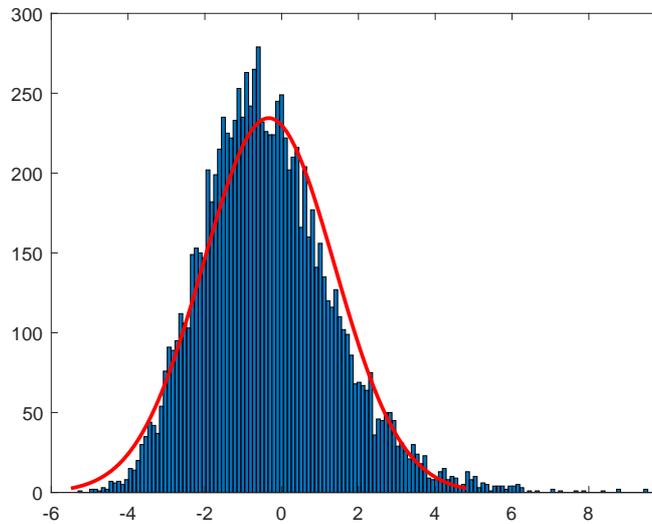


FIG. 3.6 – Comportement de l'estimateur $\hat{\beta}_{1pr}$ pour $n=500$.

On constate que plus n est grand, plus le comportement de l'estimateur préliminaire tend vers la loi gaussienne. Ce qui confirme le résultat du lemme (3.1).

Les graphes précédents ont été obtenu avec le logiciel MATLAB après exécution du programme (1) cité en annexe.

3.5.2 Etude asymptotique de l'estimateur final $\hat{\beta}$

Pour étudier le comportement asymptotique de cet estimateur, nous allons faire varier la taille de l'échantillon du modèle ARCH(1) définit par

$$X_t = \varepsilon_t h_{t-1}(\beta)$$

avec

$$h_{t-1}(\beta) = \sqrt{\beta_0 + \beta_1 X_{t-1}^2}$$

Où ε_t sont iid de moyenne nulle et de variance unitaire.

$$\hat{\beta} = \frac{\sum_{t=1}^n Z'_{t-1} Y_t / h_{t-1}^4(\hat{\beta}_{pr})}{\sum_{t=1}^n Z'_{t-1} Z_{t-1} / h_{t-1}^4(\hat{\beta}_{pr})}$$

Etude du comportement de l'estimateur $\hat{\beta}_{\text{final}}$

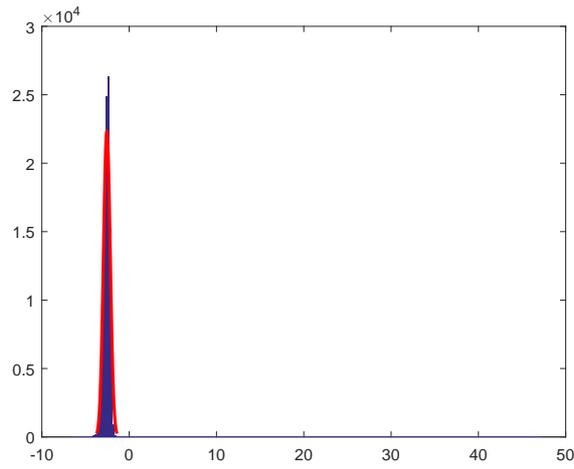


FIG. 3.7 – Comportement de l'estimateur $\hat{\beta}_{\text{final}}$ pour $n=250$.

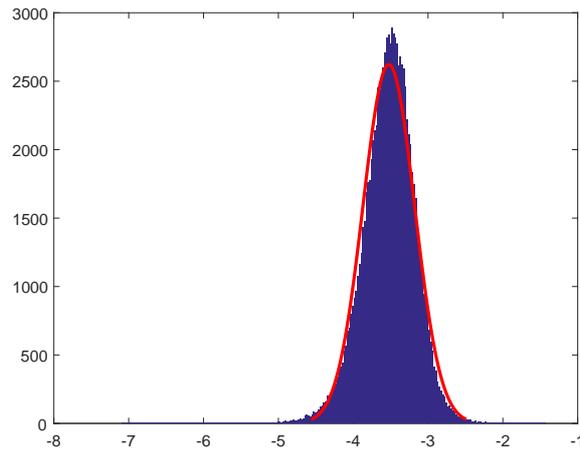


FIG. 3.8 – Comportement de l'estimateur $\hat{\beta}_{\text{final}}$ pour $n=500$.

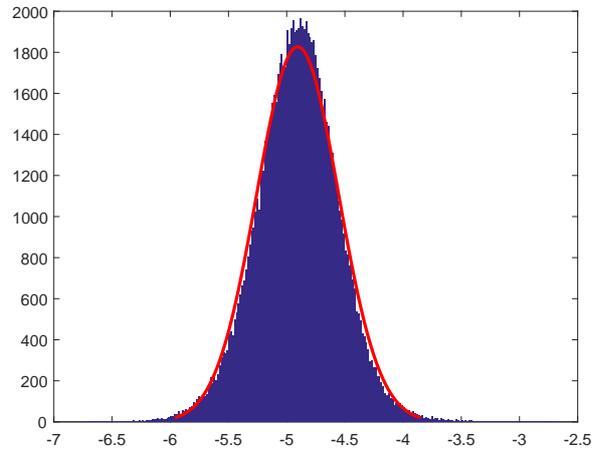


FIG. 3.9 – Comportement de l'estimateur $\hat{\beta}_{0final}$ pour $n=1000$.

Etude du comportement de l'estimateur $\hat{\beta}_{1final}$

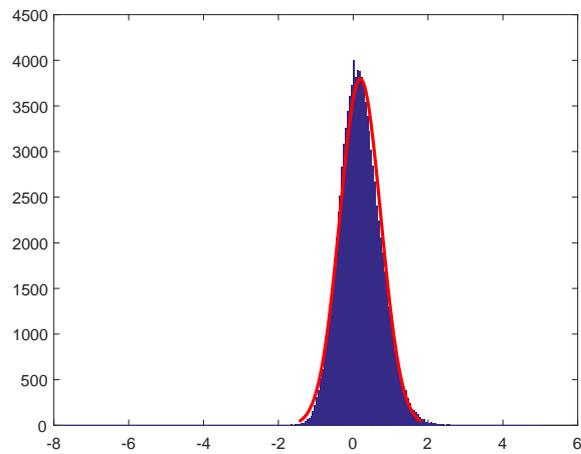


FIG. 3.10 – Comportement de l'estimateur $\hat{\beta}_{1final}$ pour $n=250$.

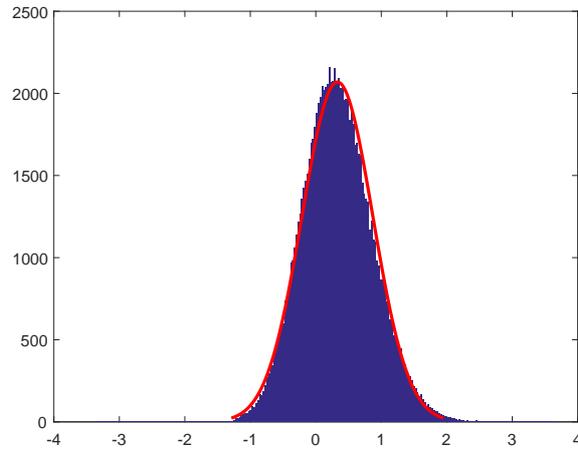


FIG. 3.11 – Comportement de l'estimateur $\hat{\beta}_{1final}$ pour $n=500$.

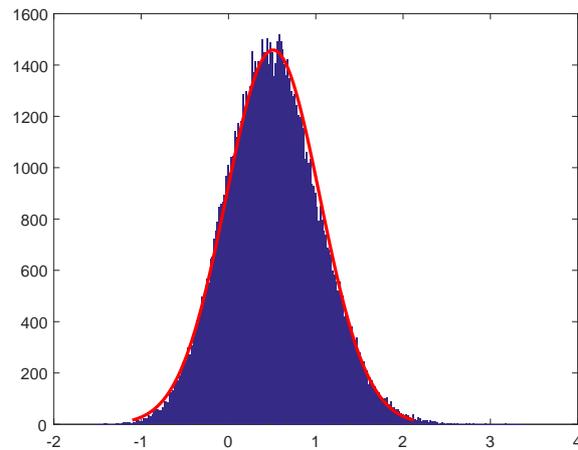


FIG. 3.12 – Comportement de l'estimateur $\hat{\beta}_{1final}$ pour $n=1000$.

On constate que plus n est grand, plus le comportement de l'estimateur préliminaire tend vers la loi gaussienne. Ce qui confirme le résultat du Théorème (3.1). Les graphes précédents ont été obtenu avec le logiciel MATLAB après exécution du programme (2) cité en annexe.

Conclusion et perspectives

Dans le premier chapitre, nous nous sommes intéressés aux modèles ARCH /GARCH introduit pour la première fois par Engle en 1982, leurs propriétés et quelques extensions IGARCH, TARCH, EGARCH... mais nous avons d'abord donné quelques outils de base de la théorie des séries chronologiques.

Dans le deuxième chapitre, nous avons étudié l'estimation des paramètres d'un modèle ARCH et cela en utilisant deux méthodes distinctes à savoir ; la méthode des moindres carrés ordinaires et l'estimation du maximum de vraisemblance, nous avons étudié les propriétés asymptotique de ce dernier, et enfin nous avons donné le test d'homoscédasticité.

Dans le dernier chapitre nous avons abordé une troisième méthode d'estimation pour les paramètres d'un modèle ARCH qui est l'estimation par la méthode en deux étapes proposé par Bose et Mukherjee en 2001 et nous avons modélisé la méthode en utilisant les commandes du Logiciel MATLAB.

Plusieurs chercheurs essayent d'étudier la raison de la volatilité du marché. La raison de son existence n'est pas seulement les informations existantes sur le marché mais aussi les comportements des investisseurs, les bulles spéculatives et plusieurs autres facteurs. De ce fait il serait plus intéressant d'introduire les méthodes Bayésienne dans l'inférence des modèles ARCH.

Bibliographie

- [1] BAILLIE, R., BOLLERSLEV, T., AND MIKKELSEN, H. Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of Economics* 74(1). (1996), 3–30.
- [2] BOLLERSLEV., ENGLE., AND NELSON. Arch models. *Handbook of Econometrics. IV.* (1994).
- [3] BOLLERSLEV, T. Generalised autoregressive conditional heteroskedasticity. *Journal of Econometrics.* 31. (1986), 307–327.
- [4] BOLLERSLEV, T., CHOU, R., JAYARAMAN, N., AND KRONER, K. Les modèles arch en finance : un point sur la théorie et les résultats empirique. *Annales d'Economie et de statistique* 24. (1991), 1–59.
- [5] BOLLERSLEV, T., CHOU, R., AND KRONER, K. Arch modeling in finance : A review of the theory and empirical evidence. *Journal of Econometrics.* 52. (1992), 5–59.
- [6] BOSE, A., AND MUKHERJEE, K. Estimating the arch parameters by solving linear equations. *Journal of Time Series Analysis* 24(2). (2003), 127–136.
- [7] BOUGEROL, P., AND PICARD, N. Stationarity of garch processes and of some nonnegative time series. *Journal of Econometrics* 52. (1992), 115–127.
- [8] BOURBONNAIS, R., AND TERRAZA, M. Analyse des séries temporelles en économie.
- [9] BOX, G., AND JENKINS, G. Times series analysis, forecasting and control, san francisco. *Holden-Day, (édition révisée)* (1976).
- [10] BÉRA, A., LEE, S., AND HIGGINGS, M. Arch models : Properties, estimation and testing. *Journal of Economic Surveys.* (1993).
- [11] DROESBEKE, J., FICHET, B., AND TASSI, P. Modelisation arch, théorie statistique et application dans le domaine de la finance. 242.
- [12] ENGLE, R. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica.* 50 (1982), 987–1007.

-
- [13] ENGLE, R. Selected readings. *Oxford university Press* (1995), 183–196.
- [14] ENGLE, R., AND BOLLERSLEV, T. Modelling the persistence of conditional variance. *Econometric reviews* 5(1). (1986), 1–50.
- [15] ENGLE, R., AND KRAFT, D. Multiperiod forecast error variances of inflation estimated from arch models. *Applied time series analysis of economic data* (1993), 293–302.
- [16] ENGLE, R., LILIEN, D., AND ROBINS, R. Estimating time varying risk premia in the term structure : The arch-m model . *Econometrica* 55. (1987), 391–407.
- [17] GOURIEROUX, C. Modèles arch : Applications financières et monétaires,paris. *Economica* (1992).
- [18] GOURIEROUX, C., AND MONFORT, A. Statistique et modèles économétriques :notions générales, estimation, prévision, algorithmes. *Journal of Econometrics* 52. (1989), 159–199.
- [19] GOURIEROUX, C., MONFORT, A., AND TROGNON, A. Pseudo maximum likelihood methods : Theory. *Econometrica*. 52. (1984), 681–700.
- [20] HURLIN, C. Econometrie pour la finance, modèles arch-garch. *Edition de l'université d'Orléans*. (2004).
- [21] LEE, J. A lagrange multiplier test for garch models. *Economics Letters* 37. (1991), 265–271.
- [22] LEE, J., AND KING, M. A locally most mean powerful score test of arch and garch regression disturbances. *Journal of Business and Economic Statistics* 11(1). (1993), 17–27.
- [23] MILHOJ, A. The moment structure of arch processes. *Scandinavian journal of statistics* 12(4). (1985), 281–292.
- [24] NELSON, D. Stationarity and persistence in the garch(1,1) model. *Econometric Theory*. 6. (1990), 318–334.
- [25] SHEPHARD, N. Statistical aspects of arch and stochastic volatility in time series models in econometric, finance and other fields. *Edition by Cox,D.R, Hinkley,D.V and Barndorff-Nielsen,O.E.* 133. (1996), 1–67.
- [26] WEISS, A. Asymptotique theory for arch models :estimation and testing. *Econometric Theory*. 2. (1986), 07–131.
- [27] ZAKOÏAN, J. Thresholds heteroskedastic model. *Journal of Economic Dynamics and Control* 18. (1990), 931–955.
- [28] ZAKOÏAN, J.-M. Models arch : Une revue de la littérature. *Journal de la société statistique de Paris*. 133. (1991), 40–57.