Remerciements

Je rends grace à Dieu qui m'a donné la volonté, la patience et le courage pour accomplir ce modeste travail.

Je remercie tout particulièrement M^r Abdelghani HAMAZ de m' avoir encadré et pour ses précieux conseils et idées.

Je remercie toutes celles et tous ceux m'ont aidé, de près ou de loin, par leurs ouvrages, leurs expériences, leurs avis et leurs opinions. Ils ont contribué considérablement à la réalisation de ce travail.

Mes remerciement sont adressés également aux membres du département de Mathématiques : Ensiegnants et camarades étudiants.

Mes respecteux remerciements aux membres de jury pour m'avoir honoré par leurs présences afin d'examiner et évaluer ce travail.

Dédicace

Je dédie ce travail à :

Mes trés chers parents qui sont sans cesse à mes cotés,

Mes trés chers frères et

mes chères soeurs qui m'ont toujours soutenue,

toute ma famille

et mes amis(es)(Baya, Romane, djedjiga et Ali) pour leur soutien morale .

Table des matières

Introduction

L'apparition des modèles ARCH et GARCH sont remplaçé dans le contexe le plus vaste du débat sur la représentation linéaire ou non linéaire des processus stochastique temporels. Les modèles ARCH autorégressifs hétéroscédastique introduit par Engle 1982 et généralisés par Bollerslev 1986, sont devenus extrêmement populaires parmi les académiques et les particiens.

Les modèles GARCH ont mené à un changement fondamental la modélisation des séries financières et sont particulièrement annoncés pour prendre en compte les caractéristiques importantes de ces séries (stationnarité, volatilité, asymétrie, saisonnalité,. . .). De plus, ces processus prennent en compte dans la modélisation la forte **Leptokurticité** observée dans la loi de distribution non conditionnelle de la plupart des séries financières. Ces processus sont proposés pour compléter l'insuffisante des modèles ARMA, l'avantage de ces modèles est expliqué par le fait qu'elle sont riches de coté théorique et simple à utilisé dans la pratique. Depuis la fin des années 80, de nombreuses extentions des modèles ARCH/GARCH ont été édités ces extensions ont provoqué des nouvelles directions pour la recherche statistique et probabilité.

Le but de notre travail est de fournir une introduction aux modèles GARCH/ARCH le plus souvent utilisé dans la modélisation en temps discret des marchés financiers. On s'intéressera plus particulièrement à l'estimation de ces modèles où la consistence et le comportement asymptotique des estimateurs sont étudiés. L'objectif n'est pas de fournir toute la théorie relative à ces modèles mais d'insister sur l'estimation des modèles ARCH/GARCH.

Dans le chapitre 1, nous présentons un rappel des principales définitions et propriétés des series financières, des processus ainsi que quelques notions utiles dans notre travail (existence des solutions stationnaires, des représentation, la notion de stationnarité faible et forte).

Dans le chapitre2, nous présentons trois méthodes d'estimation, commençant par la méthode des moindres carrés (M.C.O), ensuite la méthode de maximum de vraisemblance (M.V) et enfin la méthode des deux étapes.

Le dernier chapitre est consacré à l'estimation du maximum de vraisemblance des modèles GARCH(p, q) lorsque certains coefficients sont nuls. La distribution asymptotique les résultats sont obtenus en imposant d'autres conditions moins restrictives que les hypothèses classiques du cas général. Ce mémoire s'achève par une conclusion contenant de futures perspectives.

Enfin, l'annexe inclut les propriétés probabilistes qui sont importantes pour l'étude des modèles GARCH.

Chapitre 1

Processus conditionnellement hétéroscédastique

Les modèles ARCH et GARCH (Autoregressive Conditional Heteroscedasticity and Generalized Autoregressive Conditional Heteroscedasticity) sont des modèles qui jouent un rôle primordial au niveau de la description des séries financières vu le comportement hétéroscédastique de leur variance conditionnelle σ_t^2.
En effet, les modèles ARCH introduit par Engle (1982) et généralisé par Bollerslev (1986) leur caractérisation repose essentiellement sur le concept de variance conditionnelle σ_t^2, celle ci s'écrit comme une fonction affine des valeurs passées du carré du processus. Cette spécification particulière se révèle très fructeuse car elle permet une étude complète des propriétés des solutions tout en étant assez générale. Les modèles GARCH sont en effet susceptibles de capter les propriétés caractéristiques des séries financières.
Dans ce chapitre, nous présentons d'abord des définitions et des représentations des modèles ARCH et GARCH, nous établissons la condition de la stationnarité stricts et de second ordre, ensuite nous présentons quelques extensions du modèle ARCH, enfin nous étudions la représentation ARCH(∞) d'un GARCH.
Mais avant de présenter les modèles ARCH et GARCH commençons par introduire quelques propriétés essentielles des séries financières.

1.1 Propriétés principales des séries financiéres

Les séries financières (rentabilités d'action, taux d'intérêt, taux de change...) sont des séries de prix d'actif et de rendements présentent généralement un certain nombre de propriétés similaires suivant leur périodicité. Soit p_t le prix d'un actif à la date t. L'unité temporelle peut être le jour, la semaine ou le mois et ε_t le logarithme du rendement correspondant (c'est à dire ε_t est la différence première du logarithme du prix p_t à l'instant t) :

$$\varepsilon_t = log(p_t) - log(p_{t-1}) = log(1 + s_t),$$

où $s_t = (p_t - p_{t-1})/p_{t-1}$ désigne la variation relative des prix.

Kurtosis Soit μ_k le moment empirique centré d'ordre k du processus $\{X_t\ ,\ t = 1 \ldots T\}$

$$\begin{aligned}\mu_k &= \mathbb{E}[X_t - \mathbb{E}[X_t]]^k \\ &= \frac{1}{T}\sum_{t=1}^{T}[X_t - \overline{X}]^k.\end{aligned}$$

Définition 1.1.1. *Le Kurtosis d'une variable aléatoire X correspond à son moment centré d'ordre 4, c'est à dire :*

$$\mu_4 = \mathbb{E}[(X-\mu)^4].$$

Remarque 1.1.1. *Le Kurtosis est une mesure de "L'épaisseur" des queues de distributions. En règle générale, on exprime cette mesure en contrôlant par une fonction puissance de la variance* $\mathbf{V}(X) = \sigma^2$.
On définit ainsi une nouvelle mesure : le degré d'excès de Kurtosis :

$$\mathbf{K}_u = \mathbb{E}\left[\left(\frac{X-\mathbb{E}(X)}{\sigma_x}\right)^4\right] - 3.$$

- *Le Kurtosis de la disribution normale est 3.*
- *Si Kurtosis superieur à 3 (queues épaisses) la distribution est plutôt pointue (distribution leptokurtique) .*
- *Si le Kurtosis est inférieur à 3, la distribution est plutôt plate (distribution est dite Platikurtique).*

Propriétée 1.1.1. ***[Queue de distribution épaisses]*** *Dans les distributions empiriques des séries des rendements, on s'aperçoit que l'hypothèse de normalité est rejetée. Plus précisément, les densités de probabilité de ces séries présentent des queues épaisses (à décroissance plus lente que* $\exp(-x^2/2)$ *et des pics en zéro. On parle alors de distribution* ***Leptokurtique****. Une mesure de cet effet est obtenue à partir du coefficient de Kurtosis, rapport du moment empirique centré d'ordre 4 et du carré de la variance empirique, qui est asymptotiquement égal à 3 dans le cas gaussien et est superieur à 3 pour ces séries.*

Définition 1.1.2. *[**La volatilité**] est une mesure de l'instabilité du cours d'un actif financier. Elle mesure l'amplitude des variations d'une action, d'un produit dérivé ou d'un marché. Il s'agit d'un paramètre de quantification du risque de rendement et de prix. Les séries monétaires et financières sont caractérisées par le* ***clustering*** *de volatilité, à savoir les périodes de forte volatilité alternent avec les périodes de faible volatilité. Ce phénomène, que nous appelons aussi l'hétéroscédasticité conditionnelle, est particulièrement fréquent dans les données boursières, les taux de changes ou d'autres prix déterminés sur les marchés financiers.*

Propriétée 1.1.2. ***[Effet Levier]*** *cette propriété, notée par Brock en 1976, il existe une asymétrie entre l'effet des valeurs passées négatives et l'effet des valeurs passées positives sur la volatilité des cours ou des rendements. Les valeurs négatives (baisses du cours) tendent à provoquer une augmentation de la volatilité superieure à celle induite par des valeurs positives (hausse des cours) de même amplitude (il s'agit d'une asymétrie de la relation liant les valeurs passées des cours ou de rendements à la volatilité de ces derniers).*

Propriétée 1.1.3. ***[La Saisonnalité]*** *cet aspect nous explique un peu le lien qu'il y a entre la volatilité et l'effet du week-end et des jours fériés. C'est à dire que les marchés sont trés volatiles à la fermeture (week-end et jours fériés), (la volatilité tend à augmenter lorsque les marchés ferment).*

Les fonctions d'autocovariance et de corrélations

Définition 1.1.3. *Soit $\{X_t;\ t \in \mathbb{Z}\}$ une série stationnaire, la fonction d'autocovariance de (X_t) est définie par :*

$$\gamma_X(h) = cov(X_t, X_{t+h}) = \mathbb{E}\big[(X_t - \mathbb{E}(X_t))(X_{t+h} - \mathbb{E}(X_{t+h}))\big], \quad \forall t, h \in \mathbb{Z}.$$

Définition 1.1.4. *La fonction d'autocorrélation de $\{X_t;\ t \in \mathbb{Z}\}$ est définie par :*

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

Remarque 1.1.2. *Fonctions paires*

$$\gamma_X(h) = \gamma_X(-h), \quad \rho_X(h) = \rho_X(-h).$$

La stationnarité

Rappelons au passage quelques propriétés comme la stationnarité forte et de la stationnarité faible (ou stationnarité du second ordre).

Propriétée 1.1.4. ***[La stationnarité]*** *les processus stochastique p_t associés aux prix d'actif ne vérifie pas la stationnarité au sens de la stationnarité au second ordre, alors que les processus associés aux rendement ε_t sont compatibles avec la notion de la stationnarité au second ordre.*

Définition 1.1.5. ***[La stationnarité stricte]*** *Le processus $\{\mathrm{X_t};\ t \in \mathbb{Z}\}$ est dit strictement ou fortement stationnaire si $\forall\ t_1,\ t_2,\ \ldots,\ t_n, h \in \mathbb{Z}$,on a :*

$$\mathcal{L}(X_{t_1}, X_{t_2}, \ldots, X_{t_n})) = \mathcal{L}(X_{t_{1+h}}, X_{t_2+h} \ldots, X_{t_n+h}).$$

Définition 1.1.6. ***[La stationnarité faible au second ordre]***
Un processus $\{X_t;\ t \in \mathbb{Z}\}$ est dit stationnaire ou faiblement stationnaire ou stationnaire au second ordre si :

1. $\mathbb{E}(X_t^2) < \infty$.
2. $\mathbb{E}(X_t) = \mu, \forall t \in \mathbb{Z}$ *(ne dépend pas de temps t).*
3. $cov(X_t, X_{t+h}) = \mathbb{E}\big((X_t - \mathbb{E}(X_t))(X_{t+h} - \mathbb{E}(X_{t+h}))\big) = \gamma_X(h), \forall t, h \in \mathbb{Z}$ *(ne dépend pas de temps t, pour tout h).*

En résumé, un processus est stationnaire au second ordre si l'ensemble de ses moments sont indépendants du temps. Par conséquent, il convient de noter que la stationnarité implique que la variance $\gamma(0)$ du processus $\{X_t;\ t \in \mathbb{Z}\}$ est constante au cours du temps. Les processus P_t associés aux prix d'actif sont généralement non stationnaires au sens de la stationnarité du second ordre, tandis que les processus associés aux rendements sont compatibles avec la propriétée de stationnarité au second ordre.

Le processus Bruit Blanc (White Noise processus)

Définition 1.1.7. *Soit $\{\varepsilon_t;\ t \in \mathbb{Z}\}$ un processus stochastique, on dit que $(\varepsilon_t)_{t\in\mathbb{Z}}$ est un bruit blanc si les trois propriétés suivantes sont vérifiées :*

1. $\mathbb{E}(\varepsilon_t) = 0\ \forall t \in \mathbb{Z}$.
2. $var(\varepsilon_t) = \sigma^2\ \forall t \in \mathbb{Z}$.
3. $cov(\varepsilon_t, \varepsilon_s) = \mathbb{E}(\varepsilon_t \varepsilon_s) = 0\ ,\ \forall t \in \mathbb{Z}$.

Définition 1.1.8. *on dit que $\{\varepsilon_t;\ t \in \mathbb{Z}\}$ est un bruit blanc fort de moyenne nulle et de variance σ^2 lorsque :*
$\varepsilon_t \perp \varepsilon_s\ \forall t \neq s$ *, (les variables ε_t et ε_s sont indépendantes).*

Modèle ARMA :
Les modèles ARMA s'appuient principalement sur deux principes mis en évidence par Yule et Slutsky, le principe autorégressif (Auto Regressive) et moyenne mobile (Moving Average). Fin des années 1970, leur application à l'analyse et à la prédiction des séries temporelle fut généralisé par Box et Jenkins en combinant les deux principes ARMA, ils montrèrent que ce processus pouvait s'appliquer à de nombreux domaines et était facile à implémente.

[Modèle AR]
Un processus autorégressif est un processus dont chaque valeur est décrite comme une combinaison linéaire des valeurs précédentes plus une composante aléatoire qu'on apelle "un bruit blanc ". Le nombre de valeurs précédentes considérées est appelé l"ordre du processus".

Définition 1.1.9. *Le processus $\{X_t;\ t \in \mathbb{Z}\}$ satisfait l'équation générale d'un processus AR d'ordre p si :*

$$X_t = \delta + \sum_{i=1}^{p} \alpha_i X_{t-i} + \varepsilon_t,$$

où :

δ *est le coefficient d'accroissement.*

α_i *sont les cofficients d'autorégressifs.*

ε_t *est un bruit blanc indépendant.*

[Modèle MA] Chaque valeur est décrite par une composante d'erreur aléatoire et une combinaison linéaire des erreurs aléatoires associées aux valeurs précédentes. De même, l'ordre du processus est défini par le nombre d'erreurs précédentes prises en considération.

Définition 1.1.10. *Le processus $\{X_t;\ t \in \mathbb{Z}\}$ satisfait l'équation gènérale d'un processus MA d'ordre q :*

$$X_t = \mu + \sum_{i=1}^{q} \beta_i \varepsilon_{t-i} + \varepsilon_t,$$

où

β_i *les coefficients de moyenne mobile.*

ε_{i-1} *bruit blanc.*

Modèle mixte ARMA Le modèle de linéaire le plus courant est le modèle ARMA qui combine simplement les deux principes AR et MA.

Définition 1.1.11. *Le processus $\{X_t;\ t \in \mathbb{Z}\}$ admet l'equation générale suivante qui définit un modèle ARMA(p,q) :*

$$X_t + \sum_{i=1}^{p} \alpha_i X_{t-i} = \varepsilon_t \sum_{j=1}^{q} \beta_j \varepsilon_{t-j},$$

où p est l'ordre de processus autorégressif et q l'ordre de processus moyenne mobile.

Hétéroscédastique

Définition 1.1.12. *En statistique, on parle d'hétéroscédasticité lorsque les variances des variables examinées sont différentes. La notion d'hétéroscédasticité s'oppose à celle d'homoscédasticité, qui correspond au cas où la variance de l'erreur des variables est constante. Tandis que dans le cas d'homoscédasticité, nous avons $var(\varepsilon_t) = \sigma^2\ \forall i$, nous avons désormais $var(\varepsilon_t) = \sigma_i^2$ peut être différent de σ_j^2 , pour $i \neq j$.*

1.2 Modèle ARCH et GARCH

Le but de notre travail est de fournir une introduction aux modèles ARCH et GARCH fondée sur les deux premiers moments de ε_t conditionnels à son passé,et quelques propiétés relatives aux moments conditionnels et non conditionnels.
Dans la suite de ce chapitre, nous désignons par $(\Omega,\ \mathcal{F},\ \mathbf{P})$ un espace probabilisé, et $\mathcal{F}_{t-1}$ la σ-algèbre engendrée par tout le passé du processus ε_s, pour $s < t$, i.e $\mathcal{F}_{t-1} = \sigma(\varepsilon_s, s < t)$.

1.2.1 Modèles AutoRégressifs Conditionnellement Hétéroscédastiques ARCH

Pour bien comprendre le processus ARCH, nous allons présenter ce processus tel qu'il a été introduit par Engel (1982).

Définition 1.2.1. *Un processus $\{\varepsilon_t;\ t \in \mathbb{Z}\}$ est défini comme étant un processus ARCH(q) s'il vérifie l'équation suivante :*

$$\begin{cases} \varepsilon_t = \sigma_t \eta_t \\ \sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 = \omega + \alpha(B)\varepsilon_t^2. \end{cases} \tag{1.1}$$

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

$B^i \varepsilon_t^2 = \varepsilon_{t-i}^2$.
$\{\eta_t;\ t \in \mathbb{Z}\}$ désigne une suite de variables aléatoires indépendantes, identiquement distribuées *iid*, centrée de variance unité $\mathbb{E}(\eta_t) = 0$, $V(\eta_t) = 1$.
$\{\sigma_t;\ t \in \mathbb{Z}\}$ désigne une suite de variables telles que :

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

– η_t est indépendant de $\mathcal{F}_{t-1}$.
– $\sigma_t > 0$.

Le processus d'innovation pour ε_t^2 est par définition $\nu_t = \varepsilon_t^2 - \mathbb{E}\{\varepsilon_t^2 | \varepsilon_j, j \leq t-1\} = \varepsilon_t^2 - \sigma_t^2$, qui vérifie $\mathbb{E}(\nu_t|\mathcal{F}_{t-1}) = 0$, on remplace σ_t^2 dans l'équation(1.1) par $\varepsilon_t^2 - \nu_t$ on aura :

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

$$\Longrightarrow$$

$$\varepsilon_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \nu_t, \tag{1.2}$$

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

1.2.2 Propriétés d'un processus ARCH

Le processus $\{\varepsilon_t;\ t \in \mathbb{Z}\}$ ARCH défini par l'équation (1.1) possède les propriétés statistiques suivantes :

Propriétée 1.2.1. *Le processus $\{\varepsilon_t;\ t \in \mathbb{Z}\}$ ARCH définie par l'équation (1.1) est une différence de martingale homoscédastique :*

$$\mathbb{E}(\varepsilon_t|\mathcal{F}_{t-1}) = 0,$$

et

$$V(\varepsilon_t) = \frac{\omega}{1 - \sum_{i=1}^{q} \alpha_i}.$$

Cette propriétée signifie que le processus ARCH $\{\varepsilon_t;\ t \in \mathbb{Z}\}$ qui peut s'apparenter à un processus de bruit blanc(faible), ce qui explique notamment que l'on spécifiera des erreurs de modèles établies sous la propriété de bruit blanc des erreurs. Mais cette propriété signifie en outre que le processus ARCH $\{\varepsilon_t;\ t \in \mathbb{Z}\}$ est non conditionnellent homoscédastique.

Preuve. Pour démontrer cela, nous utilisons l'equation précédente (1.1)

1. $\mathbb{E}(\varepsilon_t|\mathcal{F}_{t-1}) = \mathbb{E}(\sigma_t\eta_t|\mathcal{F}_{t-1}) = \sigma_t\mathbb{E}(\eta_t|\mathcal{F}_{t-1}) = \sigma_t\mathbb{E}(\eta_t) = 0$,
   car $\sigma_t \in \mathcal{F}_{t-1}$,et η_t est indépendente de $\mathcal{F}_{t-1}$ et $\mathbb{E}(\eta_t) = 0$.

2. $V(\varepsilon_t) = \mathbb{E}(\varepsilon_t^2) = \mathbb{E}(\sigma_t^2\eta_t^2) = \mathbb{E}(\sigma_t^2)\mathbb{E}(\eta_t^2) = \mathbb{E}(\sigma_t^2) = \omega + \sum_{i=1}^{q} \alpha_i\mathbb{E}(\varepsilon_{t-1}^2) = \omega + \sum_{i=1}^{q} \alpha_i\mathbb{E}(\varepsilon_t^2)$, alors

$$\mathbb{E}(\varepsilon_t^2)\left(1 - \sum_{i=1}^{q} \alpha_i\right) = \omega \qquad \Longrightarrow \qquad \mathbb{E}(\varepsilon_t^2) = \frac{\omega}{1 - \sum_{i=1}^{q} \alpha_i} = V(\varepsilon_t).$$

□

Propriétée 1.2.2. *[La variance conditionnelles] La variance conditionnelle d'un processus* $\{\varepsilon_t;\ t \in \mathbb{Z}\}$ *définie par l'équation* (1.1)
est :

$$V(\varepsilon_t|\mathcal{F}_{t-1}) = \mathbb{E}(\varepsilon_t^2|\mathcal{F}_{t-1}) = \sigma^2 \qquad \forall t \in \mathbb{Z}.$$

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

Propriétée 1.2.3. *[Les auto-covariances conditionnelles] Les auto-covariances conditionnelles du processus* $\{\varepsilon_t;\ t \in \mathbb{Z}\}$ *sont nulles. C'est à dire que :*

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

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

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

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

Propriétée 1.2.4. *[Moment centré d'ordre quatre]*

- *i) Le moment conditionnelle centré d'ordre quatre (4) du processus* ε_t *verifie*

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

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

$$\mathbb{E}(\varepsilon_t^4) = 3\left[\omega^2 + \frac{2\alpha\omega^2}{1-\alpha} + \alpha^2\mathbb{E}[\varepsilon_{t-1}^4]\right]$$

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

- *iii) Le Kortosis non conditionnelle associée au processus ARCH(1) est égale à :*

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

Preuve. i)et ii)-On sait que si une variable centrée X suit une loi normale et on a : $\varepsilon_t = \eta_t\sigma_t$ centrée, alors :

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

donc,

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

iii)- D'apres les résultats précédentes

$$
\begin{aligned}
K &= \tfrac{\mathbb{E}(\varepsilon_t^4)}{\mathbb{E}(\varepsilon_t^2)^2} = \tfrac{3\omega^2(1+\alpha)}{(1-3\alpha^2)(1-\alpha)}.\tfrac{(1-\alpha)^2}{\omega^2} \\
&= 3\left(\tfrac{1-\alpha^2}{1-3\alpha^2}\right) > 3.
\end{aligned}
$$

□

Remarque 1.2.1. *Le Kortosis d'un processus ARCH est toujours supérieur à 3, la loi non conditionnelle d'un processus ARCH est donc une loi de distribution à queue épaisse, est donc plus aplatie qu'une gaussienne, on dit que cette distribution est Leptokurtique (car $k > 3$).*

Remarque 1.2.2. *La moyenne non conditionnelle et les autocovariances d'un processus ARCH étant nulles, ce qui signifie que ce processus peut être caractérisé comme étant un processus bruit blanc faible.*

Après avoir décrit le processus ARCH et ses différentes propriétés, nous allons passer en revue le processus GARCH.

1.2.3 Les modèles ARCH généralisées : GARCH

Le processus GARCH (Generalized Auto Regressive Conditional Hetermskedasticity) a été introduit en 1986 par Bollerslev. Le processus GARCH est une extension du processus ARCH, il présente les mêmes propriétés et les mêmes fondements que le processus ARCH. Disons que la seule différente en situe au niveau de la définition. Le modèle GARCH a deux dimensions (p, q) alors que le modèle ARCH en a une (q).
Nous donnons une première définition d'un processus GARCH fondée sur les deux premiers moments de ε_t conditionnels à son passé.
GARCH faible (Weak GARCH)
Drost et Nijman [1993] ont convenu d'appeler GARCH faible "weak GARCH " tout bruit blanc faible $\varepsilon_t = \eta_t\sigma_t$, tel que :
i)- $\mathbb{E}[\varepsilon_t|\mathcal{F}_{t-1}] = 0$,$\forall t \in \mathbb{Z}$ C'est la propriété différence de martingale.
ii)- Il existe des constantes $\omega, \alpha_i, \beta_j$,$i = 1, \dots, q$,$j = 1, \dots, p$ telles que :

$$\sigma_t^2 = V(\varepsilon_t|\mathcal{F}_{t-1}) = \omega + \sum_{i=1}^{q} \alpha_i\varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j\sigma_{t-j}^2, \quad \forall t \in \mathbb{Z}.$$

GARCH (p, q) semi-fort
Lorsque le processus d'innovation $\{\nu_t;\ t \in \mathbb{Z}\}$, et ε_t^2 est lui même supposé être un bruit

blanc faible, alors ils appellent GARCH semi-fort "semi-strong GARCH " le même processus $\{\varepsilon_t;\ t \in \mathbb{Z}\}$ lorsqu'il s'agit d'une différence de martingale avec un processus d'innovation $\{\nu_t;\ t \in \mathbb{Z}\}$ qui est lui même une différence de martingale. Les processus GARCH semi-forts ainsi définis coïncident bien avec l'idée initiale de Engle et Bollerslev puisqu'il est clair réciproquement que si l'on suppose que $\{\nu_t\}$ est une différence de martingale, on en déduit que :

$$\nu_t = \varepsilon_t^2 - \sigma_t^2,$$

où σ_t^2 est bien la variance de $(\varepsilon_t)_{t\in\mathbb{Z}}$ conditionnelle à l'information passée.

Définition 1.2.2. *On dit que $\{\varepsilon_t;\ t \in \mathbb{Z}\}$ est un processus GARCH(p,q) si ses deux premiers moments conditionnels existent et vérifient :*

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

$$\sigma_t^2 = V(\varepsilon_t|\varepsilon_s, s < t) = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2, \qquad t \in \mathbb{Z}. \tag{1.3}$$

L'équation (1.3) *est équivalente à :*

$$\sigma_t^2 = \omega + \alpha(B)\varepsilon_t^2 + \beta(B)\sigma_t^2, \qquad t \in \mathbb{Z}, \tag{1.4}$$

où B est l'opérateur retard, $\alpha(B) = \sum_{i=1}^{q} \alpha_i B^i$ avec $B^i\varepsilon_t^2 = \varepsilon_{t-i}^2$ et $\beta(B) = \sum_{j=1}^{p} \beta_j B^j$ avec $B^j\sigma_t^2 = \sigma_{t-j}^2$.
Si $\beta(z) = 0$ on a :

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

est le processus ARCH(q). Le processus d'innovation pour ε_t^2 est par définition $\nu = \varepsilon_t^2 - \mathbb{E}\{\varepsilon_t^2|\varepsilon_j, j \leq t-1\} = \varepsilon_t^2 - \sigma_t^2$, qui vérifie $\mathbb{E}(\nu_t|\mathcal{F}_{t-1}) = 0$, on remplace σ_t^2 dans l'équation (1.3) *par $\varepsilon_t^2 - \nu_t$ on aura la structure linéaire d'un modèle ARMA :*

$$\varepsilon_t^2 - \nu_t = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j(\varepsilon_{t-j}^2 - \nu_{t-j}),$$

$$\Longrightarrow$$

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

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

Définition 1.2.3. ***[Processus GARCH (p,q)fort (strong GARCH)]*** *soit $\{\eta_t;\ t \in \mathbb{Z}\}$ une suite de variables* iid *de loi ($\eta_t \sim \mathrm{N}(0,1)$), on dit que $\{\varepsilon_t;\ t \in \mathbb{Z}\}$ est un GARCH(p,q) au sens fort s'il vérifie l'équation suivante :*

$$\begin{cases} \varepsilon_t = \sigma_t \eta_t \\ \sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2, \end{cases} \tag{1.5}$$

où $\omega > 0, \alpha_i > 0, \beta_j > 0, i = 0, \ldots, q, j = 1, \ldots, p$.
Maintenant en remplaçant ε_{t-i} *par* $\sigma_{t-i}\eta_{t-i}$ *dans l'équation* (1.5), *on obtient une représentation autorégressive de* σ_t^2 *à coefficients aléatoires :*

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

$$\Leftrightarrow$$

$$\sigma_t^2 = \omega + \textstyle\sum_{i=1}^{r} (\alpha_i \eta_{t-i}^2 + \beta_i)\sigma_{t-i}^2 = \omega + \sum_{i=1}^{r} \alpha_i(\eta_{t-i})\sigma_{t-i}^2,$$

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

1.2.4 Propriétés des processus GARCH

Les propriétés théoriques des processus GARCH se déduisent de la même façon que nous avons développé les propriétés des processus ARCH.

Propriétée 1.2.5. *le processus* $\{\varepsilon_t\}_{t\in\mathbb{Z}}$ *est un bruit blanc, si* $\mathbb{E}[\varepsilon_t^2] < \infty$.
on a :

$$\mathbb{E}[\varepsilon_t] = \mathbb{E}[\mathbb{E}(\varepsilon_t|\mathcal{F}_{t-1})] = 0$$

et

$$cov(\varepsilon_t, \varepsilon_{t-k}) = \mathbb{E}[\varepsilon_t \varepsilon_{t-k}]$$

$$= \mathbb{E}[\varepsilon_{t-k}\mathbb{E}(\varepsilon_t|\mathcal{F}_{t-1})] = 0 \quad ,\forall k > 0.$$

Propriétée 1.2.6. *Une condition nécessaire de l'existence de la variance d'un processus GARCH*(p, q) *est :*

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

Remarque 1.2.3. *Si cette condition est vérifiée avec les contraintes de non négativité donnée ci-dessus, elle est également suffisante. Donc le processus GARCH est faiblement stationnaire ou stationnaire au second ordre. Dans le cas où l'inégalité précédente est saturée, c'est à dire que :*
$\sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} \beta_j = 1$, *on dira que le processus GARCH est intégré, et on parlera de processus* ***IGARCH****.*

Propriétée 1.2.7. *Le processus* $\{\varepsilon_t;\ t \in \mathbb{Z}\}$ *d'une représentation GARCH*(p, q) *peut être représenté sous la forme d'un processus ARMA (max*(p, q) *; q) défini dans une innovation*

$\nu_t = \varepsilon_t^2 - \sigma_t^2$ *tel que :*

$$\sigma_t^2 = \omega + \sum_{i=1}^{r} (\alpha_i + \beta_i)\varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \nu_{t-j} + \nu_t \qquad r = max(p, q),$$

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

1.2.5 Etude de la stationnarité

Nous allons chercher sous quelles conditions il existe des processus stationnaires (au sens strict et au second-ordre) vérifiant les définitions (1.3) et/ou (1.5). On s'intéresse plus particulièrement aux solutions non anticipatives du modèle (1.5), c'est-à-dire aux processus $\{\varepsilon_t;\ t \in \mathbb{Z}\}$ tel que ε_t soit une fonction mesurable des variables $\{\eta_{t-s},\ s \geq 0\}$. Nous examinons d'abord le cas du modèle GARCH(1,1) qui peut se traiter avec des techniques élémentaires. On notera, pour $x > 0$, $\log^+ x = max(\log x, 0)$.

Modèles GARCH(1,1) :

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

$$\begin{cases} \varepsilon_t = \sigma_t \eta_t \\ \sigma_t^2 = \omega + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2, \end{cases} \tag{1.6}$$

en utilisant la représentation autorégressive on obtient :

$$\begin{cases} \varepsilon_t = \sigma_t \eta_t \\ \sigma_t^2 = \omega + a(\eta_{t-1}) \sigma_{t-1}^2, \end{cases} \tag{1.7}$$

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

Théorème 1.2.1. ***[La stationnarité stricte du modèle GARCH(1.1)]***

$$-\infty \leq \gamma := \mathbb{E}\left\{log\left(\alpha\eta_t^2 + \beta\right)\right\} < 0, \tag{1.8}$$

la série

$$h_t = \left\{1 + \sum_{i=1}^{\infty} a(\eta_{t-1}) \dots a(\eta_{t-i})\right\} \omega, \tag{1.9}$$

converge presque sûrement (p.s.) et le processus $\{\varepsilon_t;\ t \in \mathbb{Z}\}$ défini par $\varepsilon_t = \sqrt{h_t \eta_t}$ est l'unique solution strictement stationnaire du modèle (1.8). Cette solution est non anticipative et ergodique. Si $\gamma \geq 0$ et $\omega > 0$, il n'existe pas de solution strictement stationnaire.

Remarque 1.2.4. *1. $\gamma = \mathbb{E}\{\log a(\eta_t)\}$ existe toujours, car*

$$\mathbb{E}\big(log^+\{a(\eta_t)\}\big) \leq \mathbb{E}\{a(\eta_t)\} = \alpha + \beta.$$

2. Si $\alpha + \beta < 1 \implies \gamma < 0$, inversement si $\gamma < 0 \implies \beta < 1$,
car si $\alpha + \beta < 1$ alors, $\mathbb{E}\left\{log\left(\alpha\eta_t^2 + \beta\right)\right\} \leq log\left\{\mathbb{E}\left(\alpha\eta_t^2 + \beta\right)\right\} = log(\alpha + \beta) < 0$.
Inversement si $\gamma < 0$, alors par absurde supposons que $\beta \geq 1$ donc :
$\mathbb{E}\{log[\alpha\eta_t^2 + \beta]\} = \mathbb{E}\left\{log\left[\beta\left(\frac{\alpha}{\beta}\eta_t^2 + 1\right)\right]\right\} = \mathbb{E}\Big\{\log(\beta) + \log\left(\frac{\alpha}{\beta}\eta_t^2 + 1\right)\Big\} > 0$, d'où la contradiction.

3. *Dans le cas ARCH(1) ($\beta = 0$), la contrainte de stationnarité stricte peut s'écrire comme suit :*
$$0 \leq \alpha < \exp\left\{-\mathbb{E}\left(\log \eta_t^2\right)\right\}. \tag{1.10}$$
En effet,
$$\begin{aligned} -\infty \leq \gamma &= \left\{\mathbb{E}(\log \alpha + log\eta_t^2)\right\} < 0 \\ &\Leftrightarrow -\infty \leq log\ \alpha < -\mathbb{E}(log\ \eta_t^2) \\ &\Leftrightarrow 0 \leq \alpha < \exp\{-\mathbb{E}(log\ \eta_t^2)\}. \end{aligned}$$

4. *Dans le cas où $\omega = 0$ et $\gamma < 0$, il est clair d'après (1.9) que la seule solution strictement stationnaire du modèle est $\varepsilon_t = 0$.*

Preuve. En utilisant la deuxième équation du modèle (1.7), et par itération, on aura pour $N \geq 0$:

$$\begin{aligned} \sigma_t^2 &= \omega + a(\eta_{t-1})\sigma_{t-1}^2 \\ &= \omega\left\{1 + \textstyle\sum_{n=1}^{N} a(\eta_{t-1})\ldots a(\eta_{t-n})\right\} + a(\eta_{t-1}\ldots a(\eta_{t-N-1})\sigma_{t-N-1}^2 \\ &:= \quad h_t(N) + a(\eta_{t-1})\ldots a(\eta_{t-N-1})\sigma_{t-N-1}^2. \end{aligned} \tag{1.11}$$

Soit $h_t = \lim_{N\to\infty} h_t(N)$ existe dans $\overline{\mathbb{R}}^+ = [0, +\infty]$, puisque les termes sont positifs. De plus $h_t(N) = \omega + a(\eta_{t-1})h_{t-1}(N-1)$, donc quand $N \to +\infty$, on obtient :
$h_t = \omega + a(\eta_{t-1})h_{t-1}$.
Maintenant montrons que la processus h_t est presque sûrement finie si et suelement si $\gamma < 0$.
Supposons $\gamma < 0$.On utilise la règle de Cauchy pour les séries à termes positifs, on a :

$$\begin{aligned} \{a(\eta_{t-1})\ldots a(\eta_{t-n})\}^{1/n} &= \exp\left\{\frac{1}{n}\log(a(\eta_{t-1})\ldots a(\eta_{t-n})\right\} \\ &= \exp\left\{\tfrac{1}{n}\textstyle\sum_{i=1}^{n}\log\left[a(\eta_{t-i})\right]\right\} \\ &\to \exp\left\{\mathbb{E}[\log\left(\alpha\eta_{t-i}^2\right) + \beta]\right\} \longrightarrow e^{(\gamma)} \quad p.s. \end{aligned} \tag{1.12}$$

Par l'application de la loi forte des grands nombres, quand $n \to +\infty$, à la suite *iid* $(\log\{a(\eta_{t-i})\})$, la série (h_t) qui définie en (1.10)converge presque sûrement dans $\mathbb{R}$, par application de la règle de Cauchy, et le processus limite, (h_t) est à valeurs réelles positives.
Par suite, le processus $\{\varepsilon_t;\ t \in \mathbb{Z}\}$ définie par $\varepsilon_t = \sqrt{h_t\eta_t}$, est strictement stationnaire et ergodique (théorème d'ergodicité), et est non anticipatif car il s'écrit comme fonction mesurabe des variables $\eta_{t-i},\ i \geq 0$, de plus $\{\varepsilon_t;\ t \in \mathbb{Z}\}$ vérifie le modèle (1.6).
Maintenant nous montrons l'unicité :
soit $\tilde{\varepsilon}_t = \sigma_t\eta_t$ une autre solution strictement stationnaire (1.11) on a :

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

par suite

$$\sigma_t^2 - h_t = \{h_t(N) - h_t\} + a(\eta_{t-1}) \dots a(\eta_{t-N-1})\sigma_{t-N-1}^2, \qquad \forall N \in \mathbb{N}.$$

Lorsque $\gamma < 0$, le premier terme qui est entre accolade tend vers 0 p.s quand $N \to \infty$ donc le second terme (qui ne dépend pas de N) est nul car la série h_t converge presque p.s et de plus la loi de σ_{t-N-1}^2 ne dépend pas aussi de N par stationnarité. On montre alors que $\sigma_t^2 = h_t$ pour tout t, p.s.
Si $\gamma > 0$, d'après (1.12) et la règle de Cauchy, $\sum_{n=1}^{N} a(\eta_{t-1}) \dots a(\eta_{t-n}) \to +\infty$, p.s lorsque $N \to \infty$. donc si $\omega > 0$, $h_t = +\infty$, p.s. D'après (1.11) il est clair que alors $\sigma_t^2 = +\infty$ p.s. Par suite il n'existe pas de solution finie de (1.6).
Dans le cas $\gamma = 0$, nous procéderons par l'absurde. Supposons qu'il existe une solution strictement stationnaire $(\varepsilon_t, \sigma_t^2)$ de (1.6). Nous avons pour $n > 0$,

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

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

$$\sum_{i=1}^{n} \log a(\eta_i) + \log \omega \to -\infty \quad p.s. \quad quand \quad n \to \infty. \tag{1.13}$$

D'après le théorème de Chung-Fuchs nous avons $\limsup \sum_{i=1}^{n} \log a(\eta_i) = +\infty$ avec probabilité 1, ce qui contredit (1.13). $\square$

Théorème 1.2.2. *[La stationnarité au second ordre du modèle GARCH(1.1)] Si $\omega > 0$ et $\alpha + \beta < 1$, le processus $\{\varepsilon_t;\ t \in \mathbb{Z}\}$ admet une solution stationnaire au second ordre, de plus elle est unique, plus précisement $\{\varepsilon_t;\ t \in \mathbb{Z}\}$ est un bruit blanc.*
Si $\alpha + \beta \geq 1$, il n'existe pas de solution non anticipative stationnaire au second ordre.

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

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

soit

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

Il faut donc $\alpha + \beta < 0$. On obtient de plus : $\mathbb{E}(\varepsilon_t^2) > 0$.
Inversement si $\alpha + \beta < 1$ on a $\gamma < 0$, alors la solution strictement stationnaire vérifie :

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

$\square$

Modèle GARCH (p,q)

Dans le cas général du GARCH(p,q) fort, la représentation vectorielle sera très utile. On a :

$$Z_t = b_t + A_t Z_{t-1}, \tag{1.14}$$

où
$Z'_t = (\varepsilon_t^2, \varepsilon_{t-1}^2, \ldots, \varepsilon_{t-q+1}^2, \sigma_t^2, \ldots, \sigma_{t-p+1}^2) \in \mathbb{R}^{p+q}$.
$b'_t = (\omega\eta_t^2, 0, \ldots, 0, \omega, 0, \ldots, 0) \in \mathbb{R}^{p+q}$.

$$A_t = \begin{pmatrix} \alpha_1\eta_t^2 & \cdots & \alpha_q\eta_t^2 & \beta_1\eta_t^2 & \cdots & \beta_p\eta_t^2 \\ & I_{q-1} & 0 & & 0 & \\ \alpha_1 & \cdots & \alpha_q & \beta_1 & \cdots & \beta_p \\ & 0 & & & I_{p-1} & 0 \end{pmatrix}, \tag{1.15}$$

A_t est une matrice de dimension $(p+q) \times (p+q)$. Dans le cas ARCH(q), Z_t ne contient que ε_t^2 et ses $q-1$ premières valeurs passées, et A_t se limite au bloc supérieur gauche de la matrice ci-dessus. L'équation (1.14) constitue un modèle vectoriel autorégressif d'ordre un, avec coefficients positifs et *iid*.
Si on itérant le modèle (1.14) on obtient :

$$Z_t = b_t + \sum_{k=1}^{\infty} A_t A_{t-1} \ldots A_{t-k+1} b_{t-k}. \tag{1.16}$$

Sous réserve que la série existe au sens presque sûr. L'objet de ce qui suit est de trouver des conditions justifiant l'existence de cette série. Lorsque le membre de droite de l'équation (1.16) a un sens, cela n'assure pas pour autant que les composantes de ce vecteur sont positives. Une condition suffisante pour que, presque sûrement,

$$b_t + \sum_{k=1}^{\infty} A_t A_{t-1} \ldots A_{t-k+1} b_{t-k} > 0, \tag{1.17}$$

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

$$\omega > 0, \quad \alpha_i \geq 0 \quad (i = 1, \ldots, q), \quad \beta_j \geq 0 \quad (j = 1, \ldots, p). \tag{1.18}$$

Stationnarité stricte

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

$$\lim_{t\to\infty} \frac{1}{t} \log \|A^t\| = \log \rho(A), \quad t \in \mathbb{N}. \tag{1.19}$$

Théorème 1.2.3. ***[Exposant de Lyapounuv]*** *Soit $\{A_t;\ t \in \mathbb{Z}\}$ une suite de matrices aléatoires, strictement stationnaire et ergodique, telle que $\mathbb{E}\log^+ \|A_t\| < \infty$. On a :*

$$\lim_{t\to+\infty} \frac{1}{t}\mathbb{E}(\log \|A_t A_{t-1} \dots A_1\|) = \gamma = \inf_{t\in\mathbb{N}^*} \frac{1}{t}\mathbb{E}(\log \|A_t A_{t-1} \dots A_1), \tag{1.20}$$

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

$$\gamma = \lim_{t\to+\infty}\ p.s. \frac{1}{t} \log \|A_t A_{t-1} \dots A_1\|. \tag{1.21}$$

γ : l'exposant de Lyapounuv.
$\exp\{\gamma\}$: Rayon spectral de la suite $\{A_t;\ t \in \mathbb{Z}\}$.

Remarque 1.2.5.
1. *On a toujours $\gamma \leq \mathbb{E}(\log \|A_1\|)$, avec égalité en dimension 1.*
2. *Si $A_t = A$ pour tout $t \in \mathbb{Z}$, on a $\gamma = \log \rho(A)$ d'après (1.19).*
3. *Toutes les normes étant équivalentes sur un espace de dimension fini, il est facile de voir que γ est indépendant du choix de la norme.*

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

Lemme 1.2.1. *Soit $\{A_t,\ t \in \mathbb{Z}\}$ une suite de matrices aléatoires* iid *telle que $\mathbb{E}\log^+ \|A_t\|$ est finie et de plus grand exposant de Lyapounuv γ.*
Alors

$$\lim_{t\to\infty} p.s. \|A_0 \dots A_{-t}\| = 0 \quad \Rightarrow \quad \gamma < 0. \tag{1.22}$$

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

Théorème 1.2.4. ***[La stationnarité stricte du modèle GARCH(p,q)]*** *Une condition nécessaire et suffisante d'exsitence d'un processus GARCH(p,q) strictement stationnaire, solution du modèle (1.5) est que*

$$\gamma < 0,$$

où γ est plus grand exposant de Lyapounuv de la suite $\{A_t; t \in \mathbb{Z}\}$ définie par (1.15).
Lorsqu'elle existe, la solution strictement stationnaire est unique, non anticipative et ergodique.

Preuve. Nous utiliserons la norme définie par $\|A\| = \sum |a_{ij}|$. Par commodité la norme sera notée de manière identique quelle que soit la dimension de A. Avec cette convention, la norme est clairement multiplicative : $\|AB\| \leq \|A\|\|B\|$ pour toutes matrices A et B telles que AB existe.
Remarquons que, les variables η_t étant de variance finie, tous les termes de la matrice A_t sont intégrables. On a donc

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

Supposons $\gamma < 0$. Alors, l'égalité (1.21) implique que la série :

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

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

$$\|Z_t\| \leq \|b_t\| + \sum_{n=1}^{\infty} \|A_t A_{t-1} \dots A_{t-n}\| \|b_{t-n-1}\| \tag{1.23}$$

et

$$\left(\|A_t A_{t-1} \dots A_{t-n}\| \|b_{t-n-1}\|\right)^{1/n} = \exp\left\{\tfrac{1}{n}\log\left(\|A_t A_{t-1} \dots A_{t-n}\|\right) + \tfrac{1}{n}\log\|b_{t-n-1}\|\right\}$$

$$\to^{p.s} \exp(\gamma) < 1 \qquad (\text{car } \ \mathbb{E}|\log\|b_{t-n-1}\|| < \infty).$$

Par application de la règle de Cauchy, $\gamma < 0$ implique Z_t est bien définie. Soit $Z_{q+1,t}$ la $q+1$-ème composante de Z_t. En posant $\varepsilon_t = \sqrt{Z_{q+1,t}}\eta_t$ on définit une solution strictement stationnaire du modèle (1.5). D'après (1.16) ε_t s'exprime comme fonction mesurable de $\eta_t, \eta_{t-1}, \dots$. La solution est donc non anticipative et ergodique puisque (η_t est ergodique.
L'unicité se démontre par le même raisonnement que dans le cas $p = q = 1$.
Supposons qu'il existe une autre solution strictement stationnaire Z_t^* du modèle (1.14).
Alors, pour tout $n \geq 0$,

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

Par absurde, on a $\mathbb{P}(\|Z_t - Z_t^*\| \neq 0) > 0$. Or on sait que $\|A_t \dots A_{t-n}\| \to 0 p.s.$ quand $n \to \infty$ car la série dans (1.23) converge.
par suite

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

ce qui implique que $\|Z_{t-n-1}\| \to \infty$ ou $\|Z_{t-n-1}^*\| \to \infty$ avec une probabilité positive. Ceci est impossible car les suites $(Z_t)_{t\in\mathbb{Z}}$ et $(Z_t^*)_{t\in\mathbb{Z}}$ sont stationnaires. On en conclut que $Z_t = Z_t^*$ pour tout t, p.s.
Finalement nous montrons la partie nécessaire du théorème. D'après le lemme(1.2.1), il suffit d'établir (1.22). Nous allons montrer que, pour $1 \leq i \leq p+q$

$$\lim_{t\to\infty} A_0 \dots A_{-t} e_i = 0 \quad p.s. \tag{1.24}$$

Où e_i est le i-ème élément de la base canonique de $\mathbb{R}^{p+q}$. Soit (ε_t) une solution strictement stationnaire de (1.5) et soit (Z_t) défini par (1.14). On a pour $t > 0$,

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

car les coefficients des matrices A_t, b_0 et z_t sont positifs. La séries $\sum_{k=0}^{t-1} A_0 \dots A_{-k} b_{-k-1}$ converge et donc $A_0 \dots A_{-k} b_{-k-1}$ tend presque sûrement vers 0 quand $k \to \infty$. Or $b_{-k-1} = \omega\eta^2_{-k-1} e_1 + \omega e_{q+1}$. Donc $A_0 \dots A_{-k} b{-}k-1$ se décompose en deux termes positifs et on a :

$$\lim_{k\to\infty} A_0 \dots A_{-k} \omega\eta^2_{-k-1} e_1 = 0, \qquad \lim_{t\to\infty} A_0 \dots A_{-k} \omega e_{q+1} = 0, \quad p.s. \tag{1.25}$$

Puisque $\omega \neq 0$, (1.24) est vraie pour $i = q+1$. En utilisant la relation suivante :

$$A_k e_{q+i} = \beta_i \eta^2_{-k} e_1 + \beta_i e_{q+1} + e_{q+i+1}, \qquad i = 1, \dots, p, \tag{1.26}$$

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

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

donc (1.24) est vraie pour $i = q+2$, et par récurrence, pour $i = q+j$, $j = 1, \dots, p$ en utilisant (1.26). Par ailleurs, on remarque que $A_{-k} e_i = \alpha_q \eta^2_{-k} e_1 + \alpha_q e_{q+1}$, ce qui permet de voir, d'après (1.25), que (1.24) est vérifiée pour $i = q$. On conclut pour les autres valeurs de i en utilisant :

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

et une récurrence ascendante. Le théorème (1.3.4) est donc démontré. $\square$

Corollaire 1.2.1. ***[Conséquences de stationnarité stricte]*** *Soit γ le plus grand exposant de Lyapounov de la suite $\{A_t;\ t \in \mathbb{Z}\}$ définie par (1.15). Si $\gamma < 0$ nous avons les propriétés suivantes équivalentes :*

- *i) $\sum_{j=1}^{p} \beta_j < 1$,*
- *ii) $1 - \beta_1 z - \dots - \beta_p z^p = 0 \Rightarrow |z| > 1$,*
- *iii) $\rho(B) < 1$, où B est la sous-matrice de A_t définie par :*

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

Preuve. Comme tous les termes des matrices A_t sont positifs, il est clair que γ est supérieur au coefficient de Lyapounov de la suite obtenue en remplaçant les coefficients des q premières lignes et des q premières colonnes par 0 dans les matrices A_t. En utilisant la Remarque 2 du Théorème (1.3.3) on voit que

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

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

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

où $\mathcal{B} = 1 - \beta_1 z - \cdots - \beta_p z^p$. On déduit que si $\gamma < 0$ alors $\mathcal{B}(z) = 0$ a toutes ses racines en dehors du cercle unité, d'où l'équivalence entre (ii)et (iii).
Maintenant, on va montrer que $(i) \Leftrightarrow (ii)$. On a $\mathcal{B}(0) = 1 - \sum_{j=1}^{p} \beta_j$ donc si $\sum_{j=1}^{p} \beta_j \geq 1$, alors $\mathcal{B}(1) \leq 0$ et par continuité, il existe une dans [0,1]. Alors $(ii) \Rightarrow (i)$.
Inversement si $\sum_{j=1}^{p} \beta_j < 1$ et si $\mathcal{B}(z_0) = 0$ pour un z_0 de module inférieur ou égal à 1 alors,

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

Exemples 1. *Dans le cas GARCH(1,1), on retrouve bien la condition de stationnarité stricte obtenue directement. La matrice A_t s'écrit dans ce cas*

$$A_t = \begin{pmatrix} \eta_t^2 \\ 1 \end{pmatrix} \begin{pmatrix} \alpha_1 & \beta_1 \end{pmatrix}.$$

On a donc

$$A_t A_{t-1} \dots A_1 = \prod_{k=1}^{t-1} (\alpha_1 \eta_{t-k}^2 + \beta_1) A_t.$$

par suite

$$\log \|A_t A_{t-1} \dots A_1\| = \sum_{k=1}^{t-1} \log (\alpha_1 \eta_{t-k}^2 + \beta_1) + \log \|A_t\|$$

et d'après (2.21) et la loi forte des grandes nombres $\gamma = E \log (\alpha_1 \eta_t^2 + \beta_1)$. La condition nécessaire et suffisante de stationnarité stricte s'écrit donc $E \log (\alpha_1 \eta_t^2 + \beta_1) < 0$, comme nous le savions déjà.

1.3 Existence des moments d'ordre 2s

Nous concluons cette partie avec un résultats établissant que la condition de stationnarité stricte implique également l'existence de certains moments. Nous montrons au préable le lemme suivant :

Lemme 1.3.1. *Soit X une v.a.r. presque sûrement positive. Si $\mathbb{E}(X^r) < \infty$ pour un $r > 0$ et si $\mathbb{E}(\log X) < 0$ alors il existe $s > 0$ tel que $\mathbb{E}(X^s) < 1$.*

Preuve. La fonction génératrice des moments de $Y = \log X$ est définie par $M(u) = \mathbb{E}e^{uY} = \mathbb{E}X^u$. La fonction M est continuement dérivable sur $[0, r]$ et on a, pour $u > 0$

$$\frac{M(u) - M(0)}{u} = \int \frac{e^{uy} - 1}{u} dP_Y(y). \tag{1.27}$$

Remarquons que :

$$\forall \tau > 0, \quad \forall u \in [0, \tau], \quad \left| \frac{e^{uy} - 1}{u} \right| \leq \frac{e^{\tau|y|}}{\tau}. \tag{1.28}$$

Ce résultat s'obtient par exemple en introduisant la fonction définie par $g(v) = \frac{e^v - 1}{v}$ pour $v \neq 0$ et $g(0) = 1$. La fonction g étant croissante sur $\mathbb{R}$, on a pour $y \geq 0$,

$$\frac{e^{uy} - 1}{u} \leq \frac{e^{\tau y} - 1}{\tau} \leq \frac{e^{\tau|y|}}{\tau},$$

et pour $y < 0$,

$$\frac{1 - e^{uy}}{u} \leq -y \leq \frac{e^{-\tau y}}{\tau},$$

ce qui prouve (1.28). le membre de droite de cette inégalité est clairement P_Y-intégrable quand $\tau \in]0, r]$. Par suite, par le théorème de Lebesgue, la dérivée à droite de M en 0 est d'après (1.27)

$$\int y dP_Y(y) = \mathbb{E}(\log X) < 0.$$

Comme $M(0) = 1$, il existe $s > 0$ tel que $M(s) = \mathbb{E}(X^s) < 1$. □

Lemme 1.3.2. *Soit $\{A_t\}$ est une suite de matrices positives, γ l'exposant de Lyapounov. Alors*

$$\gamma < 0 \quad \Leftrightarrow \quad \exists s > 0, \quad \exists k_0 > 1, \delta := \mathbb{E}(\|A_{k_0} A_{k_0-1} \ldots A_1\|^s) < 1.$$

Preuve. Supposons $\gamma < 0$. Comme $\gamma = \inf_{k \in \mathbb{Z}} \frac{1}{k} \mathbb{E}(\log \|A_k A_{k-1} \ldots A_1\|) < 0$, il existe $k_0 \geq 1$ tel que $\mathbb{E}(\log \|A_{k_0} A_{k_0-1} \ldots A_1\|) < 0$. De plus,

$$\begin{aligned}
\mathbb{E}(\|A_{k_0} A_{k_0-1} \ldots A_1\|) &= \|\mathbb{E}(A_{k_0} A_{k_0-1} \ldots A_1)\| \\
&= \|(\mathbb{E}A_1)^{k_0}\| \\
&= (\mathbb{E}\|A_1\|)^{k_0} < \infty,
\end{aligned}$$

en utilisant la norme multiplicative $\|A_\| = \sum_{i,j} |A(i,j)|$, la positivité des éléments des A_i, indépendance et l'équidistribution des A_i. Le lemme (1.4.1) entraîne donc l'existence d'un $s > 0$ et $k_0 \geq 1$ tel que $\delta < 1$, en appliquant l'inégalité de Jensen, on obtient :

$$\gamma \leq \frac{1}{k_0} \mathbb{E}(\log \|A_{k_0} A_{k_0-1} \ldots A_1\|) \leq \frac{1}{sk_0} \log \delta < 0.$$

□

Corollaire 1.3.1. *Soit l'exposant de Lyapounov,*

$$\gamma < 0 \Rightarrow \exists s > 0, \quad \mathbb{E}(\sigma_t^{2s}) < \infty, \quad \mathbb{E}(\varepsilon_t^{2s}) < \infty,$$

où $\varepsilon_t = \sigma_t \eta_t$ est un processus GARCH(p,q) solution strictement stationnaire.

Preuve. Pour $s \in]0,1[,\ a, b > 0$, on a $\left(\frac{a}{a+b}\right)^s + \left(\frac{b}{a+b}\right)^s \geq 1$, et par conséquent $\left(\sum_i u_i\right)^s \leq \sum_i {u_i}^s$ pour toute suite de nombres positifs u_i. En le fait que la norme est multiplicative. La solution stationnaire est définie par (1.16) est satisfait

$$\mathbb{E}\|Z_t\|^s \leq \|\mathbb{E}b_1\|^s \left\{1 + \sum_{k=0}^{\infty} \delta^k \sum_{i=1}^{k_0} \{\mathbb{E}\|A_1\|^s\}^i \right\} < \infty.$$

On conclut en remarquant que $\sigma_t^{2s} \leq \|Z_t\|^s$ et $\varepsilon_t^{2s} \leq \|Z_t\|^s$. □

1.4 Extension des modèles ARCH et GARCH

Cette section represente quelques processus de type GARCH améliorés qui sont très utilisés dans le domaine finance.

1.4.1 Modèles GARCH/ ARCH Asymétrie

Cette partie couvre les modèles ARCH non-linéaires et plus particulièrement la prise en compte des phénomènes asymétrie. l'idée est tous simplement est que l'effet hétéroscédastique est sans doute pas le même suivant que l'erreur est positive ou négative. Deux grandes classes de modèles ont été proposées :
-Nelson [1990] s'est intéressé aux évolutions asymétriques de la variance à l'aide des modèles EGARCH .
-Engle et Bollerslev [1986] ont étudié les modèles ARCH à seuils (TGARCH) où la variance est une fonction linéaire définie par morceaux qui permet différentes fonctions de volatilité selon le signe et la valeur des chocs. Rabemananjara et Zakoïan [1991] ont proposé une généralisation avec les modèles TGARCH.

Définition 1.4.1. ***EGARCH*** *[Exponential GARCH] Un processus $\{\varepsilon_t;\ t \in \mathbb{Z}\}$ satisfait une representation EGARCH si et seulement si :*

$$\begin{cases} \varepsilon_t = \sigma_t \eta_t \\ \log(\sigma_t^2) = \omega + \sum_{i=1}^{q} \alpha_i g(\eta_{t-i}) + \sum_{j=1}^{p} \beta_j \log(\sigma_{t-j}^2), \end{cases} \tag{1.29}$$

où $\{\eta_t;\ t \in \mathbb{Z}\}$ est une suite de variables iid *telles que $\mathbb{E}(\eta_t) = 0$ et $\mathbf{V}(t) = 1$, et la fonction $g(.)$ vérifie :*

$$g(\eta_{t-i}) = \theta \eta_{t-i} + \gamma[|\eta_{t-i}| - \mathbb{E}|\eta_{t-i}|], \tag{1.30}$$

tel que :

ω, α_i, β_j, θ et γ sont des réels.
Si on remplace la fonction g dans l'équation (1.29), si on pose $a_i = \alpha_i\theta$ et $b_i = \alpha_i\gamma$, la variance conditionnelle de ε_t s'écrire sous la forme :

$$\log(\sigma_t^2) = \omega + \sum_{i=1}^{q} a_i\eta_{t-i} + \sum_{i=1}^{q} b_i(|\eta_{t-i}| - \mathbb{E}(|\eta_{t-i}|)) + \sum_{j=1}^{p} \beta_j \log(\sigma_{t-j}^2). \tag{1.31}$$

Définition 1.4.2. ***[Modèles GARCH à seuil (TGARCH)]*** *Soit $\{\eta_t;\ t \in \mathbb{Z}\}$ une suite de variable* iid *telles que $\mathbb{E}(\eta_t) = 0$ et $var(\eta_t) = 1$. On dit que $\{\varepsilon_t;\ t \in \mathbb{Z}\}$ est un processus GARCH à seuil (Threshold GARCH (p, q)) s'il vérifie une équation de la forme :*

$$\begin{cases} \varepsilon_t = \sigma_t\eta_t \\ \sigma_t = \omega + \sum_{i=1}^{q} \alpha_{i,+}\mathbb{I}_{\varepsilon_{t-i}\geq 0}\varepsilon_{t-i} + \alpha_{i,-}\mathbb{I}_{\varepsilon_{t-i}<0}\varepsilon_{t-i} + \sum_{j=1}^{p} \beta_j\sigma_{t-j}, \end{cases} \tag{1.32}$$

où $\omega, \alpha_{i,+}, \alpha_{i,-}$ et β_j sont des réels,
$\mathbb{I}_{\varepsilon_{t-i}<0}$ désigne la fonction indicatrice telle que $\mathbb{I}_{\varepsilon_{t-i}<0} = 1$ si $\varepsilon_{t-i} < 0$, et $\mathbb{I}_{\varepsilon_{t-i}<0} = 0$ sinon, et $\mathbb{I}_{\varepsilon_{t-i}\geq 0} = 1$ si $\varepsilon_{t-i} \geq 0$, et $\mathbb{I}_{\varepsilon_{t-i}\geq 0} = 0$ sinon.

Remarque 1.4.1. *Sous les contraintes $\omega > 0$, $\alpha_{i,+} \geq 0$, $\alpha_{i,-} \geq 0$ et $\beta_j \geq 0$ la variable σ_t est toujours strictement positive et s'interprète comme l'écart-type conditionnel de $\{\varepsilon_t;\ t \in \mathbb{Z}\}$.*

1.4.2 Processus ARCH-M

Le processus ARCH-M (ARCH dans Mean), proposé par Engle , Lilien et Robins (1987) pour estimer les variances conditionnelles dans GARCH et permet de prendre en compte l'influence de la volatilité du processus dans les variations en niveau du processus. En fait, la variance conditionnelle se trouve être une variable explicative de la moyenne conditionnelle du processus. C'est à dire que le modèle linéaire s'écrira ici :

$$\begin{cases} Y_t = X_t\beta + \delta g(\sigma_t^2) + \varepsilon_t \\ \sigma_t^2 = \omega + \alpha\varepsilon_{t-1}^2, \end{cases} \tag{1.33}$$

où $\{X_t;\ t \in \mathbb{Z}\}$ est un vecteur de variables explicatives, δ est un scalaire, et $\{\varepsilon_t;\ t \in \mathbb{Z}\}$ est un processus ARCH(q) tel que $\varepsilon_t = \sigma_t\eta_t$, de variance conditionnelle σ_t^2, et $Y_t|\mathcal{F}_{t-1} \rightarrow \mathcal{N}(\mu_t, \sigma_t^2)$, avec $\mu_t = X_t\beta + \delta g(\sigma_t^2)$.
La fonction g est souvent choisie parmi la fonction identité, la fonction carré ou la fonction logarithme.

Remarque 1.4.2. *Il existe de nombreuses autres extensions des processus ARCH/GARCH. On citera, entre autres, le processus IGARCH(p, q) [GARCH Intégré], introduit et étudié par Engle et Bollerslev (1986) et Nelson (1990), FIGARCH(p, d, q) qui est introduit par Baillie, Bollerslev et Mikkelsen (1996), qui sont des processus de type longue mémoire, on citere encore les processus GJR-ARCH et GJR-GARCH, et Q-GARCH (Q pour Quadratique) il a été introduit par Engle et Ng (1993) et sentana (1995), qui sont des modèles asymétriques.*

1.5 Représentation ARCH(∞)

On dit qu'un processus $\{\varepsilon_t;\ t \in \mathbb{Z}\}$ est un ARCH(∞) s'il existe une suite des variables $\{\eta_t;\ t \in \mathbb{Z}\}$ *iid* tels que $\mathbb{E}(\eta_t) = 0$ et $\mathbb{E}(\eta_t^2) = 1$ et une suite de constante $\phi_i \geq 0 \quad i = 1, \ldots,$ et $\phi_0 > 0$ tel que :

$$\begin{cases} \varepsilon_t = \sigma_t \eta_t \\ \sigma_t^2 = \phi_0 + \sum_{i=1}^{\infty} \phi_i \varepsilon_{t-i}^2. \end{cases} \tag{1.34}$$

Cette classe contient évidement le processus ARCH(q). En effet, il suffit de poser $\phi_i = 0$ pour $i \geq q+1$.
Cette classe est plus générale et peut contenir les GARCH(p, q).

1.5.1 Conditions d'existence des processus ARCH(∞)

L'existence des processus ARCH(∞) exige des hypothèses sur la suite des scalaires (ϕ_i) et (η_t). Le théorème suivant donne les conditions d'existence.

Théorème 1.5.1. ***[Existence d'une solution stationnaire d'un ARCH(∞)]*** *Pour tout $s \in]0, 1]$, posons*

$$A_s = \sum_{i=1}^{\infty} \phi_i^s \ , \qquad \mu_{2s} = \mathbb{E}|\eta_t|^{2s}.$$

S'il existe $s \in]0, 1]$ tel que

$$A_s \mu_{2s} < 1,$$

alors (1.34) admet une solution strictement stationnaire et non anticipative donnée par :

$$\begin{cases} \varepsilon_t = \sigma_t \eta_t \\ \sigma_t^2 = \phi_0 + \phi_0 \sum_{k=1}^{\infty} \ \sum_{i_1, i_2, \ldots, i_k \geq 1} \phi_{i_1} \ldots \phi_{i_k} \eta_{t-i_1}^2 \ldots \eta_{t-i_1-\cdots-i_k}^2. \end{cases} \tag{1.35}$$

Le processus $\{\varepsilon_t;\ t \in \mathbb{Z}\}$ défini par (1.35) est unique solution strictement stationnaire et non anticipative du modèle (1.34) tel que : $\mathbb{E}|\varepsilon_t|^{2s} < \infty$.

Preuve. Soit S_t la variable aléatoire définie par :

$$S_t = \phi_0 + \phi_0 \sum_{k=1}^{\infty} \sum_{i_1, i_2, \ldots, i_k \geq 1} \phi_{i_1} \ldots \phi_{i_k} \eta_{t-i_1 \ldots}^2 \eta_{t-i_1}^2 \ldots \eta_{t-i_1-\ldots-i_k}^2. \tag{1.36}$$

Comme tous les termes de (1.36) sont positifs en utilisant l'inégalité $(a+b)^s \leq a^s + b^s$, pour $a, b \geq 0$, on obtient :

$$S_t^s \leq \phi_0^s + \phi_0^s \sum_{k=1}^{\infty} \sum_{i_1, i_2, \ldots i_k} \phi_{i_1}^s \ldots \phi_{i_k}^s \eta_{t-i_1}^{2s} \ldots \eta_{t-i_1-\cdots-i_k}^{2s}.$$

L'indépendances de η_t nous assure

$$\mathbb{E}(S_t^2) \leq \phi_0^s + \phi_0^s \sum_{k=1}^{\infty} \sum_{i_1,i_2,\ldots,i_k \geq 1} \phi_{i_1}^s \ldots \phi_{i_k}^s \mathbb{E}(\eta_{t-i_1}^{2s} \ldots \eta_{t-i_1-\cdots-i_k}^{2s})$$

$$\leq \phi_0^s \left[1 + \textstyle\sum_{k=1}^{\infty} (A_s \mu_{2s})^k \right] = \tfrac{\phi_0^s}{1-A_s\mu_{2s}},$$

ce qui montre que $S_t < \infty$ p.s.
Tous les termes de la somme étant positifs, on a :

$$\sum_{i=1}^{\infty} \phi_i S_{t-i} \eta_{t-i}^2 = \sum_{i_0=1}^{\infty} \phi_{i_0} \eta_{t-i_0}^2 \left(\phi_0 + \phi_0 \sum_{k=1}^{\infty} \sum_{i_1,i_2,\ldots,i_k \geq 1} \phi_{i_1} \ldots \phi_{i_k} \eta_{t-i_0-i_1}^2 \ldots \eta_{t-i_0-i_1-\cdots-i_k}^2 \right)$$

$$= \phi_0 \textstyle\sum_{i_0=1}^{\infty} \phi_{i_0} \eta_{t-i_0}^2 + \phi_0 \sum_{i_0=1}^{\infty} \phi_{i_0} \eta_{t-i_0}^2 \sum_{k=1}^{\infty} \sum_{i_1,i_2,\ldots,i_k \geq 1} \phi_{i_1} \ldots \phi_{i_k} \eta_{t-i_0-i_1}^2 \ldots \eta_{t-i_0-i_1-\cdots-i_k}^2$$

$$= \phi_0 \textstyle\sum_{k=0}^{\infty} \sum_{i_0,i_1,\ldots,i_k \geq 1} \phi_{i_0} \ldots \phi_{i_k} \eta_{t_{i_0}}^2 \eta_{t-i_0-i_1}^2 \ldots \eta_{t_{i_0}-i_1-\cdots-i_k}^2.$$

Par conséquent, $S_t = \phi_0 + \sum_{i=1}^{\infty} S_{t-i} \eta_{t-i}^2$.
La solution strictement stationnaire du modèle (1.34) est obtenue par $\varepsilon_t = (S_t)^{1/2} \eta_t$, de plus $\mathbb{E}|\varepsilon_t|^{2s} \leq \frac{\mu_{2s}\phi_0^s}{1-A_s\mu_{2S}}$.
Pour la deuxième partie du théorème, on a :
$\mathbb{E}|\varepsilon_t|^{2s} = \mathbb{E}|(S_t)^{1/2}\eta_t|^{2s} \leq \mathbb{E}|S_t^s| \ \mathbb{E} \ |\eta_t^{2s}|$

$$\leq \mu_{2s} \mathbb{E}|S_t^s| \leq \tfrac{\mu_{2s}\phi_0^s}{1-A_s\mu_{2s}}.$$

A présent, montrons l'unicité de la solution.
Soit (ε_t) la solution strictement stationnaire et non anticipative du modèle (1.34) sachant que $\mathbb{E}|\varepsilon_t|^{2s} < \infty$, pour $q \geq 1$ et par itération des ε_{t-i}^2 on obtient :

$$\sigma_t^2 = \phi_0 + \sum_{i=1}^q \phi_i\varepsilon_{t-i}^2$$

$$= \phi_0 + \sum_{i=1}^q \phi_i\eta_{t-i}^2\sigma_{t-i}^2$$

$$= \phi_0 + \sum_{i_1=1}^q \phi_{i_1}\eta_{t-i_1}^2 \left(\phi_0 + \sum_{i_2=1}^q \phi_{i_2}\eta_{t-i_1-i_2}^2\sigma_{t-i_1-i_2}^2\right)$$

$$= \phi_0 + \phi_0\sum_{i_1=1}^q\phi_{i_1}\eta_{t-i_1}^2 + \sum_{i_1=1}^q \phi_{i_1}\eta_{t-i_1}^2 \sum_{i_2=1}^q \phi_{i_2}\eta_{t-i_1-i_2}^2\sigma_{t-i_1-i_2}^2$$

$$= \phi_0+\phi_0\sum_{i_1=1}^q\phi_{i_1}\eta_{t-i_1}^2+\sum_{i_1=1}^q \phi_{i_1}\eta_{t-i_1}^2 \sum_{i_2=1}^q \phi_{i_2}\eta_{t-i_1-i_2}^2\left(\phi_0+\sum_{i_3=1}^q \phi_{i_3}\eta_{t-i_1-i_2-i_3}^2\sigma_{t-i_1-i_2-i_3}^2\right)$$

$$= \phi_0 + \phi_0 \sum_{i_1=1}^q \phi_{i_1}\eta_{t-i_1}^2 + \phi_0 \sum_{i_1=1}^q \phi_{i_1}\eta_{t-i_1}^2 \sum_{i_2=1}^q \phi_{i_2}\eta_{t-i_1-i_2}^2$$

$$+ \sum_{i_1=1}^q \phi_{i_1}\eta_{t-i_1}^2 \sum_{i_2=1}^q \phi_{i_2}\eta_{t-i_1-i_2}^2 \sum_{i_3=1}^q \phi_{i_3}\eta_{t-i_1-i_2-i_3}^2\sigma_{t-i_1-i_2-i_3}^2$$

$$\vdots$$

$$= \phi_0 + \phi_0 \sum_{k=1}^q \sum_{i_1,i_2,\ldots,i_k\geq 1} \phi_{i_1}\phi_{i_2}\ldots\phi_{i_k}\eta_{t-i_1}\ldots\eta_{t-i_1-i_2-\cdots-i_k}$$

$$+ \sum_{i_1,i_2,\ldots,i_k\geq 1} \phi_{i_1}\phi_{i_2}\ldots\phi_{i_{q+1}}\eta_{t-i_1}\ldots\eta_{t-i_1-i_2-\cdots-i_q}\varepsilon_{t-i_1-i_2-\cdots-i_{q+1}}$$

$$= S_{t,q} + R_{t,q}.$$

Quand $q \to \infty$, $S_{t,q} \to S_t$ p.s. De plus, la solution non anticipative, ε_t est indépendant de η_t', pour tout $t' > t$, par conséquent :

$$\mathbb{E}\big(R_{t,q}^s\big) \leq \sum_{i_1,i_2,\ldots,i_{q+1}\geq 1} \phi_{i_1}^s\phi_{i_2}^s\ldots\phi_{i_{q+1}}^s\mathbb{E}\big(|\eta_{t-i_1}|^{2s}\ldots|\eta_{t-i_1-i_2-\cdots-i_q}|^{2s}|\varepsilon_{t-i_1-i_2-\cdots-i_{q+1}}|^{2s}\big)$$

$$\leq (A_s\mu_{2s})^q A_s\mathbb{E}|\varepsilon_t|^{2s} < \infty.$$

Ainsi, $\sum_{q\geq 1}\mathbb{E}\big(R_{t,q}^s\big) < \infty$, car $A_s\mu_{2s} < 1$ et $\mathbb{E}|\varepsilon_t|^{2s} < \infty$. Finalement, $R_{t,q} \to 0$ p.s. quand $q \to \infty$.

D'où $\sigma_t^2 = S_t$. □

1.5.2 Représentation ARCH(∞) d'un GARCH

Il est parfois utile de considérer la représentation ARCH(∞) pour les processus GARCH(p, q), cette représentation permet d'écrire la variance conditionnelle σ_t^2 de ε_t comme une fonction affine de son passé, sous les conditions $\omega > 0$, $\alpha_i \geq 0$, $\beta_j \geq 0$, $i = \overline{1,q}$, $j = \overline{1,p}$.

Représentation d'un GARCH(1.1) :

Théorème 1.5.2. *Si $\beta < 1$, alors le modèle GARCH(1.1) admet une représentation ARCH(∞) tels que :*

$$\sigma_t^2 = \frac{1}{1-\beta} + \alpha \sum_{i=1}^{\infty} \beta^{i-1} \varepsilon_{t-i}^2. \tag{1.37}$$

dans ce cas, on a :

$$A_s = \alpha^s \sum_{i=1}^{\infty} \beta^{(i-1)s} = \frac{\alpha^s}{1-\beta^s}.$$

La condition $A_s\mu_{2s} < 1$ (la condition de la stationnarité) peut s'écrire sous la forme suivante :

$$\alpha^s \mu_{2s} + \beta^s < 1, \quad pour \quad s \in]0,1].$$

Preuve. 1-

$$A_s = \alpha^s \sum_{i=1}^{\infty} \beta^{(i-1)s} = \frac{\alpha^s}{1-\beta^s} < 1 \Leftrightarrow A_s\mu_{2s} = \frac{\alpha^s \mu_{2s}}{1-\beta^s} < 1 \Leftrightarrow \alpha^s \mu_{2s} < 1 - \beta^s \Leftrightarrow \alpha^s \mu_{2s} + \beta^s < 1.$$

2-Si $\alpha + \beta < 1$ cette condition est satisfait pour $s = 1$ toutefois. La stationnarité au seconde ordre n'est pas necessaire pour la validité de (1.37).
3-Effectivement, si $\{\varepsilon_t;\ t \in \mathbb{Z}\}$ représente la solution stationnaire et non-anticipative du modèle GARCH(1.1), alors pour $q \geq 1$,

$$\begin{aligned}
\sigma_t^2 &= \omega + \alpha\varepsilon_{t-1}^2 + \beta(\omega + \alpha\varepsilon_{t-2}^2 + \beta\sigma_{t-2}^2) \\
&= \omega + \alpha\varepsilon_{t-1}^2 + \omega\beta + \alpha\beta\varepsilon_{t-2}^2 + \beta^2\sigma_{t-2}^2 \\
&= \omega + \alpha\varepsilon_{t-1}^2 + \omega\beta + \alpha\beta\varepsilon_{t-2}^2 + \beta^2(\omega + \alpha\varepsilon_{t-3}^2 + \beta\sigma_{t-3}^2) \\
&= \omega + \alpha\varepsilon_{t-1}^2 + \omega\beta + \alpha\beta\varepsilon_{t-2}^2 + \omega\beta^2 + \alpha\beta^2\varepsilon_{t-3}^2 + \beta^3\sigma_{t-3}^2 \\
&\vdots \\
&= \omega\textstyle\sum_{i=1}^{q}\beta^{i-1} + \alpha\sum_{i=1}^{q}\beta^{i-1}\varepsilon_{t-i}^2 + \beta^q\sigma_{t-q}^2.
\end{aligned}$$

D'après le corrollaire (1.4.1), il existe $s \in]0,1]$ tel que $\mathbb{E}(\sigma_t^{2s}) = c < \infty$, alors :

$$\sum_{q\geq 1} \mathbb{E}\big(\beta^q\sigma_{t-q}^2\big)^s = \sum_{q\geq 1} \beta^{qs}\mathbb{E}\big(\sigma_{t-q}^{2s}\big) = \frac{c\beta^s}{1-\beta^s} < \infty.$$

$\beta^q\sigma_{t-q}^2 \to 0 p.s.$ quand $q \to \infty$.
D'où on obtient :

$$\sigma_t^2 = \frac{\omega}{1-\beta} + \alpha \sum_{i=1}^{\infty} \beta^{i-1} \varepsilon_{t-i}^2.$$

□

Représentation ARCH(∞) d'un GARCH(p,q) :

Théorème 1.5.3. *Si $\{\varepsilon_t;\ t \in \mathbb{Z}\}$ est une solution strictement stationnaire et non anticipative du modèle (1.5), alors (ε_t) admet une représentation ARCH(∞) sous la forme (1.34) où les coefficients ϕ_i sont donnés par*

$$\phi_0 = \frac{\omega}{\mathcal{B}(1)}, \quad \sum_{i=1}^{\infty} \phi_i Z^i = \frac{\mathcal{A}(Z)}{\mathcal{B}(Z)}, \quad Z \in \mathbb{C}, \quad |Z| \leq 1, \tag{1.38}$$

où $\mathcal{A}(Z) = \alpha_1 z + \cdots + \alpha_q z^q$ et $\mathcal{B}(z) = 1 - \beta_1 z - \cdots - \beta_p z^p$.

Preuve. On réécrire le modèle (1.5) sous la forme vectorielle

$$\underline{\sigma}_t^2 = B\underline{\sigma}_{t-1}^2 + \underline{c}_t, \tag{1.39}$$

où $\underline{\sigma}_t^2 = (\sigma_t^2, \ldots, \sigma_{t-p+1})'$, $\underline{c}_t = (\omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2, 0, \ldots, 0)'$ et B une matrice définie dans le corollaire (1.3.1)

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

Le corollaire (1.3.1), prouve que la condition de stricte stationnarité implique que $\rho(B) < 1$. De plus, $\mathbb{E}(\|\underline{c}_t\|^s) < \infty$ d'après le corollaire (1.4.1). En itérant (1.39) on obtient :

$$\begin{aligned} \underline{\sigma}_t^2 &= B\underline{\sigma}_{t-1}^2 + \underline{c}_t \\ &= \underline{c}_t + B(\underline{c}_{t-1} + B\underline{\sigma}_{t-2}^2) \\ &= \underline{c}_t + B\underline{c}_{t-1} + B^2\sigma_{t-2}^2 \\ &\vdots \\ &= \underline{c}_t + B\underline{c}_{t-1} + B^2\underline{c}_{t-2} + \cdots + B^{t-1}\underline{c}_1 + B^t\underline{\sigma}_0^2 \\ &= \textstyle\sum_{k=0}^{\infty} B^k \underline{c}_{t-k}. \end{aligned}$$

Par consequent, les composantes du vecteur $\sum_{k=0}^{\infty} B^k \underline{c}_{t-k}$ sont presque sûrement des valeurs réelles.
On a ainsi

$$\sigma_t^2 = e' \sum_{k=0}^{\infty} B^k \underline{c}_{t-k}, \qquad e' = (1, 0, \ldots, 0).$$

Les coefficients obtenus en cette représentation ARCH(∞) coïncide avec ceux de (1.38). $\square$

Chapitre 2

Estimation des modèles GARCH

Dans ce chapitre nous présentons trois méthodes d'estimations des modèles GARCH : La méthode de moindres carrés ordinaire (M.C.O), maximum de vraisemblance (M.V) et la méthode des deux phases. Ainsi, nous établissons les propriétés asymptotiques (convergence presque sûr et la normalité asymptotique) de ces estimateurs en imposant des conditions et des hypothèses sur les moments.

2.1 Méthode des moindres carrées ordinaire(M.C.O)

Dans cette partie nous présentons la méthode moindre carée ordinaire pour un modéle ARCH.
Nous considérons l'estimation par les moindres carées ordinaires (M.C.O) du modèle ARCH(q) :

$$\begin{cases} \varepsilon_t = \sigma_t \eta_t \\ \sigma_t^2 = \omega_0 + \sum_{i=1}^q \alpha_{0i} \varepsilon_{t-i}^2. \end{cases} \tag{2.1}$$

Avec $\omega_0 > 0$, $\alpha_{0i} \geq 0$, $1 \leq i \leq q$,
et où $\{\eta_t;\ t \in \mathbb{Z}\}$ est une suite de variables *iid*, $\mathbb{E}(\eta_t) = 0$, $\mathbb{V}(\eta_t) = 1$.
La vraie valeur du vecteur des paramètres est noté $\theta_0 = (\omega, \alpha_{01}, \ldots, \alpha_{0q})'$ et nous noterons θ une valeur quelconque.
On déduit de (2.1) la représentation AR(q) données par :

$$\varepsilon_t^2 = \omega_0 + \sum_{i=1}^{q} \alpha_{01} \varepsilon_{t-i}^2 + u_t \tag{2.2}$$

où $u_t = \varepsilon_t^2 - \sigma_t^2 = (\eta_t^2 - 1)\sigma_t^2$.
La suite $(u_t, \mathcal{F}_{t-1})_t$ constitue donc une différence de martingale.
On suppose que l'on dispose d'observations $\varepsilon_1, \ldots, \varepsilon_n$, réalisations partielles du processus $\{\varepsilon_t;\ t \in \mathbb{Z}\}$, et de valeurs initiales $\varepsilon_0, \ldots, \varepsilon_{1-q}$ Par exemple ces valeurs initiales peuvent être choisis nulles. Soit le vecteur

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

on déduit de (2.2) le système

$$\varepsilon_t^2 = Z'_{t-1}\theta_0 + u_t, \qquad t = 1, \ldots, n, \tag{2.3}$$

Soit

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

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

$$X = \begin{pmatrix} Z'_{n-1} \\ \vdots \\ Z'_0 \end{pmatrix}, Y = \begin{pmatrix} \varepsilon_n^2 \\ \vdots \\ \varepsilon_1^2 \end{pmatrix}, U = \begin{pmatrix} u_n \\ \vdots \\ u_1 \end{pmatrix}.$$

Supposons que la matrice XX' soit inversible (nous verrons que c'est le cas asymptotiquement, donc aussi pour n assez grand). On en déduit l'estimateur des M.C.O de θ :

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

2.1.1 Propriétés asymptotiques de l'estimateur M.C.O

Nous serons amenés, pour établir la convergence, à considérer les hypothèses suivantes :
H1 :ε_t est solution non anticipative strictement stationnaire du modèle (2.1).
H2 :$\mathbb{E}_{\theta_0}(\varepsilon_t^4) < \infty$.
H3 :$P[\eta_t^2 = 1] \neq 1$.

Théorème 2.1.1. ***Convergence des estimateurs M.C.O pour un ARCH*** *Soit* $(\widehat{\theta}_n)$ *une suite d'estimateur du M.C.O satisfaisant (2.4). Sous H1, H2 et H3,*

$$\widehat{\theta}_n \xrightarrow{p.s} \theta_0.$$

Preuve. La preuve comporte plusieurs étapes :

étape 1 : Nous avons vu (Théorème2.1.1) que l'unique solution stationnaire non anticipative (ε_t) est ergodique. Le processus (Z_t) est également car Z_t s'écrit comme fonction mesurable des ε_{t-i}. le théorème ergodique appliqué au processus strictement stationnaire (Z_t) entraîne

$$\frac{1}{n}X'X = \frac{1}{n}\sum_{t=1}^{n} Z_{t-1}Z'_{t-1} \xrightarrow{p.s} \mathbb{E}_{\theta_0}(Z_{t-1}Z'_{t-1}) \quad quand \quad n \to \infty \tag{2.5}$$

L'existence de l'espérance est assuré par l'hypothèse **H3**. On a de même

$$\frac{1}{n}X'Y = \frac{1}{n}\sum_{t=1}^{n} Z_{t-1}\varepsilon_t^2 \xrightarrow{p.s} \mathbb{E}_{\theta_0}(Z_{t-1}\varepsilon_t^2) \quad quand \quad n \to \infty.$$

étape 2 : Montrons par absurde l'inversibilité de la matrice $\mathbb{E}_{\theta_0}Z_{t-1}Z'_{t-1} = \mathbb{E}_{\theta_0}Z_tZ'_t$.
Supposons qu'il existe c vecteur nul de $\mathbb{R}^{q+1}$ tel que $c'\mathbb{E}_{\theta_0}Z_tZ'_t = 0$.
Donc $\mathbb{E}_{\theta_0}\{c'Z_t(c'Z_t)'\} = 0$, d'où l'on déduit que $c'Z_t$ est p.s.constant. Par suite, il existe une combinaison linéaire p.s égale à une constante des variabes $\varepsilon_t^2, \ldots, \varepsilon_{t-q+1}^2$.
On peut supposer sans perte de généralité que, dans cette combinaison, le coefficients de $\varepsilon_t^2 = \eta_t^2\sigma_t^2$ est 1.

Donc η_t s'exprime p.s comme fonction mesurable des variables $\varepsilon_{t-1}, \ldots, \varepsilon_{t-q}$.
Ceci implique que η_t^2 est p.s égale à une constante. Cette constante ne peut être que 1, mais on aboutit alors à une contradiction avec **H3**.
Donc $\mathbb{E}_{\theta_0} Z_{t-1} Z_{t-1}^{'}$ est inversible.

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

$$\widehat{\theta}_n = \left(\frac{X'X}{n} \right)^{-1} \frac{X'Y}{n} \longrightarrow \left\{ \mathbb{E}_{\theta_0} Z_{t-1} Z_{t-1}^{'} \right\}^{-1} \mathbb{E}_{\theta_0} (Z_{t-1} \varepsilon_t^2).$$

étape 4 : Rappelons que le processus u_t est l'innovation forte de (ε_t^2). On a donc, en particulier, les relations d'orthogonalité

$$\mathbb{E}_{\theta_0}(u_t) = \mathbb{E}_{\theta_0}(u_t \varepsilon_{t-1}^2) = \cdots = \mathbb{E}(u_t \varepsilon_{t-q}^2) = 0.$$

c'est-à-dire

$$\mathbb{E}_{\theta_0}(Z_{t-1} u_t) = 0,$$

d'où l'on déduit, en utilisant (2.3),

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

D'après étape2 et étape3, $\widehat{\theta}_n$ converge p.s. vers θ_0.

Pour démontrer la normalité asymptotique de l'estimateur des M.C.O, nous devons faire l'hypothèse supplémentaire :
H4 :$\mathbb{E}_{\theta_0}(\varepsilon_t^8) < +\infty$.
Introduisons les matrices carrées symétriques de taille $q+1$,

$$\mathbf{A} = \mathbb{E}_{\theta_0}(Z_{t-1} Z_{t-1}^{'}), \quad \mathbf{I} = \mathbb{E}_{\theta_0}(\sigma_t^4 Z_{t-1} Z_{t-1}^{'}).$$

L'inversibilité de **A** a été établie dans la preuve du Théorème (2.1.1), celle de **I** sera montrée de la même façon, qui établit la normalité asymptotique de l'estimateur des M.C.O. On note $\mu_4 = \mathbb{E}(\eta_t^4)$. □

Théorème 2.1.2. *Sous les hypothèses **H1-H4**,*

$$\sqrt{n}(\widehat{\theta}_n - \theta_0) \xrightarrow{\mathcal{L}} \mathcal{N}(0, (\mu_4 - 1) \mathbf{A}^{-1} \mathbf{I} \mathbf{A}^{-1}).$$

Preuve. On a, d'après (2.3)

$$\widehat{\theta}_n = \left(\frac{1}{n} \sum_{t=1}^{n} Z_{t-1} Z_{t-1}^{'} \right)^{-1} \left(\frac{1}{n} \sum_{t=1}^{n} Z_{t-1} \varepsilon_t^2 \right)$$

$$= \left(\frac{1}{n} \sum_{t=1}^{n} Z_{t-1} Z_{t-1}^{'} \right)^{-1} \left\{ \frac{1}{n} \sum_{t=1}^{n} Z_{t-1} (Z_{t-1}^{'} \theta_0 + u_t) \right\}$$

$$= \theta_0 + \left(\frac{1}{n}\sum_{t=1}^{n} Z_{t-1}Z_{t-1}^{'}\right)^{-1}\left\{\frac{1}{n}\sum_{1} n \sum_{t=1}^{n} Z_{t-1}^{n}u_t\right\}.$$

Donc,

$$\sqrt{n}(\widehat{\theta}_n - \theta_0) = \left(\frac{1}{n}\sum_{t=1}^{n} Z_{t-1}Z_{t-1}^{'}\right)^{-1}\left\{\frac{1}{\sqrt{n}}\sum_{t=1}^{n} Z_{t-1}u_t\right\}. \tag{2.6}$$

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 $(\mathbb{E}_{\theta_0}(\lambda' Z_{t-1}u_t|\mathcal{F}_{t-1}) = 0)$, de variance

$$\begin{aligned}\mathbf{V}_{\theta_0}(\lambda' Z_{t-1}u_t) &= \lambda'\mathbb{E}_{\theta_0}\left\{Z_{t-1}Z_{t-1}^{'}u_t^2\right\}\lambda = \lambda'\mathbb{E}_{\theta_0}\left\{Z_{t-1}Z_{t-1}^{'}(\eta_t^2-1)^2\sigma_t^4\right\}\lambda \\ &= (\mu_4 - 1)\lambda'\mathbf{I}\lambda.\end{aligned}$$

Par application d'un TCL pour différence de martingale stationnaire, on en déduit que, pour tout $\lambda \neq 0$

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

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

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

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(2.1.1).
Par suite, on déduit de (2.5),(2.6) et (2.7), par raisonnement classique, que $\sqrt{n}(\widehat{\theta}_n - \theta_0)$ est asymptotiquement normal, de moyenne le vecteur nul, et de variance de la matrice du théorème. □

2.2 Méthode du maximum de vraisemblance

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

$$\begin{cases} \varepsilon_t = \sqrt{h_t}\sigma_t \\ h_t = \omega_0 + \sum_{i=1}^{q} \alpha_{0i}\varepsilon_{t-1}^2 + \sum_{j=1}^{p} \beta_{0j}h_{t-j}^2, \end{cases} \tag{2.8}$$

où $\omega > 0,\ i = 1, \cdots, q,\ \beta_j \geq 0,\ j = 1, \cdots, p,\ (\eta_t)$ est une suite de variables *iid*.
Les ordres p et q sont supposés connus. Afin de définir l'estimateur du M.V du modèle (2.8) introduisons

$$\theta = (\theta_1, \theta_2, \cdots, \theta_{p+q+1})' = (\omega, \alpha_1, \cdots, \alpha_q, \beta_1, \cdots, \beta_p)', \tag{2.9}$$

le vecteur des paramètres à estimer, qui appartient à un espace de paramètres $\Theta \subset]0, +\infty[\times[0, +\infty[^{p+q}$. la vraie valeur du paramètre est inconnue. Elle est notée :

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

Pour écrire la vraisemblance du modèle, il faut specifier une distribution particulière pour les variables *iid* η_t. On considère généralement la vraisemblance gaussienne, i.e. la vraisemblance obtenue à partir d'une loi normale centrée réduite pour les η_t. Nous ne ferons cependant pas l'hypothèse que cette loi constitue la vraie distribution du processus *iid*. La spécification d'une distribution gaussienne pour les variables η_t ne permet pas d'en déduire simplement la loi de l'échantillon. On travaille avec la vraisemblance de $(\varepsilon_1, \cdots, \varepsilon_n)$ conditionnellement à certaines valeurs initiales.
Étant données des valeurs initiales $\varepsilon_0, \cdots, \varepsilon_{1-q}, \quad \widetilde{\sigma}_0^2, \cdots, \widetilde{\sigma}_{1-p}^2$ que nous allons préciser la vraisemblance conditionnelle gaussienne $L_n(\theta)$ s'écrit :

$$L_n(\theta) = L_n(\theta; \varepsilon_1, \cdots, \varepsilon_n) = \prod_{t=1}^{n} \sqrt{\frac{1}{2\Pi\widetilde{\sigma}_t^2}} \exp\Big(-\frac{\varepsilon_t^2}{2\widetilde{\sigma}_t^2}\Big), \tag{2.10}$$

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

$$\widetilde{\sigma}_t^2 = \widetilde{\sigma}_t^2(\theta) = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \widetilde{\sigma}_{t-j}^2. \tag{2.11}$$

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 :

$$\varepsilon_0^2 = \cdots = \varepsilon_{1-q}^2 = \frac{1}{1 - \sum_{i=1}^{q} \alpha_i - \sum_{j=1}^{p} \beta_j}. \tag{2.12}$$

On peut prendre comme valeurs initiales

$$\varepsilon_0^2 = \cdots = \varepsilon_{1-q}^2 = \widetilde{\sigma}_0^2 = \widetilde{\sigma}_{1-p}^2 = \omega, \tag{2.13}$$

ou encore

$$\varepsilon_0^2 = \cdots = \varepsilon_{1-q}^2 = \widetilde{\sigma}_0^2 = \widetilde{\sigma}_{1-p}^2 = \varepsilon_1^2. \tag{2.14}$$

Un estimateur du M.V de θ est défini comme toute solution mesurable $\widehat{\theta}_n$ de :

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

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

$$\widetilde{I}_n(\theta) = \frac{1}{n} \sum_{t=1}^{n} \widetilde{\ell}_t, \qquad \widetilde{\ell}_t = \widetilde{\ell}_t(\theta) = \frac{\varepsilon_t^2}{\widetilde{\sigma}_t^2} + \log \widetilde{\sigma}_t^2. \tag{2.15}$$

et $\widetilde{\sigma}_t^2$ est définie en (2.11). Un estimateur du maximum de vraisemblance est donc une solution mesurable de l'équation :

$$\widehat{\theta}_n = \arg\min_{\theta \in \Theta} \widetilde{I}_n(\theta). \tag{2.16}$$

EQUATIONS DE VRAISEMBLANCE :

Définition 2.2.1. *On obtient les equations de vraisemblance en annulant la dérivée par rapport à θ du critère $\widetilde{I}_n(\theta)$, ce qui donne :*

$$\frac{\partial \log L(\theta)}{\partial\theta} = \frac{\partial \widetilde{I}_n(\theta)}{\partial\theta} = 0, \tag{2.17}$$

avec

$$\frac{\partial \widetilde{I}_n(\theta)}{\partial\theta} = \frac{1}{n}\sum_{t=1}^{n} \frac{\varepsilon_t^2}{\widetilde{\sigma}_t^4}\frac{\partial\widetilde{\sigma}_t^2}{\partial\theta} - \frac{\widetilde{\sigma}_t^2}{\widetilde{\sigma}_t^4}\frac{\partial\widetilde{\sigma}_t^2}{\partial\theta} = \frac{1}{n}\sum_{t=1}^{n}\left\{\varepsilon_t^2 - \widetilde{\sigma}_t^2\right\}\frac{1}{\widetilde{\sigma}_t^4}\frac{\partial\widetilde{\sigma}_t^2}{\partial\theta} = 0. \tag{2.18}$$

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

$$\frac{1}{n}\sum_{t=1}^{n}\left\{\varepsilon_t^2 - \sigma_t^2\right\}\frac{1}{\sigma_t^4}\frac{\partial\sigma_t^2}{\partial\theta} = 0. \tag{2.19}$$

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

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

car $\frac{1}{\sigma_t^4(\theta_0)}\frac{\partial\sigma_t^2(\theta_0)}{\partial\theta}$ est une fonction mesurable des ε_{t-i}, $i > 0$. Ce résultat n'est autre que la version asymptotique de (2.18) en θ_0, en utilisant le théorème ergoqique.

2.2.1 Propriétés asymptotiques de l'estimateur de M.V

La convergence forte

Le modèle (2.8) possède une solution strictement stationnaire si et seulement si le coefficient de Lyapounov de la suite de matrices $A_0 = (A_{0t})$ est strictement négatif $\gamma(A_0) < 0$ (ce qu'on a déjà vue dans le premier chapitre) tel que :

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

avec

$$\gamma(A_0) = \inf_{t\in\mathbb{N}^*} \frac{1}{t}\mathbb{E}\left\{\log\|A_t \dots A_1\|\right\} = \lim_{t\to+\infty} p.s. \frac{1}{t}\log\|A_t A_{t-1}\dots A_1\|. \tag{2.20}$$

Notons

$$\mathcal{A}_0(z) = \sum_{i=1}^{q} \alpha_i z^i, \quad et \quad \mathcal{B}_0(z) = 1 - \sum_{j=1}^{q} \beta_j z^j,$$

a toutes ces racines à l'extérieur du cercle unité, et $\alpha_i \geq 0, \ 1 \leq i \leqq$. Pour montrer la convergence forte, les hypothèses suivantes serons faites :
H1 :$\theta_0 \in \Theta$, et Θ est compacte.
H2 :$\gamma(A_0) < 0$ et $\forall \theta \in \Theta, \ \sum_{j=1}^{q} \beta_j < 1$.
H3 :η_t^2 a une loi non dégénérée et $\mathbb{E}(\eta_t^2) = 1$.
H4) :Si $p > 0$, $\mathcal{A}_0(z)$ et $\mathcal{B}_0(z)$ n'ont pas de racine commune, $\mathbf{A}_0(1) \neq 0$, et $\alpha_{0q} + \beta_{0p} \neq 0$.

Théorème 2.2.1. ***[convergence forte de l'estimateur du M.V]*** *Soit $(\widehat{\theta}_n)$ une suite d'estimateurs du M.V satisfaisant (2.16), avec les conditions initiales (2.13) ou (2.14). Sous les hypothèses **H1-H4***

$$\widehat{\theta}_n \xrightarrow{p.s} \theta_0, \quad quand \quad n \to \infty.$$

Preuve. On note K et ρ des constantes génériques dont la valeur pourra changer en cours de preuve. A titre d'exemple, nous pourrons écrire pour $0 < \rho_1 < 1$ et $0 < \rho_2 < 1, \ i_1 \geq 0, \ i_2 \geq 0,$

$$0 < K \sum_{i \geq i_1} \rho_1^i + \sum_{i \geq i_2} i \rho_2^i \leq K \rho^{min(i_1, i_2)}.$$

□

Schéma de la preuve

La démonstration repose sur une représentation vectorielle autorégressive d'ordre 1 du vecteur $\underline{\sigma}_t^2 = (\sigma_t^2, \sigma_{t-1}^2, \cdots, \sigma_{t-p+1}^2)$, analogue à celle utilisée pour l'étude de la stationnarité. L'hypothèse **H2** permet d'exprimer $\underline{\sigma}_t^2$ sous forme d'une série dépendant du passé infini de la variable ε_t^2. On montre que les valeurs initiales n'ont pas d'importance asymptotiquement en utilisant le fait que, sous l'hypothèse de stationnarité stricte, ε_t^2 admet nécessairement un moment d'ordre s, avec $s > 0$. Cette propriété permet également de vérifier que $\mathbb{E}_{\theta_0}[\ell_t(\theta_0)]$ est bien définie sur $\mathbb{R}$ et que $\mathbb{E}_{\theta_0}(\ell_t(\theta)) - \mathbb{E}_{\theta_0}(\ell_t(\theta_0)) \geq 0$, ce qui assure que le critère limite est minimisé en la vraie valeur. La difficulté provient du faite que $\mathbb{E}_{\theta_0}(\ell_t^+(\theta))$ peut être égal à $+\infty$. Les hypothéses **H3** et **H4** sont cruciales pour établir l'identifiabilité : la première exclut l'existence d'une combinaison linéaire constante entre les $\varepsilon_{t-j}^2, \ j \geq 0$. On utilise également l'hypothèse d'absence de racines communes.
L'ergodicité de $\ell_t(\theta)$ est un argument de compacité permettent de conclure.
Preuve de théorème(2.2.1) : réécrire l'équation :

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2, \quad \forall t \in \mathbb{Z}, \tag{2.21}$$

sous forme matricielle on a :

$$\underline{\sigma}_t^2 = \underline{c}_t + B \underline{\sigma}_{t-1}^2 \tag{2.22}$$

où

$$\underline{\sigma}_t^2 = \begin{pmatrix} \sigma_t^2 \\ \sigma_{t-1}^2 \\ \vdots \\ \sigma_{t-p+1}^2 \end{pmatrix}, \quad \underline{c}_t^2 = \begin{pmatrix} \omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad B = \begin{pmatrix} \beta_1 & \beta_2 & \cdots & & \beta_p \\ 1 & 0 & \cdots & & 0 \\ 0 & 1 & \cdots & & 0 \\ \vdots & \ddots & \ddots & & \vdots \\ 0 & & \cdots & 1 & 0 \end{pmatrix}. \tag{2.23}$$

Nous allons établir les résultats intermédiaires suivants :

i) $\lim_{n\to\infty} \sup_{\theta\in\Theta} |I_n(\theta) - \widetilde{I}_n(\theta)| = 0, \quad p.s.$

ii) $(\exists t \in \mathbb{Z}$ tel que $\sigma_t^2(\theta) = \sigma_t^2(\theta_0) P_{\theta_0} p.s.) \Longrightarrow \theta = \theta_0$.

iii) $\mathbb{E}_{\theta_0}|\ell_t(\theta_0)| < \infty$ si $\theta = \theta_0$, et si $\theta \neq \theta_0$, $\mathbb{E}_{\theta_0}(\ell_t(\theta)) > \mathbb{E}_{\theta_0}(\ell_t(\theta_0))$.

iv) Pour tout $\theta \neq \theta_0$ il existe un voisinage $V(\theta)$ tel que $\liminf_{n\to\infty} \inf_{\theta^*\in V(\theta)} \widetilde{I}_n(\theta^*) > \mathbb{E}_{\theta_0}(\ell_1(\theta_0))$, p.s.

(i)

Pour montrer la condition (i) on utilise le corrollaire (1.2.1) et la condition $\sum_{j=1}^p \beta_j < 1$ de l'hypothèse **H2** implique $\rho(B) < 1$.

La compacité de Θ implique que :

$$\sup_{\theta\in\Theta} \rho(B) < 1. \tag{2.24}$$

En itérant (2.22), on obtient :

$$\underline{\sigma}_t^2 = \underline{c}_t + B\underline{c}_{t-1} + B^2\underline{c}_{t-2} + \ldots + B^{t-1}\underline{c}_1 + B^t\underline{\sigma}_0^2 = \sum_{k=0}^{\infty} B^k \underline{c}_{t-k}. \tag{2.25}$$

Soit $\widetilde{\underline{\sigma}}_t^2$ le vecteur obtenu en ramplaçant σ_{t-i}^2 par $\widetilde{\underline{\sigma}}_{t-i}^2$ dans $\underline{\sigma}_{t-i}^2$. Soit $\widetilde{\underline{c}}$ le vecteur obtenu en remplaçant $\varepsilon_0^2, \ldots, \varepsilon_{1-q}^2$ par les valeurs initiales (2.13) ou (2.14). nous avons :

$$\widetilde{\underline{\sigma}}_t^2 = \underline{c}_t + B\underline{c}_{t-1} + \ldots + B^{t-q-1}\underline{c}_{q+1} + B^{t-q}\widetilde{\underline{c}}_q + \ldots + B^{t-1}\widetilde{\underline{c}}_1 + B^t\widetilde{\underline{\sigma}}_0^2. \tag{2.26}$$

On déduit de (2.24)-(2.26) que presque sûrement :

$$\sup_{\theta\in\Theta} \|\underline{\sigma}_t^2 - \widetilde{\underline{\sigma}}_t^2\| = \sup_{\theta\in\Theta} \left\| \left\{ \sum_{k=1}^{q} B^{t-k}(\underline{c}_k - \widetilde{\underline{c}}_k) + B^t(\underline{\sigma}_0^2 - \widetilde{\underline{\sigma}}_0^2) \right\} \right\| \leq K\rho^t, \quad \forall t \in \mathbb{Z}. \tag{2.27}$$

Pour $x > 0$ on a $\log x \leq x - 1$. Par suite, pour $x, y > 0$, $\left|\log \frac{x}{y}\right| \leq \frac{|x-y|}{\min(x,y)}$. On a donc presque sûrement, en utilisant (2.27),

$$\begin{aligned} \sup_{\theta\in\Theta} |I_n(\theta) - \widetilde{I}_n(\theta)| &\leq n^{-1} \sum_{t=1}^n \sup_{\theta\in\Theta} \left\{ \left| \frac{\widetilde{\sigma}_t^2 - \sigma_t^2}{\widetilde{\sigma}_t^2 \sigma_t^2} \right| \varepsilon_t^2 + \left| \log\left(1 + \frac{\sigma_t^2 - \widetilde{\sigma}_t^2}{\widetilde{\sigma}_t^2}\right) \right| \right\} \\ &\leq \left\{ \sup_{\theta\in\Theta} \frac{1}{\omega^2} \right\} K n^{-1} \sum_{t=1}^n \rho^t \varepsilon_t^2 + \left\{ \sup_{\theta\in\Theta} \frac{1}{\omega} \right\} K n^{-1} \sum_{t=1}^n \rho^t. \end{aligned} \tag{2.28}$$

Pour déduire (i) est suffisant de utiliser le lemme de Cesàro tel que $\rho^t\varepsilon_t^2 \to 0$p.s. Cette convergence est obtenu par le lemme de Borel-Cantelli, l'inégalité de Markov et l'existence d'un moment d'ordre $s > 0$ pour ε_t^2 :
soit $\delta > 0$,

$$\sum_{t=0}^{\infty}\mathbb{P}(\rho^t\varepsilon_t^2 > \delta) \leq \sum_{t=0}^{\infty}\frac{\mathbb{E}(\rho^t\varepsilon_t^2)^s}{\delta^s} = \frac{\mathbb{E}(\varepsilon_t^{2s})}{(1-\rho)\delta^s} < \infty.$$

(ii) Identifiabilité du paramètre.
Supposons que $\sigma_t^2(\theta) = \sigma_t^2(\theta_0)P_{\theta_0}$ p.s. On note que $\mathcal{B}_\theta(B)$ est inversible sous l'hypothèse **H2**, par le corollaire(1.2.1). Si $\mathcal{A}_\theta(1) \neq 0$ d'après (2.21), on obtient :

$$\frac{\mathcal{A}_\theta(B)}{\mathcal{B}_\theta(B)}\left\{\varepsilon_t^2 + \frac{\omega}{\mathcal{A}_\theta(1)}\right\} = \frac{\mathcal{A}_{\theta_0}(B)}{\mathcal{B}_{\theta_0}(B)}\left\{\varepsilon_t^2 + \frac{\omega_0}{\mathcal{B}_{\theta_0}(1)}\right\} \quad p.s. \quad \forall t \in \mathbb{Z}.$$

$\Longrightarrow$

$$\left\{\frac{\mathcal{A}_\theta(B)}{\mathcal{B}_\theta(B)} - \frac{\mathcal{A}_{\theta_0}(B)}{\mathcal{B}_{\theta_0}(B)}\right\}\varepsilon_t^2 = \frac{\omega_0}{\mathcal{B}_{\theta_0}(1)} - \frac{\omega}{\mathcal{B}_\theta(1)} \quad p.s. \quad \forall t \in \mathbb{Z}.$$

Si la série en B entre accolades était non nulle, cela signifierait qu'il existerait une combinaison linéaire des ε_{t-j}^2, $j \geq 0$, égale à une constante. Donc l'innovation linéaire du processus (ε_t^2) serait nulle. Or, la loi de η_t étant non dégénérée d'après l'hypothèse **H3**,
$\varepsilon_t^2\mathbb{E}_{\theta_0}(\varepsilon_t^2|\varepsilon_{t-1}^2, \cdots) = \sigma_t^2(\theta_0)(\eta_t^2 - 1) \neq 0$, avec probabilité positive. On a donc :

$$\frac{\mathcal{A}_\theta(z)}{\mathcal{B}_\theta(z)} = \frac{\mathcal{A}_{\theta_0}(z)}{\mathcal{B}_{\theta_0}(z)}, \quad \forall |z| \leq 1 \quad et \quad \frac{\omega}{\mathcal{B}_\theta(1)} = \frac{\omega_0}{\mathcal{B}_{\theta_0}(1)}. \tag{2.29}$$

Sous l'hypothèse **H4**, ceci entraîne $\mathcal{A}_\theta(z) = \mathcal{A}_{\theta_0}(z)$, $\mathcal{B}_\theta(z) = \mathcal{B}_{\theta_0}(z)$ et $\omega = \omega_0$, ce qui preuve (ii).

(iii) Le critère limite est minimisé en la vraie valeur :
Le critère que l'on minimise n'est pas intégrable en tout point, mais remarquons que $\mathbb{E}_{\theta_0}(I_n(\theta)) = \mathbb{E}_{\theta_0}(\ell_t(\theta))$ est bien défini dans $\mathbb{R} \cup \{+\infty\}$ car, en notant $x^- = \min(x, 0)$ et $x^+ \max(x, 0)$. En utilisant le fait que $(f+g)^- \leq g^-$ pour $f \geq 0$, et que si $f \leq g$ alors $f^- \geq g^-$, on aura :

$$\mathbb{E}_{\theta_0}(\ell_t^-(\theta)) \leq \mathbb{E}_{\theta_0}(\log^- \sigma_t^2 \leq \max\{0, -\log\omega\} < \infty.$$

Il reste à montrer que $\mathbb{E}_{\theta_0}(\ell_t^+(\theta_0)) < \infty$, en utilisant l'inégalité de Jensen et à l'existence d'un moment d'ordre $s > 0$ pour ε_t^2, il vient :

$$\mathbb{E}_{\theta_0}(\log^+ \sigma_t^2(\theta_0)) < \infty, \tag{2.30}$$

car

$$\mathbb{E}_{\theta_0}(\log\sigma_t^2(\theta_0)) = \mathbb{E}_{\theta_0}\big(\frac{1}{s}\log\{\sigma_t^2(\theta_0)\}^s\big) \leq \frac{1}{s}\log\mathbb{E}_{\theta_0}(\{\sigma_t^2(\theta_0)\}^s) < \infty.$$

d'où

$$\mathbb{E}_{\theta_0}(\ell_t(\theta_0)) = \mathbb{E}_{\theta_0}\left\{\frac{\sigma_t^2(\theta_0)\eta_t^2}{\sigma_t^2(\theta_0)} + \log\sigma_t^2(\theta_0)\right\} = 1 + \mathbb{E}_{\theta_0}(\log\sigma_t^2(\theta_0)) < \infty.$$

Ayant déjà établi que $\mathbb{E}_{\theta_0}(\ell_t^-(\theta)) < \infty$, on en déduit que $\mathbb{E}_{\theta_0}(\ell_t(\theta_0))$ est bien défini dans $\mathbb{R}$.
Puisque pour tout $x > 0$, $\log x \leq x - 1$, avec égalité si et seulement si $x = 1$, on a :

$$\begin{aligned}
\mathbb{E}_{\theta_0}(\ell_t(\theta)) - \mathbb{E}_{\theta_0}(\ell_t(\theta_0)) &= \mathbb{E}_{\theta_0}\left(\log \tfrac{\sigma_t^2(\theta)}{\sigma_t^2(\theta_0)}\right) + \mathbb{E}_{\theta_0}\left(\tfrac{\sigma_t^2(\theta_0)\eta_t^2}{\sigma_t^2(\theta)}\right) - \mathbb{E}_{\theta_0}(\eta_t^2) \\
&= \mathbb{E}_{\theta_0}\left(\log \tfrac{\sigma_t^2(\theta)}{\sigma_t^2(\theta_0)}\right) + \mathbb{E}_{\theta_0}\left(\tfrac{\sigma_t^2(\theta_0)}{\sigma_t^2(\theta)}\right) - 1 \\
&\geq \mathbb{E}_{\theta_0}\left\{\log \frac{\sigma_t^2(\theta)}{\sigma_t^2(\theta_0)} + \log \frac{\sigma_t^2(\theta_0)}{\sigma_t^2(\theta)}\right\} = 0, \qquad (2.31)
\end{aligned}$$

avec l'égalité si et seulement si $\sigma_t^2(\theta_0)/\sigma_t^2(\theta) = 1$ P_{θ_0} -p.s. c'est-à-dire, étant donné (ii), si et seulement si $\theta = \theta_0$.

(iv)Utilisation de compacité de Θ et l'ergodicité de $(\ell_t(\theta))$.

Pour tout $\theta \in \Theta$ et tout entier positif k, soit $V_k(\theta)$ la boule ouverte de centre θ et de rayon $1/k$. En raison de (i),

$$\limsup_{n\to\infty} \sup_{\theta\in\Theta} |I_n(\theta) - \widetilde{I}_n(\theta)| \geq \liminf_{n\to\infty} \inf_{\theta^*\in V_k(\theta)\cap\Theta} I_n(\theta^*) - \liminf_{n\to\infty} \inf_{\theta^*\in V_k(\theta)\cap\Theta} \widetilde{I}_n(\theta^*)$$

$$\begin{aligned}
\liminf_{n\to\infty} \inf_{\theta^*\in V_k(\theta)\cap\Theta} \widetilde{I}_n(\theta^*) &\geq \liminf_{n\to\infty} \inf_{\theta^*\in V_k(\theta)\cap\Theta} I_n(\theta^*) - \limsup_{n\to\infty} \sup_{\theta\in\Theta} |I_n(\theta) - \widetilde{I}_n(\theta)| \\
&\geq \liminf_{n\to\infty} \inf_{\theta^*\in V_k(\theta)\cap\Theta} I_n(\theta^*) \\
&\geq \liminf_{n\to\infty} \frac{1}{n}\sum_{t=1}^{n} \inf_{\theta^*\in V_k(\theta)\cap\Theta} \ell_t(\theta^*).
\end{aligned}$$

On peut utiliser le théorème ergodique standard pour obtenir la convergence de cette moyenne empirique, car on a déjà vue $\ell_t(\theta^*)$ n'est pas nécessairement intégrable, sauf en θ_0. Une adaptation de ce théorème pour une suite strictement stationnaire et ergodique de variables admettant une espérance dans $\mathbb{R} \cup \{+\infty\}$ ce qui est le cas de $\{\ell_t(\theta^*)\}$ et par suite de $\left\{\inf_{\theta^*\in V_k(\theta)\cap\Theta} \ell_t(\theta^*)\right\}$, permet d'affirmer que :

$$\liminf_{n\to\infty} n^{-1}\sum_{t=1}^{n} \inf_{\theta^*\in V_k(\theta)\cap\Theta} \ell_t(\theta^*) = \mathbb{E}_{\theta_0}\left(\inf_{\theta^*\in V_k(\theta)\cap\Theta} \ell_1(\theta^*)\right).$$

D'après le théorème de beppo-levi, $\mathbb{E}_{\theta_0}\left(\inf_{\theta^*\in V_k(\theta)\cap\Theta} \ell_1(\theta^*)\right)$ tend en croissant vers $\mathbb{E}_{\theta_0}(\ell_1(\theta))$ quand $k \to \infty$. Étant donné (2.31), nous avons montré (iv). La fin de la preuve du théorème utilise en argument de compacité. Remarquons que pour tout voisinage $V(\theta_0)$ de θ_0

$$\limsup_{n\to\infty} \inf_{\theta^*\in V(\theta_0)} \widetilde{I}_n(\theta^*) \leq \lim_{n\to\infty} \widetilde{I}_n(\theta_0) = \lim_{n\to\infty} I_n(\theta_0) = \mathbb{E}_{\theta_0}(\ell_1(\theta_0)). \qquad (2.32)$$

Le compact Θ est recouvert par la réunion d'un voisinage quelconque $V(\theta_0)$ de θ_0 et de l'ensemble des $V(\theta)$, $\theta \in \Theta|V(\theta_0)$, où $V(\theta)$ vérifie (iv). Il existe donc un sous recouvrement fini de Θ par $V(\theta_0), V(\theta_1), \ldots, V(\theta_k)$, d'où l'on déduit que

$$\inf_{\theta\in\Theta} \widetilde{I}_n(\theta) = \min_{i=0,\ldots,k} \inf_{\theta\in\Theta\cap V(\theta_i)} \widetilde{I}_n(\theta).$$

les relations (iv) et (2.32) montrent que, presque sûrement, $\widehat{\theta}_n$ appartient à $V(\theta_0)$ pour n assez grand. Ceci étant vrai pour tout voisinage $V(\theta_0)$, le résultat est montré.

Normalité asymptotique

Pour montrer la normalité asymptotique, on considère les hypothèses supplémentaire suivantes :

H5 :$\theta_0 \in \overset{\circ}{\Theta}$, où $\overset{\circ}{\Theta}$ est l'intérieur de Θ.

H6 :$k_\eta = \mathbb{E}(\eta_t^4) < \infty$.

La loi limite de $\widehat{\theta}_n$ est donnée par le résultat suivant :

Théorème 2.2.2. ***[Normalité asymptotique des estimateurs du M.V]*** *Sous les hypothèses **H1-H6**,*

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

où

$$J := \mathbb{E}_{\theta_0}\left(\frac{\partial^2 \ell_t(\theta_0)}{\partial\theta\partial\theta'}\right) = \mathbb{E}_{\theta_0}\left(\frac{1}{\sigma_t^4(\theta_0)}\frac{\partial\sigma_t^2(\theta_0)}{\partial\theta}\frac{\partial\sigma_t^2(\theta_0)}{\partial\theta'}\right), \tag{2.33}$$

est une matrice définie positive.

Preuve. La preuve de ce théorème repose classiquement sur un développement de Taylor du critère (2.15) en θ_0.

On a :

$$\begin{aligned} 0 &= n^{-1/2}\sum_{t=1}^{n}\frac{\partial}{\partial\theta}\widetilde{\ell}_t(\widehat{\theta}_n) \\ &= n^{-1/2}\sum_{t=1}^{n}\frac{\partial}{\partial\theta}\widetilde{\ell}_t(\theta_0) + \left(\frac{1}{n}\sum_{t=1}^{n}\frac{\partial^2}{\partial\theta_i\partial\theta_j}\widetilde{\ell}_t(\theta_{ij}^*)\right)\sqrt{n}(\widehat{\theta}_n - \theta_0), \end{aligned} \tag{2.34}$$

où les θ_{ij}^* sont entre $\widehat{\theta}_n$ et θ_0. Nous montrerons que :

$$n^{-1/2}\sum_{t=1}^{n}\frac{\partial}{\partial\theta}\widetilde{\ell}_t(\theta_0) \xrightarrow{\mathcal{L}} \mathcal{N}(0, (k_\eta - 1)J), \tag{2.35}$$

et

$$\left(\frac{1}{n}\sum_{t=1}^{n}\frac{\partial^2}{\partial\theta_i\partial\theta_j}\widetilde{\ell}_t(\theta_{ij}^*)\right) \xrightarrow{p} J(i,j). \tag{2.36}$$

Maintenant, nous allons a nouveau décomposer la démonstration en plusieurs points.

- (i) $\mathbb{E}_{\theta_0}\left\|\frac{\partial\ell_t(\theta_0)}{\partial\theta}\frac{\partial\ell_t(\theta_0)}{\partial\theta'}\right\| < \infty, \qquad \mathbb{E}_{\theta_0}\left\|\frac{\partial_t^2(\theta_0)}{\partial\theta\partial\theta'}\right\| < \infty.$
- (ii)J est inversible et $V_{\theta_0}\left\{\frac{\partial\ell_t(\theta_0)}{\partial\theta}\right\} = (k_\eta - 1)J$.
- (iii) Il existe un voisinage $V(\theta_0)$ de θ_0 tel que, pour tous $i, j, k \in \{1, \ldots, p+q+1\}$,
$$\mathbb{E}_{\theta_0}\sup_{\theta\in V(\theta_0)}\left|\frac{\partial^3\ell_t(\theta)}{\partial\theta_i\partial\theta_j\partial\theta_k}\right| < \infty.$$
- (iv) $\left\|n^{-1/2}\sum_{t=1}^{n}\left\{\frac{\partial\ell_t(\theta_0)}{\partial\theta} - \frac{\partial\widetilde{\ell}_t(\theta_0)}{\partial\theta}\right\}\right\|$ et $\sup_{\theta\in V(\theta_0)}\left\|n^{-1}\sum_{t=1}^{n}\left\{\frac{\partial^2\ell_t(\theta_0)}{\partial\theta\partial\theta'} - \frac{\partial^2\widetilde{\ell}_t(\theta_0)}{\partial\theta\partial\theta'}\right\}\right\|$

  tendent en probabilité vers 0 quand $n \to \infty$.

– (v) $n^{-1/2}\sum_{t=1}^{n}\frac{\partial}{\partial\theta}\ell_t(\theta_0)\xrightarrow{\mathcal{L}}\mathcal{N}(0,(k_\eta-1)J)$.
– (vi) $n^{-1}\sum_{t=1}^{n}\frac{\partial^2}{\partial\theta_i\partial\theta_j}\ell_t(\theta^*_{ij})\longrightarrow J(i,j)p.s.$

-i) **Intégrabilité des dérivées du critère en θ_0 :**
On a $\ell_t(\theta)=\varepsilon_t^2/\sigma_t^2+\log\sigma_t^2$, nous avons :

$$\frac{\partial\ell_t(\theta)}{\partial\theta}=\left\{1-\frac{\varepsilon_t^2}{\sigma_t^2}\right\}\left\{\frac{1}{\sigma_t^2}\frac{\partial\sigma_t^2}{\partial\theta}\right\}. \tag{2.37}$$

$$\frac{\partial^2\ell_t(\theta)}{\partial\theta\partial\theta'}=\left\{1-\frac{\varepsilon_t^2}{\sigma_t^2}\right\}\left\{\frac{1}{\sigma_t^2}\frac{\partial^2\sigma_t^2}{\partial\theta\partial\theta'}\right\}+\left\{2\frac{\varepsilon_t^2}{\sigma_t^2}-1\right\}\left\{\frac{1}{\sigma_t^2}\frac{\partial\sigma_t^2}{\partial\theta}\right\}\left\{\frac{1}{\sigma_t^2}\frac{\partial\sigma_t^2}{\partial\theta'}\right\}. \tag{2.38}$$

Pour $\theta=\theta_0$, $\varepsilon_t^2/\sigma_t^2=\eta_t^2$ est indépendant des termes en σ_t^2 ou en ses dérivées. Pour prouver (i) il suffira donc de montrer que

$$\mathbb{E}_{\theta_0}\left\|\frac{1}{\sigma_t^2}\frac{\partial\sigma_t^2}{\partial\theta}(\theta_0)\right\|<\infty,\quad \mathbb{E}_{\theta_0}\left\|\frac{1}{\sigma_t^2}\frac{\partial^2\sigma_t^2}{\partial\theta\partial\theta'}(\theta_0)\right\|<\infty,\quad \mathbb{E}_{\theta_0}\left\|\frac{1}{\sigma_t^4}\frac{\partial\sigma_t^2}{\partial\theta}\frac{\partial\sigma_t^2}{\partial\theta'}(\theta_0)\right\|<\infty. \tag{2.39}$$

D'après (2.25) nous avons :

$$\frac{\partial\underline{\sigma}_t^2}{\partial\omega}=\sum_{k=0}^{\infty}B^k\underline{1},\qquad \frac{\partial\underline{\sigma}_t^2}{\partial\alpha_i}=\sum_{k=0}^{\infty}B^k\underline{\varepsilon}_{t-k-i}^2, \tag{2.40}$$

$$\frac{\partial\underline{\sigma}_t^2}{\partial\beta_(j)}=\sum_{k=1}^{\infty}\left\{\sum_{i=1}^{k}B^{i-1}B^jB^{k-i}\right\}\underline{c}_{t-k}. \tag{2.41}$$

où $\underline{1}=(1,0,\cdots,0)'$, $\underline{\varepsilon}_t^2=(\varepsilon_t^2,0,\cdots0)'$, $B^(j)$ est une matrice $p\times p$ avec $(1,j)$ est un 1 et des 0 ailleurs. Remarquons que, d'après la positivité des coefficients et (2.40)-(2.41), les dérivées de σ_t^2 sont pas négative. D'après (2.40), il est claire que $\frac{\partial\sigma_t^2}{\partial\omega}$ est bornée. Puisque $\sigma_t^2\geq\omega>0$, il en est de même pour $\{\partial\sigma_t^2/\partial\omega\}/\sigma_t^2$. Donc cette variable possède des moments de tous ordres. D'après la seconde égalité de (2.40) et la positivité de tous les termes considérés, nous avons :

$$\alpha_i\frac{\partial\underline{\sigma}_t^2}{\partial\alpha_i}=\sum_{k=1}^{\infty}B^k\alpha_i\underline{\varepsilon}_{t-k-1}^2\leq\sum_{k=1}^{\infty}B^k\underline{c}_{t-k}=\underline{\sigma}_t^2.$$

On en déduit que

$$\frac{1}{\sigma_t^2}\frac{\partial\sigma_t^2}{\partial\alpha_i}\leq\frac{1}{\alpha_i}. \tag{2.42}$$

La variable $(\partial\sigma_t^2/\partial\alpha_i)\sigma_t^{-2}$ possède donc des moments de tous ordres en $\theta=\theta_0$. D'après (2.41) et $\beta_jB^{(j)}\leq B$, nous avons :

$$\beta_j\frac{\partial\underline{\sigma_t}^2}{\partial\beta_j}\leq\sum_{k=1}^{\infty}\left\{\sum_{i=1}^{\infty}B^{i-1}BB^{k-i}\right\}\underline{c}_{t-k}=\sum_{k=1}^{\infty}kB^k\underline{c}_{t-k}. \tag{2.43}$$

En utilisant (2.24), nous avons $\|B^K\|\leq K\rho^k$ pour tout k. De plus ε_t^2 possède un moment d'ordre $s\in]0,1[$, il en est donc de même pour $\underline{c}_t(1)=\omega+\sum_{i=1}^{q}\alpha_i\varepsilon_{t-i}^2$. En utilisant de plus

(2.43), les minorations $\sigma_t^2 \geq \omega + B^k(1,1)\underline{c}_{t-k}(1)$ et la relation $x/(1+x) \leq x^s$ pour tout $x \geq 0$, on obtient :

$$\begin{aligned}
\mathbb{E}_{\theta_0} \frac{1}{\sigma_t^2} \frac{\partial \sigma_t^2}{\partial \beta_j} &\leq \mathbb{E}_{\theta_0} \frac{1}{\beta_j} \sum_{k=1}^{\infty} \frac{k B^k(1,1)\underline{c}_{t-k}(1)}{\omega + B^k(1,1)\underline{c}_{t-k}(1)} \\
&\leq \frac{1}{\beta_j} \sum_{k=1}^{\infty} k \mathbb{E}_{\theta_0} \left\{ \frac{B^k(1,1)\underline{c}_{t-k}(1)}{\omega} \right\}^s \\
&\leq \frac{K^s}{\omega^s \beta_j} \mathbb{E}_{\theta_0} \{\underline{c}_{t-k}(1)\}^s \sum_{k=1}^{\infty} k \rho^{sk} \leq \frac{K}{\beta_j}. \qquad (2.44)
\end{aligned}$$

Sous l'hypothèse **H5** on a $\beta_{0j} > 0$ pour tout j, ce qui permet de conclure que la première espérance dans (2.39) existe.
Regardons maintenant les dérivées d'ordre supérieur de σ_t^2. D'après la première égalité de (2.40), on a :

$$\frac{\partial^2 \underline{\sigma}_t^2}{\partial \omega} = \frac{\partial^2 \underline{\sigma}_t^2}{\partial \omega \partial \alpha_i} = 0, \quad \frac{\partial^2 \underline{\sigma}_t^2}{\partial \omega \partial \beta_j} = \sum_{k=1}^{\infty} \left\{ \sum_{i=1}^{k} B^{i-1} B^{(j)} B^{k-i} \right\} \underline{1}. \qquad (2.45)$$

On a donc :

$$\beta_j \frac{\partial^2 \underline{\sigma}_t^2}{\partial \omega \partial \beta_j} \leq \sum_{k=1}^{\infty} k B^k \underline{1},$$

qui est un vecteur de constantes finies (puisque $\rho(B) < 1$). On en déduit que $\partial^2 \sigma_t^2(\theta_0)/\partial \omega \partial \theta_i$ est bornée et admet donc des moments de tous ordres. Il en est biensûr de même pour $\{\partial^2 \sigma_t^2(\theta_0)/\partial \omega \partial \theta_i\}/\sigma_t^2(\theta_0)$. la seconde égalité de (2.40) donne :

$$\frac{\partial^2 \underline{\sigma}_t^2}{\partial \alpha_i \partial \alpha_j} = 0, \quad \frac{\partial^2 \underline{\sigma}_t^2}{\partial \alpha_i \partial \alpha_j} = \sum_{k=1}^{\infty} \left\{ \sum_{i=1}^{k} B^{i-1} B^{(j)} B^{k-i} \right\} \underline{\varepsilon}_{t-k-i}. \qquad (2.46)$$

Les arguments utilisés pour montrer (2.43) donnent :

$$\mathbb{E}_{\theta_0} \frac{\partial^2 \sigma_t^2 / \partial \alpha_i \partial \beta_j}{\sigma_t^2} \leq \frac{K^*}{\beta_j}.$$

Ceci implique que $\{\partial^2 \sigma_t^2(\theta_0)/\partial \alpha_i \partial \theta\}/\sigma_t^2(\theta_0)$ est intégrable. La dérivation de (2.41) par rapport à $\beta_{j'}$ donne :

$$\begin{aligned}
\beta_j \beta_{j'} \frac{\partial^2 \underline{\sigma}_t^2}{\partial \beta_j \partial \beta_{j'}} &= \beta_j \beta_{j'} \sum_{k=2}^{\infty} \left[\sum_{i=2}^{k} \left\{ \left(\sum_{\ell=1}^{i-1} B^{\ell-1} B^{(j')} B^{i-1-\ell} \right) B^{(j)} B^{k-i} \right\} \right. \\
&\quad \left. + \sum_{i=1}^{k-1} \left\{ B^{i-1} B^(j) \left(\sum_{\ell=1}^{k-i} B^{\ell-1} B^{(j')} B^{k-i-\ell} \right) \right\} \right] \underline{c}_{t-k} \\
&\leq \sum_{k=2}^{\infty} \left[\sum_{i=2}^{k} (i-1) B^k + \sum_{i=1}^{k-1} (k-i) B^k \right] \underline{c}_{t-k} = \sum_{k=2}^{\infty} k(k-1) B^k \underline{c}_{t-k}, \qquad (2.47)
\end{aligned}$$

car $\beta_j B^{(j)} \leq B$. En utilisant le même argument pour (2.43), on déduit :

$$\mathbb{E}_{\theta_0} \frac{1}{\sigma_t^2} \frac{\partial^2 \sigma_t^2}{\partial \beta_j \partial \beta_j'} \leq \frac{K^*}{\beta_j \beta_j'},$$

et l'existence de la deuxième espérance dans (2.39) est prouvée. Par ailleurs, puisque $\{\partial^2 \sigma_t^2 / \partial \omega\} / \sigma_t^2$ est bornée, et par (2.42), les variables $\{\partial \sigma_t^2 / \partial \alpha_i\} / \sigma_t^2$ sont bornées en θ_0 il est claire que :

$$\mathbb{E}_{\theta_0} \left\| \frac{1}{\sigma_t^4(\theta_0)} \frac{\partial \sigma_t^2(\theta_0)}{\partial \theta_i} \frac{\partial \sigma_t^2(\theta_0)}{\partial \theta} \right\| < \infty,$$

pour $i = 1, \cdots, q+1$. Avec des notations et arguments déjà utilisés pour montrer (2.43), et en utilisant l'inégalité élémenàtaire $x/(1+x) \leq x^{s/2}$ pour tout $x \geq 0$, l'inégalité de Minkowski donne :

$$\left\{ \mathbb{E}_{\theta_0} \left(\frac{1}{\sigma_t^2(\theta_0)} \frac{\partial \sigma_t^2(\theta_0)}{\partial \beta_j} \right)^2 \right\}^{1/2} \leq \frac{1}{\beta_{0j}} \sum_{k=1}^{\infty} k \left\{ \mathbb{E}_{\theta_0} \left(\frac{B^k(1,1) \underline{c}_{t-k}(1)}{\omega_0} \right)^s \right\}^{1/2} < \infty.$$

Finalement, l'inégalité de Cauchy-schwartz permet de conclure que la troisième espérance de (2.39) existe.

-ii) **Inversibilité de J et lien avec la variance de la dérivée du critère.**

D'après (2.37) et (i) on a :

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

tel que : $\varepsilon_t^2 / \sigma_t^2(\theta_0) = \eta_t^2$ est indépendant des termes en σ_t^2 ou en ses dérivées.
En utilisant (2.38) et (i), J existe et (2.33). Nous avons également :

$$\begin{aligned} V_{\theta_0} \left\{ \frac{\partial \ell_t(\theta_0)}{\partial \theta} \right\} &= \mathbb{E}_{\theta_0} \left\{ \frac{\partial \ell_t(\theta_0)}{\partial \theta} \frac{\partial \ell_t(\theta_0)}{\partial \theta'} \right\}. \\ &= \mathbb{E}_{\theta_0} \{ (1 - \eta_t^2)^2 \} \mathbb{E}_{\theta_0} \left\{ \frac{\partial \sigma_t^2(\theta_0) / \partial \theta}{\sigma_t^2(\theta_0)} \frac{\partial \sigma_t^2(\theta_0) / \partial \theta'}{\sigma_t^2(\theta_0)} \right\} \\ &= (k_\eta - 1) J. \end{aligned} \tag{2.48}$$

Maintenant supposons que :

$$\lambda' J \lambda = \mathbb{E} \left[\frac{1}{\sigma_t^4(\theta_0)} \left(\lambda' \frac{\partial \sigma_t^2(\theta_0)}{\partial \theta} \right)^2 \right] = 0,$$

pour $\lambda \in \mathbb{R}^{p+q+1}$, tel que $\lambda' \{ \partial \sigma_t^2(\theta_0) / \partial \theta \} = 0$ *p.s.* D'après (2.21) et la stationnarité de $\{ \partial \sigma_t^2(\theta_0) / \partial \theta \}$, on a :

$$0 = \lambda' \frac{\partial \sigma_t^2(\theta_0)}{\partial \theta} = \lambda' \begin{pmatrix} 1 \\ \varepsilon_{t-1}^2 \\ \vdots \\ \varepsilon_{t-q}^2 \\ \sigma_{t-1}^2(\theta_0) \\ \vdots \\ \sigma_{t-p}^2(\theta_0) \end{pmatrix} + \sum_{j=1}^{p} \beta_j \lambda' \frac{\partial \sigma_{t-j}^2(\theta_0)}{\partial \theta} = \lambda' \begin{pmatrix} 1 \\ \varepsilon_{t-1}^2 \\ \vdots \\ \varepsilon_{t-q}^2 \\ \sigma_{t-1}^2(\theta_0) \\ \vdots \\ \sigma_{t-p}^2(\theta_0) \end{pmatrix}.$$

Posons $\lambda = (\lambda_0, \lambda_1, \ldots, \lambda_{q+p})$. Il est clair que $\lambda_1 = 0$, sinon ε_{t-1}^2 serait mesurable par rapport à la tribu engendrée par $\{\eta_u, u < t-1\}$. Pour la même raison on a : $\lambda_2 = \cdots = \lambda_{2+i} = 0$ si $\lambda_{q+1} = \cdots = \lambda_{q+i} = 0$. Par conséquent, $\lambda \neq 0$ implique une représentation GARCH$(p-1, q-1)$. Ceci est impossible en raison de **H4** en argumentant comme pour établir (2.29). Par suite $\lambda' J \lambda = 0$ implique $\lambda = 0$, donc on a prouvé ii).
-iii) **Intégrabilité uniforme des dérivées d'ordre 3 du critère.**
En dérivant (2.38), nous obtenons :

$$\begin{aligned}
\frac{\partial^3 \ell_t(\theta)}{\partial\theta_i \partial\theta_j \partial\theta_k} &= \left\{1 - \frac{\varepsilon_t^2}{\sigma_t^2}\right\}\left\{\frac{1}{\sigma_t^2}\frac{\partial^3 \sigma_t^2}{\partial\theta_i \partial\theta_j \partial\theta_k}\right\} \qquad (2.49)\\
&+\left\{2\frac{\varepsilon_t^2}{\sigma_t^2} - 1\right\}\left\{\frac{1}{\sigma_t^2}\frac{\partial \sigma_t^2}{\partial\theta_i}\right\}\left\{\frac{1}{\sigma_t^2}\frac{\partial^2 \sigma_t^2}{\partial\theta_j \partial\theta_k}\right\}\\
&+\left\{2\frac{\varepsilon_t^2}{\sigma_t^2} - 1\right\}\left\{\frac{1}{\sigma_t^2}\frac{\partial \sigma_t^2}{\partial\theta_j}\right\}\left\{\frac{1}{\sigma_t^2}\frac{\partial^2 \sigma_t^2}{\partial\theta_i \partial\theta_k}\right\}\\
&+\left\{2\frac{\varepsilon_t^2}{\sigma_t^2} - 1\right\}\left\{\frac{1}{\sigma_t^2}\frac{\partial \sigma_t^2}{\partial\theta_k}\right\}\left\{\frac{1}{\sigma_t^2}\frac{\partial^2 \sigma_t^2}{\partial\theta_i \partial\theta_j}\right\}\\
&+\left\{2 - 6\frac{\varepsilon_t^2}{\sigma_t^2}\right\}\left\{\frac{1}{\sigma_t^2}\frac{\partial \sigma_t^2}{\partial\theta_i}\right\}\left\{\frac{1}{\sigma_t^2}\frac{\partial \sigma_t^2}{\partial\theta_j}\right\}\left\{\frac{1}{\sigma_t^2}\frac{\partial \sigma_t^2}{\partial\theta_k}\right\}.
\end{aligned}$$

Commençons de montrer que $\{1 - \varepsilon_t^2/\sigma_t^2\}$ est intégrable. C'est le terme le plus déficile à traiter. En effet nous n'avons pas intégrabilité uniformément de $\varepsilon_t^2/\sigma_t^2$ sur Θ : en $\theta = (\omega, 0, \ldots, 0)'$, le rapport $\varepsilon_t^2/\sigma_t^2$ n'est pas intégrable si $\mathbb{E}(\varepsilon_t^2) = \infty$. Nous allons cependant montrer que $\{1 - \varepsilon_t^2/\sigma_t^2\}$ est uniformément intégrable au voisinage de θ_0. Soit Θ^* un compact contenant θ_0 et contenu dans l'intérieur de Θ. Notons B_0 la matice B (définie en (2.23)) évaluée au point $\theta = \theta_0$. Pour tout $\delta > 0$, il existe un voisinage $V(\theta_0)$ de θ_0, tel que $V(\theta_0) \subseteq \Theta^*$, tel que $B_0 \leq (1+\delta)B$ pour tout $\theta \in V(\theta_0)$.
D'après (2.25), on obtient :

$$\sigma_t^2 = \omega \sum_{k=0}^{\infty} B^k(1,1) + \sum_{k=1}^{q} \alpha_i \left\{\sum_{k=0}^{\infty} B^k(1,1)\varepsilon_{t-k-i}^2\right\}.$$

Puisque $V(\theta_0) \subset \Theta^*$, on a $\sup_{\theta \in V(\theta_0)} 1/\alpha_i < \infty$. En utilisant encore $x/(1+x) \leq x^s$ pour tout $x \geq 0$ et tout $s \in \geq]0,1[$, on obtient :

$$\begin{aligned}
\sup_{\theta \in V(\theta_0)} \frac{\sigma_t^2(\theta_0)}{\sigma_t^2} &\leq \sup_{\theta \in V(\theta_0)} \left\{\frac{\omega_0 \sum_{k=0}^{\infty} B_0^k(1,1)}{\omega} + \sum_{i=1}^{q} \alpha_{0i}\left(\sum_{k=0}^{\infty} \frac{B_0^k(1,1)\varepsilon_{t-k-i}^2}{\omega + \alpha_i B^k(1,1)\varepsilon_{t-k-i}^2}\right)\right\}\\
&\leq K + \sum_{i=1}^{q} \sup_{\theta \in V(\theta_0)} \left\{\frac{\alpha_{0i}}{\alpha_i}\sum_{k=0}^{\infty} \frac{B_0^k(1,1)}{B^k(1,1)}\left(\frac{\alpha_i B^k(1,1)\varepsilon_{t-k-i}^2}{\omega}\right)^s\right\}\\
&\leq K + K\sum_{i=1}^{q}\sum_{k=0}^{\infty}(1+\delta)^k \rho^{ks}\varepsilon_{t-k-1}^{2s}. \qquad (2.50)
\end{aligned}$$

si on choisit s tel que $\mathbb{E}(\varepsilon_t^{2s}) < \infty$ et par exemple, $\delta = (1-\rho^s)/(2\rho^s)$ et $V(\theta_0)\theta$ voisinage de θ_0, on obtient :

$$\mathbb{E}_{\theta_0} \sup_{\theta \in V(\theta_0)} \frac{\varepsilon_t^2}{\sigma_t^2} = \mathbb{E}_{\theta_0} \sup_{\theta \in V(\theta_0)} \frac{\sigma_t^2(\theta_0)}{\sigma_t^2} < \infty.$$

Avec même choix de δ, en choisissant s tel que $\mathbb{E}(\varepsilon_t^{4s}) < \infty$ et en utilisant (2.50), on obtient :

$$\begin{aligned}\left\| \sup_{\theta \in V(\theta_0)} \frac{\varepsilon_t^2}{\sigma_t^2} \right\|_2 &= k_\eta^{1/2} \left\| \sup_{\theta \in V(\theta_0)} \frac{\sigma_t^2(\theta_0)}{\sigma_t^2} \right\|_2 \\ &\leq k_\eta^{1/2} K + k_\eta^{1/2} K_q \sum_{k=0}^{\infty} (1+\delta)^k \rho^{ks} \|\varepsilon_t^{2s}\|_2 < \infty. \qquad (2.51)\end{aligned}$$

Maintenant nous étudions le deuxième terme entre accolades dans (2.49). En dérivant (2.45), (2.46), à l'aide des utilisés pour montrer (2.42) on obtient :

$$\sup_{\theta \in \Theta*} \frac{1}{\sigma_t^2} \frac{\partial^3 \sigma_t^2}{\partial\theta_{i_1} \partial\theta_{i_2} \partial\theta_{i_3}} \leq K,$$

lorsque les indices i_1, i_2 et i_3 ne sont pas tous dans $\{q+1, q+2, \ldots, q+1+p\}\}$ (i.e. quand on dérive par au moins un paramètre autre qu'un des β_j). En reprenant les arguments utilisés pour montrer (2.43) et (2.44), on obtient :

$$\beta_i \beta_j \beta_k \frac{\partial^3 \sigma_t^2}{\partial\beta_i \partial\beta_j \partial\beta_k} \leq \sum_{k=3}^{\infty} k(k-1)(k-2) B^k(1,1) \underline{c}_{t-k}(1),$$

$$\sup_{\theta \in \Theta*} \frac{1}{\sigma_t^2} \frac{\partial^3 \sigma_t^2}{\partial\beta_i \partial\beta_j \partial\beta_k} \leq K \left\{ \sup_{\theta \in \Theta*} \frac{1}{\omega^s \beta_i \beta_j \beta_k} \right\} \sum_{k=3}^{\infty} k(k-1)(k-2) \rho^{ks} \left\{ \sup_{\theta \in \Theta*} \underline{c}_{t-k}(1) \right\}^s,$$

pour tout $s \in]0,1[$. Tel que $\mathbb{E}_{\theta_0} \{\sup_{\theta \in \Theta*} \underline{c}_{t-k}(1)\}^{2s} < \infty$ pour $s > 0$, on déduit :

$$\mathbb{E}_{\theta_0} \sup_{\theta \in \Theta^*} \left| \frac{1}{\sigma_t^2} \frac{\partial^3 \sigma_t^2}{\partial\theta_i \partial\theta_j \partial\theta_k} \right|^2 < \infty. \qquad (2.52)$$

Généralement, on peut remplacer la puissance 2 par la puissance d arbitrairement grande :

$$\mathbb{E}_{\theta_0} \sup_{\theta \in \Theta^*} \left| \frac{1}{\sigma_t^2} \frac{\partial^3 \sigma_t^2}{\partial\theta_i \partial\theta_j \partial\theta_k} \right|^d < \infty. \qquad (2.53)$$

Par l'inégalité de Cauchy-Schwartz, (2.51) et (2.52) on obtient :

$$\mathbb{E}_{\theta_0} \sup_{\theta \in V(\theta_0)} \left| \left\{ 1 - \frac{\varepsilon_t^2}{\sigma_t^2} \right\} \left\{ \frac{1}{\sigma_t^2} \frac{\partial^3 \sigma_t^2}{\partial\theta_i \partial\theta_j \partial\theta_k} \right\} \right| < \infty.$$

Les autres terme entre parenthèses in (2.49) se traitent similaire à (2.53) on montre que :

$$\mathbb{E}_{\theta_0} \sup_{\theta \in \Theta^*} \left| \frac{1}{\sigma_t^2} \frac{\partial^2 \sigma_t^2}{\partial\theta_i \partial\theta_j} \right|^d < \infty, \quad \mathbb{E}_{\theta_0} \sup_{\theta \in \Theta^*} \left| \frac{1}{\sigma_t^2} \frac{\partial \sigma_t^2}{\partial\theta_i} \right|^d < \infty, \qquad (2.54)$$

pour tout entier d. Ceci permet d'établir à l'aide de l'inégalité de Hölder que :

$$\mathbb{E}_{\theta_0} \sup_{\theta \in V(\theta_0)} \left| \left\{ 2 - 6\frac{\varepsilon_t^2}{\sigma_t^2} \right\} \left\{ \frac{1}{\sigma_t^2} \frac{\partial \sigma_t^2}{\partial \theta_i} \right\} \left\{ \frac{1}{\sigma_t^2} \frac{\partial \sigma_t^2}{\partial \theta_j} \right\} \left\{ \frac{1}{\sigma_t^2} \frac{\partial \sigma_t^2}{\partial \theta_k} \right\} \right|$$

$$\leq \left\| \sup_{\theta \in V(\theta_0)} \left| 2 - 6\frac{\varepsilon_t^2}{\sigma_t^2} \right| \right\|_2 \max_i \left\| \sup_{\theta \in \Theta^*} \left| \frac{1}{\sigma_t^2} \frac{\partial \sigma_t^2}{\partial \theta_i} \right| \right\|_6^3 < \infty.$$

Les autres termes de la somme dans (2.49) se traitent de même façon. donc on a montré (iii).
-iv) **Oubli asymptotique des valeurs initiales :**
Pour prouver (iv), en utilisant (2.26), on obtient les equations analogues à (2.40)et(2.41) pour les dérivées de $\underline{\widetilde{\sigma}}_t^2$:

$$\frac{\partial \underline{\widetilde{\sigma}}_t^2}{\partial \omega} = \sum_{k=0}^{t-1-q} B^k \underline{1} + \sum_{k=1}^{q} B^{t-k} \frac{\partial \underline{\widetilde{c}}_k}{\partial \omega} + B^t \frac{\partial \underline{\widetilde{\sigma}}_0^2}{\partial \omega}, \tag{2.55}$$

$$\frac{\partial \underline{\widetilde{\sigma}}_t^2}{\partial \alpha_i} = \sum_{k=0}^{t-1-q} B^k \underline{\varepsilon}_{t-k-i}^2 + \sum_{k=1}^{q} B^{t-k} \frac{\partial \underline{\widetilde{c}}_k}{\partial \alpha_i} + B^t \frac{\partial \underline{\widetilde{\sigma}}_0^2}{\partial \alpha_i}, \tag{2.56}$$

$$\frac{\partial \underline{\widetilde{\sigma}}_t^2}{\partial \beta_j} = \sum_{k=1}^{t-1-q} \left\{ \sum_{i=1}^{k} B^{i-1} B^{(j)} B^{k-i} \right\} \underline{c}_{t-k} + \sum_{k=1}^{q} \left\{ \sum_{i=1}^{t-k} B^{i-1} B^{(j)} B^{t-k-i} \right\} \underline{\widetilde{c}}_k. \tag{2.57}$$

Où $\partial \underline{\widetilde{\sigma}}_0^2 / \partial \omega = (0, \dots, 0)'$ quand les conditions initiales sont données par (2.14) et vant $(1, \dots, 1)'$ quand les conditions initiales sont donnée par (2.13). Les dérivées secondes ont des expressions similaires. la compacité de Θ, et le fait que $\rho(B) < 1$ permettent d'affirmer que, presque sûrement

$$\sup_{\theta \in \Theta} \left\| \frac{\partial \sigma_t^2}{\partial \theta} - \frac{\partial \widetilde{\sigma}_t^2}{\partial \theta} \right\| < K\rho^t, \quad \sup_{\theta \in \Theta} \left\| \frac{\partial^2 \theta_t^2}{\partial \theta \partial \theta'} - \frac{\partial^2 \widetilde{\sigma}_t^2}{\partial \theta \partial \theta'} \right\| < K\rho^t, \quad \forall\, t \in \mathbb{Z} \tag{2.58}$$

d'aprés (2.27), on obtient :

$$\left| \frac{1}{\sigma_t^2} - \frac{1}{\widetilde{\sigma}_t^2} \right| = \left| \frac{\widetilde{\sigma}_t^2 - \sigma_t^2}{\sigma_t^2 \widetilde{\sigma}_t^2} \right| \leq \frac{K\rho^t}{\sigma_t^2}, \quad \frac{\sigma_t^2}{\widetilde{\sigma}_t^2} \leq 1 + K\rho^t, \tag{2.59}$$

puisque

$$\frac{\partial \widetilde{\ell}_t(\theta)}{\partial \theta} = \left\{ 1 - \frac{\varepsilon_t^2}{\widetilde{\sigma}_t^2} \right\} \left\{ \frac{1}{\widetilde{\sigma}_t^2} \frac{\partial \widetilde{\sigma}_t^2}{\partial \theta} \right\} \quad et \quad \frac{\partial \ell_t(\theta)}{\partial \theta} = \left\{ 1 - \frac{\varepsilon_t^2}{\sigma_t^2} \right\} \left\{ \frac{1}{\sigma_t^2} \frac{\partial \sigma_t^2}{\partial \theta} \right\},$$

en utilisant (2.59) et la première inégalité dans (2.58),

$$\left| \frac{\partial \ell_t(\theta_0)}{\partial \theta_i} - \frac{\partial \widetilde{\ell}_t(\theta_0)}{\partial \theta_i} \right| = \left| \left\{ \frac{\varepsilon_t^2}{\widetilde{\sigma}_t^2} - \frac{\varepsilon_t^2}{\sigma_t^2} \right\} \left\{ \frac{1}{\sigma_t^2} \frac{\partial \sigma_t^2}{\partial \theta_i} \right\} + \left\{ 1 - \frac{\varepsilon_t^2}{\widetilde{\sigma}_t^2} \right\} \left\{ \frac{1}{\sigma_t^2} - \frac{1}{\widetilde{\sigma}_t^2} \right\} \left\{ \frac{\partial \sigma_t^2}{\partial \theta_i} \right\} \right.$$

$$\left. + \left\{ 1 - \frac{\varepsilon_t^2}{\widetilde{\sigma}_t^2} \right\} \left\{ \frac{1}{\widetilde{\sigma}_t^2} \right\} \left\{ \frac{\partial \sigma_t^2}{\partial \theta_i} - \frac{\partial \widetilde{\sigma}_t^2}{\partial \theta_i} \right\} \right| (\theta_0)$$

$$\leq K\rho^t (1 + \eta_t^2) \left| 1 + \left\{ \frac{1}{\sigma_t^2(\theta_0)} \frac{\partial \sigma_t^2(\theta_0)}{\partial \theta_i} \right\} \right|.$$

Par suite

$$\left|n^{-1/2}\sum_{t=1}^{n}\left\{\frac{\partial \ell_t(\theta_0)}{\partial\theta_i}-\frac{\partial\widetilde{\ell}_t(\theta_0)}{\partial\theta_i}\right\}\right|\leq K^*n^{-1/2}\sum_{t=1}^{n}\rho^t(1+\eta_t^2)\left|1+\left\{\frac{1}{\sigma_t^2(\theta_0)}\frac{\partial\sigma_t^2(\theta_0)}{\partial\theta_i}\right|.\qquad(2.60)$$

L'inégalité de Markov ,(2.39) et l'indépendance entre η_t et $\sigma_t^2(\theta_0)$ implique que, pour tout $\epsilon>0$,

$$\mathbb{P}\left(n^{-1/2}\sum_{t=1}^{n}\rho^t(1+\eta_t^2)\left|1+\frac{1}{\sigma_t^2(\theta_0)}\frac{\partial\sigma_t^2(\theta_0)}{\partial\theta}\right|>\epsilon\right)$$
$$\leq\frac{2}{\epsilon}\left(1+\mathbb{E}_{\theta_0}\left|\frac{1}{\sigma_t^2(\theta_0)}\frac{\partial\sigma_t^2(\theta_0)}{\partial\theta}\right|\right)n^{-1/2}\sum_{t=1}^{n}\rho^t\to 0,$$

tel que, par (2.60) montre la première partie de (iv).
Maintenant regardond l'influence asymptotique des valeurs initiales sur les dérivées secondes-ordres du critère en un voisinage de θ_0. D'après (2.38) et les majorations précédentes , nous trouvons :

$$\sup_{\theta\in V(\theta_0)}\left|n^{-1}\sum_{t=1}^{n}\left\{\frac{\partial^2\ell_t(\theta_0)}{\partial\theta_i\partial\theta_j}-\frac{\partial^2\widetilde{\ell}_t(\theta_0)}{\partial\theta_i\partial\theta_j}\right\}\right|\leq n^{-1}\sum_{t=1}^{n}\sup_{\theta\in V(\theta_0)}\left\{\frac{\varepsilon_t^2}{\widetilde{\sigma}_t^2}-\frac{\varepsilon_t^2}{\sigma_t^2}\right\}\left\{\frac{1}{\sigma_t^2}\frac{\partial^2\sigma_t^2}{\partial\theta_i\partial\theta_j}\right\}$$
$$+\left\{1-\frac{\varepsilon_t^2}{\widetilde{\sigma}_t^2}\right\}\left\{\left(\frac{1}{\sigma_t^2}-\frac{1}{\widetilde{\sigma}_t^2}\right)\frac{\partial^2\sigma_t^2}{\partial\theta_i\partial\theta_j}+\frac{1}{\widetilde{\sigma}_t^2}\left(\frac{\partial^2\sigma_t^2}{\partial\theta_i\partial\theta_j}-\frac{\partial^2\widetilde{\sigma}_t^2}{\partial\theta_i\partial\theta_j}\right)\right\}$$
$$+\left\{2\frac{\varepsilon_t^2}{\sigma_t^2}-2\frac{\varepsilon_t^2}{\widetilde{\sigma}_t^2}\right\}\left\{\frac{1}{\sigma_t^2}\frac{\partial\sigma_t^2}{\partial\theta_i}\right\}\left\{\frac{1}{\sigma_t^2}\frac{\partial\sigma_t^2}{\partial\theta_j}\right\}$$
$$+\left\{2\frac{\varepsilon_t^2}{\widetilde{\sigma}_t^2}-1\right\}\left\{\left(\frac{1}{\sigma_t^2}-\frac{1}{\widetilde{\sigma}_t^2}\right)\frac{\partial\sigma_t^2}{\partial\theta_i}+\frac{1}{\widetilde{\sigma}_t^2}\left(\frac{\partial\sigma_t^2}{\partial\theta_i}-\frac{\partial\widetilde{\sigma}_t^2}{\partial\theta_i}\right)\right\}\left\{\frac{1}{\sigma_t^2}\frac{\partial\sigma_t^2}{\partial\theta_j}\right\}$$
$$+\left\{2\frac{\varepsilon_t^2}{\widetilde{\sigma}_t^2}-1\right\}\left\{\frac{1}{\widetilde{\sigma}_t^2}\frac{\partial\widetilde{\sigma}_t^2}{\partial\theta_i}\right\}\left\{\left(\frac{1}{\sigma_t^2}-\frac{1}{\widetilde{\sigma}_t^2}\right)\frac{\partial\sigma_t^2}{\partial\theta_j}+\frac{1}{\widetilde{\sigma}_t^2}\left(\frac{\partial\sigma_t^2}{\partial\theta_j}-\frac{\partial\widetilde{\sigma}_t^2}{\partial\theta_j}\right)\right\}$$
$$\leq Kn^{-1}\sum_{t=1}^{n}\rho^t\Upsilon_t,$$

où

$$\Upsilon_t=\sup_{\theta\in V(\theta_0)}\left\{1+\frac{\varepsilon_t^2}{\sigma_t^2}\right\}\left\{1+\frac{1}{\sigma_t^2}\frac{\partial^2\sigma_t^2}{\partial\theta_j\partial\theta_i}+\frac{1}{\sigma_t^2}\frac{\partial\sigma_t^2}{\partial\theta_i}\frac{1}{\sigma_t^2}\frac{\partial\sigma_t^2}{\partial\theta_j}\right\}.$$

D'après (2.51), (2,54) et l'énégalité de Hölder, on voit que, pour un certain voisinage $V(\theta_0)$, l'espérance de Υ_t est une constante finie. En utilisant à nouveau l'inégalité de Markov on montre alors la seconde convergence de (iv).
(v)**Utilisation d'un TCL pour accroissements de martingale.**
Rappelons que ε_u, $u<t$ désigne la tribu engendrée par les variables ε_{t-i}, $i\geq 0$. Le vecteur des scores conditionnels est évidemment centré, ce qui peut se retrouve directement à partir de (2.36), en utilisant le fait que $\sigma_t^2(\theta_0)\in\varepsilon_u$, $u<t$ et $\mathbb{E}_{\theta_0}(\varepsilon_t^2|\varepsilon_u,u<t)=\sigma_t^2(\theta_0)$:

$$\mathbb{E}_{\theta_0}\left(\frac{\partial}{\partial\theta}\ell_t(\theta_0)|\varepsilon_u,u<t\right)=\frac{1}{\sigma_t^4(\theta_0)}\left(\frac{\partial}{\partial\theta}\sigma_t^2(\theta_0)\right)\mathbb{E}_{\theta_0}(\sigma_t^2(\theta_0)-\varepsilon_t^2|\varepsilon_u,u<t)=0.$$

Notons également, d'après (2.48), $V_{\theta_0}(\partial \ell_t(\theta_0)/\partial\theta) = (k_\eta - 1)J$ est finie. D'après l'invérsibilité de J et les hypothèses sur la loi de η_t (qui entraînent $0 < k_\eta - 1 < \infty$), cette matrice de variance est non dégénérée. Nous déduisons que $\forall\ \lambda \in \mathbb{R}^{p+q+1}$, la suite $\{\lambda' \frac{\partial}{\partial\theta}\ell_t(\theta_0), \varepsilon_t\}_t$ est une différence de martingale stationnaire ergodique de carré intégrable :

$$\begin{aligned} V_{\theta_0}\left(\lambda' \frac{\partial \ell_t(\theta_0)}{\partial\theta}\right) &= \lambda' \mathbb{E}_{\theta_0}\left(\frac{\partial \ell_t(\theta_0)}{\partial\theta}\right)\lambda \\ &= \lambda' \mathbb{E}_{\theta_0}\left\{(1-\eta_t^2)^2 \frac{1}{\sigma_t^4(\theta_0)} \frac{\partial \sigma_t \sigma^2(\theta_0)}{\partial\theta} \frac{\partial \sigma_t \sigma^2(\theta_0)}{\partial\theta'}\right\}\lambda \\ &= (k_\eta - 1)\lambda' J \lambda. \end{aligned}$$

En utilisant le (T.C.L) pour les différences de martingale, on aura :

$$n^{-1/2} \sum_{t=1}^{n} \left\{\lambda' \frac{\partial \ell_t(\theta_0)}{\partial\theta}\right\} \xrightarrow{\mathcal{L}} \mathcal{N}(0, (k_\eta - 1)\lambda' J \lambda),$$

par suite, en aplliquant le théorème de Wold-Cramer

$$n^{-1/2} \sum_{t=1}^{n} \left\{\frac{\partial \ell_t(\theta_0)}{\partial\theta}\right\} \xrightarrow{\mathcal{L}} \mathcal{N}(0, (k_\eta - 1)J.$$

iv)**Utilisation d'un second développement limité et du théorème ergodique.**
Reprenons le développement de Taylor (2.33) du critère en θ_0. On a, pour tout i et j

$$n^{-1} \sum_{t=1}^{n} \frac{\partial^2}{\partial\theta_i \partial\theta_j} \ell_t(\theta_{ij}^*) = n^{-1} \sum_{t=1}^{n} \frac{\partial^2}{\partial\theta_i \partial\theta_j} \ell_t(\theta_0) + n^{-1} \sum_{t=1}^{n} \frac{\partial}{\partial\theta'}\left\{\frac{\partial^2}{\partial\theta_i \partial\theta_j} \ell_t(\widetilde{\theta}_{ij})\right\}(\theta_{ij}^* - \theta_0), \quad (2.61)$$

où $\widetilde{\theta}_{ij}$ est entre θ_{ij}^* et θ_0. La convergence presque sûre de $\widetilde{\theta}_{ij}$ vers θ_0, le théorème ergodique et (iii) impliquent que p.s.

$$\begin{aligned} \limsup_{n\to\infty} \left\| n^{-1} \sum_{t=1}^{n} \frac{\partial}{\partial\theta'}\left\{\frac{\partial^2}{\partial\theta_i \partial\theta_j} \ell_t(\widetilde{\theta}_{ij})\right\}\right\| &\leq \limsup_{n\to\infty} n^{-1} \sum_{t=1}^{n} \sup_{\theta\in V(\theta_0)} \left\| \frac{\partial}{\partial\theta'}\left\{\frac{\partial^2}{\partial\theta_i \partial\theta_j} \ell_t(\theta)\right\}\right\| \\ &= \mathbb{E}_{\theta_0} \sup_{\theta\in V(\theta_0)} \left\| \frac{\partial}{\partial\theta'}\left\{\frac{\partial^2}{\partial\theta_i \partial\theta_j} \ell_t(\theta)\right\}\right\| < \infty. \end{aligned}$$

Puisque $\|\theta_{ij}^* - \theta_0\| \longrightarrow 0$ presque sûrement, le seconde terme du membre de droite de (2.61) converge vers 0 avec probabilité 1. La convergence dans (vi) résulte du théorème ergodique appliqué au premier terme du membre droite de (2.61). Pour achever la preuve du théorème (2.2.2) il suffit d'appliquer le lemme de slutsky. Par (iv), (v) et (vi) nous obtenons (2.35) et (2.36).

□

2.3 Méthode des deux étapes :

Le but de cette section est d'une part l'élaboration d'une méthode en deux étapes, pour l'estimation des modèles GARCH , et d'une part, l'analyse des propriétés statistiques des estimateurs fourni par cette méthode.
L'avantage de cet estimateur est de posséder une formule explicite, contrairement à la méthode Q.M.V déjà vue. Toutefois, il est important de signaler que l'étude du comportement asymptotique nécessite des moments d'ordre élevés.
Afin d'introduire l'estimateur préliminaire, posons :

$$\varepsilon_t = \sigma_{t-1}(\beta)\eta_t, \quad 1 \leq t \leq n, \tag{2.62}$$

$$Y_t = \varepsilon_t^2, \quad pour \quad 1-p \leq t \leq n,$$

$$\mathbf{Z}_{t-1} = [1, Y_{t-1}, \ldots, Y_{t-p}]' = [1, \varepsilon_{t-1}^2, \ldots, \varepsilon_{t-p}^2]',$$

et $u_t = \eta_t^2 - 1, \quad 1 \leq t \leq n$. Ainsi :

$$\sigma_{t-1}^2(\beta) = \mathbf{Z}_{t-1}^{'}\beta, \tag{2.63}$$

où $\beta = [\beta_0, \beta_1, \ldots, \beta_p]$.
Étape1 : Construction de l'estimateur préliminaire
En remplaçant $\sigma_{t-1}^2(\beta)$ dans (2.62), on aura :

$$Y_t = \mathbf{Z}_{t-1}^{'}\beta + \sigma_{t-1}r(\beta)u_t, \quad 1 \leq t \leq n, \tag{2.64}$$

où $\mathbb{E}[\sigma_{t-1}^2(\beta)u_t] = \mathbb{E}[\sigma_{t-1}^2(\beta)]\mathbb{E}[u_t] = 0, \quad 1 \leq t \leq n.$
On retrouve ainsi dans l'équation (2.64) la structure linéaire du modèle autorégressif d'ordre p dont l'innovation est nécessairement centrée.
En ignorant la partie aléatoire dans $\sigma_{t-1}^2(\beta)$ et la présence de β dans son expression, on obtient par les M.C.O un estimateur préliminaire que nous notons

$$\widehat{\beta}_{pr} = (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{Y}, \tag{2.65}$$

où $\mathbf{Z}$ est une matrice d'ordre $n \times (1+p)$ dont la t-ième ligne est $\mathbf{Z}_{t-1}^{'}$ et $\mathbf{Y}$ est le vecteur composé de $Y_t, \quad 1 \leq t \leq n$.
La démonstration de la normalité asymptotique de cet estimateur comme, celle des autres théorèmes de ce chapitre, repose sur l'application du théorème centrale limite approprié.

Étape2 : Construction de l'estimateur d'intéret
A présent, nous allons utiliser $\widehat{\beta}_{pr}$ pour construire un estimateur d'intérêt $\widehat{\beta}$ de β.
En divisant (2.64) par $\sigma_{t-1}^2(\beta)$, on obtient :

$$\frac{Y}{\sigma_{t-1}^2(\beta)} = \left\{\frac{\mathbf{Z}_{t-1}}{\sigma_{t-1}^2(\beta)}\right\}'\beta + u_t.$$

Dans cette expression, si nous remplaçons $\sigma_{t-1}^2(\beta)$ par $\sigma_{t-1}^2(\widehat{\beta}_{pr})$ on aura :

$$\frac{Y}{\sigma_{t-1}^2(\widehat{\beta}_{pr})} \approx \left\{\frac{\mathbf{Z}_{t-1}}{\sigma_{t-1}^2(\widehat{\beta}_{pr})}\right\}'\beta + u_t. \tag{2.66}$$

Par suite, en ignorant la présence de l'aléatoire dans $\sigma^2_{t-1}(\widehat{\beta}_{pr})$, alors l'equation (2.66) est similaire à la structure linéaire du modèle autorégressif. Ainsi, nous estimons β par les M.C.O, on obtient l'estimateur d'intérêt suivant :

$$\widehat{\beta} = \left[\sum_{t=1}^{n}\left\{\frac{\mathbf{Z}_{t-1}\mathbf{Z}'_{t-1}}{\sigma^4_{t-1}(\widehat{\beta}_{pr})}\right\}\right]^{-1}\left[\sum_{t=1}^{n}\left\{\frac{\mathbf{Z}_{t-1}\mathbf{Y}_t}{\sigma^4_{t-1}(\widehat{\beta}_{pr})}\right\}\right]. \tag{2.67}$$

Avant d'étudier les propriétés de l'estimateur d'intérêt, nous énoncerons d'abord un lemme qui traite l'estimateur préliminaire $\widehat{\beta}_{pr}$. Nous supposons la condition suivante sur les moments : pour tout $1 \leq j,\ k, l, m \leq p$,

$$\mathbb{E}[Y_j Y_k Y_l Y_m] < \infty. \tag{2.68}$$

L'équation (2.68) nous assure $\mathbb{E}[\mathbf{Z}_0\mathbf{Z}'_0] < \infty$ et $\mathbb{E}[(\beta'\mathbf{Z}_0)^2\mathbf{Z}_0\mathbf{Z}'_0] < \infty$.
Quand les erreurs suivent une loi gaussienne centrée réduite, les conditions nécessaires st suffisantes d'existences des moments d'ordre r (r élevé) de Y sont donnés en fonction du paramètre β.

Lemme 2.3.1. *Sous l'équation du modèle (2.62) et l'hypothèse dans (2.68), on a :*

$$\sqrt{n}(\widehat{\beta}_{pr} - \beta)\xrightarrow{\mathcal{L}}\mathcal{N}\left(0, V(\eta_1^2)[\mathbb{E}(\mathbf{Z}_0\mathbf{Z}'_0)]^{-1}\mathbb{E}[(\beta'\mathbf{Z}_0)^2\mathbf{Z}_0\mathbf{Z}'_0]\mathbb{E}(\mathbf{Z}_0\mathbf{Z}'_0)]^{-1}\right). \tag{2.69}$$

La normalité asymptotique de $\widehat{\beta}_{pr}$ se démontre en s'appuyant sur les arguments utilisés dans le théorème (2.1.2).
Afin d'énoncer le théorème qui donne la distribution asymptotique de l'estimateur d'intérêt $\widehat{\beta}$, on suppose que :

$$\sqrt{n}(\widehat{\beta}_{pr} - \beta) = O_p(1), \tag{2.70}$$

et

$$\mathbb{E}\left\{\frac{Y_j Y_{-k} Y_{-l}}{(\beta'\mathbf{Z}_0)^3}\right\} < \infty. \tag{2.71}$$

Notons les conditions du lemme (2.3.1)implique (2.70).

Théorème 2.3.1. *Sous l'équation du modèle (2.62), et sous les hypothèses (2.70) et (2.71),*

$$\sqrt{n}(\widehat{\beta} - \beta)\xrightarrow{\mathcal{L}}\mathcal{N}\left(0, \mathbf{V}(\eta_1^2)\left\{\mathbb{E}\left\{\mathbf{Z}_0\mathbf{Z}'_0(\beta'\mathbf{Z}_0)^{-2}\right\}\right\}^{-1}\right). \tag{2.72}$$

Preuve. De (2.67),

$$\begin{aligned}
\widehat{\beta} &= \left[\frac{1}{n}\sum_{t=1}^{n}\left\{\frac{\mathbf{Z}_{t-1}\mathbf{Z}'_{t-1}}{\sigma^4_{t-1}(\widehat{\beta}_{pr})}\right\}\right]^{-1}\left[\frac{1}{n}\sum_{t=1}^{n}\left\{\frac{\mathbf{Z}_{t-1}\mathbf{Y}_t}{\sigma^4_{t-1}(\widehat{\beta}_{pr})}\right\}\right] \\
&= \left[\frac{1}{n}\sum_{t=1}^{n}\left\{\frac{\mathbf{Z}_{t-1}\mathbf{Z}'_{t-1}}{\sigma^4_{t-1}(\widehat{\beta}_{pr})}\right\}\right]^{-1}\left[\frac{1}{n}\sum_{t=1}^{n}\left\{\frac{\mathbf{Z}_{t-1}(\mathbf{Z}'_{t-1}\beta + \sigma^2_{t-1}(\beta)u_t)}{\sigma^4_{t-1}(\widehat{\beta}_{pr})}\right\}\right] \\
&= \beta + \left[\frac{1}{n}\sum_{t=1}^{n}\left\{\frac{\mathbf{Z}_{t-1}\mathbf{Z}'_{t-1}}{\sigma^4_{t-1}(\widehat{\beta}_{pr})}\right\}\right]^{-1}\left[\frac{1}{n}\sum_{t-1}^{n}\left\{\frac{\mathbf{Z}_{t-1}\sigma^2_{t-1}(\beta)u_t}{\sigma^4_{t-1}(\widehat{\beta}_{pr})}\right\}\right]
\end{aligned}$$

$$\sqrt{n}(\widehat{\beta}-\beta)=\left[\frac{1}{n}\sum_{t=1}^{n}\left\{\frac{\mathbf{Z}_{t-1}\mathbf{Z}'_{t-1}}{\sigma^4_{t-1}(\widehat{\beta}_{pr})}\right\}\right]^{-1}\left[\frac{1}{\sqrt{n}}\sum_{t-1}^{n}\left\{\frac{\mathbf{Z}_{t-1}\sigma^2_{t-1}(\beta)u_t}{\sigma^4_{t-1}(\widehat{\beta}_{pr})}\right\}\right].$$

Montrons que si

$$\frac{1}{n}\sum_{t=1}^{n}\left\{\frac{1}{\sigma^4_{t-1}(\widehat{\beta}_{pr})}-\frac{1}{\sigma^4_{t-1}(\widehat{\beta})}\right\}\mathbf{Z}_{t-1}\mathbf{Z}'_{t-1}=o_p(1), \tag{2.73}$$

et

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\sigma^2_{t-1}(\beta)\left\{\frac{1}{\sigma^4_{t-1}(\widehat{\beta}_{pr})}-\frac{1}{\sigma^4_{t-1}(\beta)}\right\}\mathbf{Z}_{t-1}u_t=o_p(1), \tag{2.74}$$

on aura nécessairement

$$\sqrt{n}(\widehat{\beta}-\beta)=\left[\frac{1}{n}\sum_{t=1}^{n}\left\{\frac{\mathbf{Z}_{t-1}\mathbf{Z}'_{t-1}}{\sigma^4_{t-1}(\widehat{\beta})}\right\}\right]^{-1}\left[\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\frac{\mathbf{Z}_{t-1}\sigma^2_{t-1}(\beta)u_t}{\sigma^4_{t-1}(\widehat{\beta})}\right]+o_p(1).$$

Pour montrer (2.73) et (2.74), nous utilisons les égalités suivantes pour $u,v>0$
(i)

$$\frac{1}{u^2}-\frac{1}{v^2}=\frac{-2(u-v)}{\chi^3}, \tag{2.75}$$

où

$$0<1/\chi\leq(1/v)\{1+(v/u)\} \tag{2.76}$$

(ii)

$$\frac{1}{u^2}-\frac{1}{v^2}=\frac{-2(u-v)}{v^3}+\frac{3(u-v)^2}{\zeta^4}, \tag{2.77}$$

où

$$0<1/\zeta\leq(1/v)\{1+(v/u)\}, \tag{2.78}$$

(iii) Si $U=[u_1,\ldots,u_k]', V=[v_1,\ldots,v_k]'$, et W un vecteur dont toutes les composantes sont positives alors

$$\frac{W'V}{W'U}\leq 1+\frac{v_1}{u_1}+\ldots+\frac{v_k}{u_k}, \tag{2.79}$$

nous définissons $v_j/u_j=0$ si $u_j=0=v_j$.
En particulier, quand (2.76) et (2.78) sont utilisés avec $u=\widehat{\beta}'_{pr}\mathbf{Z}_{t-1}$ et $v=\beta'\mathbf{Z}_{t-1}$ alors par (2.79) on obtient :

$$\begin{aligned}\frac{1}{\chi_{t,n}}&\leq\frac{1}{\mathbf{Z}'_{t-1}\beta}\left[1+\frac{\mathbf{Z}'_{t-1}\beta}{\mathbf{Z}'_{t-1}\widehat{\beta}_{pr}}\right]\\&\leq\frac{1}{\mathbf{Z}'_{t-1}\beta}\left[1+\left\{1+\frac{\beta_0}{\widehat{\beta}_{0pr}}+\ldots+\frac{\beta_p}{\widehat{\beta}_{ppr}}\right\}\right],\end{aligned} \tag{2.80}$$

où $\widehat{\beta}_{jpr}$ est la jème entrée de $\widehat{\beta}_{pr}, 0\leq j\leq p$ et par l'hypothèse (2.70),

$$\left[1+\left\{1+\frac{\beta_0}{\widehat{\beta}_{0pr}}+\ldots+\frac{\beta_0}{\widehat{\beta}_{ppr}}\right\}\right]=O_p(1). \tag{2.81}$$

Soit $\mathbf{B}_n = [b_{n0}, \ldots, b_{np}]' = \sqrt{n}(\widehat{\beta}_{pr} - \beta) = O_p(1)$. En utilisant l'équation (2.75) et en remplaçant $\sigma^2_{t-1}(\beta)$ par (2.63) et $\mathbf{Z}_{t-1}$ par $[1, Y_{t-1}, \ldots, Y_{t-p}]'$ pour démontrer (2.73), on obtient :

$$\begin{aligned}
\frac{1}{n}\sum_{t=1}^{n}\left\{\frac{1}{\sigma^4_{t-1}(\widehat{\beta}_{pr})} - \frac{1}{\sigma^4_{t-1}(\beta)}\right\}\mathbf{Z}_{t-1}\mathbf{Z}^{'}_{t-1} &= \frac{1}{n}\sum_{t=1}^{n}\left\{\frac{1}{(\mathbf{Z}^{'}_{t-1}\widehat{\beta}_{pr})^2} - \frac{1}{(\mathbf{Z}^{'}_{t-1}\beta)^2}\right\}\mathbf{Z}_{t-1}\mathbf{Z}^{'}_{t-1} \\
&= \frac{1}{n}\sum_{t=1}^{n}\left\{\frac{-2(\mathbf{Z}^{'}_{t-1}\widehat{\beta}_{pr} - \mathbf{Z}^{'}_{t-1}\beta)}{\chi^3_{t,n}}\right\}\mathbf{Z}_{t-1}\mathbf{Z}^{'}_{t-1} \\
&= \frac{1}{n}\sum_{t=1}^{n}\left\{\frac{-2(\widehat{\beta}_{pr} - \beta)'\mathbf{Z}_{t-1}}{\chi^3_{t,n}}\right\}\mathbf{Z}_{t-1}\mathbf{Z}^{'}_{t-1} \\
&= \frac{-2}{n^{3/2}}\sqrt{n}(\widehat{\beta}_{pr} - \beta)'\sum_{t=1}^{n}\left\{\frac{\mathbf{Z}_{t-1}\mathbf{Z}_{t-1}\mathbf{Z}^{'}_{t-1}}{\chi^3_{t,n}}\right\} \\
&= \frac{-2}{n^{3/2}}\mathbf{B}^{'}_n\sum_{t=1}^{n}\left\{\frac{\mathbf{Z}_{t-1}\mathbf{Z}_{t-1}\mathbf{Z}^{'}_{t-1}}{\chi^3_{t,n}}\right\} \\
&= \frac{-2}{n^{3/2}}b_{n0}\sum_{t=1}^{n}\left\{\frac{\mathbf{Z}_{t-1}\mathbf{Z}^{'}_{t-1}}{\chi^3_{t,n}}\right\} \\
&= \frac{-2}{n^{3/2}}\sum_{j=1}^{p}b_{nj}\sum_{t=1}^{n}\left\{\frac{\mathbf{Y}_{t-j}\mathbf{Z}_{t-1}\mathbf{Z}^{'}_{t-1}}{\chi^3_{t,n}}\right\} \\
&= -2\mathbf{T}_1 - 2\mathbf{T}_2.
\end{aligned}$$

De (2.79)

$$\frac{1}{\chi^3_{t,n}} \leq \beta_0^{-3}\left[1 + \left\{1 + \frac{\beta_0}{\widehat{\beta}_{0pr}} + \ldots + \frac{\beta_0}{\widehat{\beta}_{ppr}}\right\}\right]^3,$$

Comme l'unique solution stationnaire non anticipative de (ε_t) est ergodique, et $\forall\, t$, $\mathbf{Z}_{t-1}$ s'ecrit comme fonction mesurable des ε_{t-i}, alors le processus $(\mathbf{Z}_{t-1})$ est également stationnaire et ergodique, alors $\mathbf{T}_1 = o_p(1)$.
Pour $\mathbf{T}_2$,

$$\begin{aligned}
\sum_{t=1}^{n}\frac{\mathbf{Y}_{t-j}\mathbf{Y}_{t-k}\mathbf{Y}_{t-l}}{\chi^3_{t,n}} &\leq \sum_{t=1}^{n}\left\{\frac{\mathbf{Y}_{t-j}\mathbf{Y}_{t-k}\mathbf{Y}_{t-l}}{(\mathbf{Z}^{'}_{t-1}\beta)^3}\right\}\left[1 + \left\{\frac{\mathbf{Z}^{'}_{t-1}\beta}{\mathbf{Z}^{'}_{t-1}\widehat{\beta}_{pr}}\right\}\right]^3 \\
&\leq \left[1 + \left\{1 + \frac{\beta_0}{\widehat{\beta}_{0pr}} + \ldots + \frac{\beta_0}{\widehat{\beta}_{ppr}}\right\}\right]^3\sum_{t=1}^{n}\left\{\frac{\mathbf{Y}_{t-j}\mathbf{Y}_{t-k}\mathbf{Y}_{t-l}}{(\mathbf{Z}^{'}_{t-1}\beta)^3}\right\}.
\end{aligned}$$

Puisque

$$n^{-3/2}\mathbb{E}\left\{\sum_{t=1}^{n}\frac{\mathbf{Y}_{t-j}\mathbf{Y}_{t-k}\mathbf{Y}_{t-l}}{(\mathbf{Z}^{'}_{t-1}\beta)^3}\right\} = o_p(1),$$

on obtient $\mathbf{T}_2 = o_p(1)$ Pour démontrer (2.74), on utilise l'équation (2.77) et en remplaçant $\sigma^2_{t-1}(\beta)$ par (2.63), on aura :

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\sigma_{t-1}^{2}(\beta)\left\{\frac{1}{\sigma_{t-1}^{4}(\widehat{\beta}_{pr})}-\frac{1}{\sigma_{t-1}^{4}(\beta)}\right\}\mathbf{Z}_{t-1}u_t = \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\sigma_{t-1}^{2}(\beta)\left\{\frac{1}{(\mathbf{Z}_{t-1}'\widehat{\beta}_{pr})^2}-\frac{1}{\mathbf{Z}_{t-1}'\beta)^2}\right\}\mathbf{Z}_{t-1}u_t$$
$$= \frac{1}{\sqrt{n}}\sum_{t=1}^{2}\sigma_{t-1}^{n}(\beta)\frac{-2(\mathbf{Z}_{t-1}'\widehat{\beta}_{pr}-\mathbf{Z}_{t-1}'\beta)}{(\mathbf{Z}_{t-1}'\beta)^3}\mathbf{Z}_{t-1}u_t$$
$$+\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\sigma_{t-1}^{2}(\beta)\frac{3(\mathbf{Z}_{t-1}'\widehat{\beta}_{pr}-\mathbf{Z}_{t-1}'\beta)^2}{\zeta_{t,n}^4}\mathbf{Z}_{t-1}u_t$$
$$= \frac{-2}{n}\sum_{t=1}^{n}\sigma_{t-1}^{2}(\beta)\sqrt{n}(\widehat{\beta}_{pr}-\beta)'\frac{\mathbf{Z}_{t-1}\mathbf{Z}_{t-1}u_t}{(\mathbf{Z}_{t-1}'\beta)^3}$$
$$+\frac{3}{n^{3/2}}\sum_{t=1}^{n}\sigma_{t-1}^{2}(\beta)\{\sqrt{n}(\widehat{\beta}_{pr}-\beta)'\mathbf{Z}_{t-1}\}^2\frac{\mathbf{Z}_{t-1}u_t}{\zeta_{t,n}^4}$$
$$= -2\mathbf{T}_3+3\mathbf{T}_4.$$

Comme $\mathbb{E}(u_t)=0$, on utilise des techniques similaires à celle utilisée pour démontrer $\mathbf{T}_1 = o_p(1)$ et $\mathbf{T}_2 = o_p(1)$, on trouve $\mathbf{T}_3 = o_p(1)$.
On pose $\mathbf{T}_4 = \mathbf{T}_{41} + \mathbf{T}_{42}$, où

$$\mathbf{T}_{41} = b_{n0}^2 n^{-3/2}\sum_{t=1}^{n}\frac{\mathbf{Z}_{t-1}\mathbf{Z}_{t-1}'\beta}{\zeta_{t,n}^4}$$
$$= o_p(1),$$

et

$$\mathbf{T}_{42} = \sum_{j=1}^{p}\sum_{k=1}^{p} b_{nj}b_{nk}n^{-3/2}\sum_{t=1}^{n}\frac{\mathbf{Y}_{t-j}\mathbf{Y}_{t-k}\mathbf{Z}_{t-1}u_t\mathbf{Z}_{t-1}'\beta}{\zeta_{t,n}^4}$$
$$\leq O_p(1)\times n^{-3/2}\sum_{t=1}^{n}\frac{\mathbf{Y}_{t-j}\mathbf{Y}_{t-k}\mathbf{Y}_{t-l}|u_t|}{(\mathbf{Z}_{t-1}'\beta)^3}.$$

De (2.71), on conclut

$$n^{-3/2}\sum_{t=1}^{n}\mathbb{E}\left\{\frac{\mathbf{Y}_{t-j}\mathbf{Y}_{t-k}\mathbf{Y}_{t-l}|u_t|}{(\mathbf{Z}_{t-1}'\beta)^3}\right\} = \mathbb{E}(|u_t|)n^{-1/2}\mathbb{E}\left\{\frac{\mathbf{Y}_{t-j}\mathbf{Y}_{t-k}\mathbf{Y}_{t-l}}{(\mathbf{Z}_{t-1}'\beta)^3}\right\}$$
$$= o_p(1),$$

finnalement $\mathbf{T}_{42} = o_p(1)$. □

2.4 Conclusion

Dans ce chapitre, nous avons présenté en trois parties les diverses méthodes d'estimation des modèles GARCH en mettent l'accent sur les propriétés asymptotiques de chaque estimateur.
Pour la méthode des moindres carrés et celle de Mukherjee (2003) l'existence des moments d'ordre huit, pour la normalité asymptotique est exigé. Cette hypothèse, très forte d'ailleurs, induit une restriction sur l'espace des paramètres, et donc nuit à la modélisation des processus à queue épaisse, auxquels sont préconisé les modèles ARCH. Cependant, leurs expression sont explicites et simple à obtenir.
Les estimateurs de maximum de vraisemblance des paramètres du modèle GARCH sont convergents et asymptotiquement normaux. La précision de ces estimateurs s'exprime en fonction da la matrice J. Il est important de noter que lorsque la vraie densité conditionnelle est effectivement normale, les estimateurs de la moyenne et ceux de la variance (conditionnelles) sont asymptotiquement non corrélés : ils peuvent ainsi être estimés séparément sans perte d'efficacité. Par ailleurs, l'estimateur du M.V est plus précis que celui des M.C.O, en plus la variance asymptotique de cet estimateur coincide avec celui de l'estimateur des deux phases, mais la normalité asymptotique est obtenue sans aucune hypothèse sur les solides, cela explique pourquoi la méthode de M.V est la meilleure.

Chapitre 3

Estimation M.V dans les processus GARCH lorsque certains coefficients sont égaux à zéro

3.1 Introduction

Dans ce chapitre nous étudions les conditions pour lesquelles les estimateurs des paramètres du modèles GARCH soient consistents et asymptotiquements gaussiens lorsque certains coefficients sont égaux à zéro. Ces résultats sont obtenues sous certains conditions "moyennes".
Ces résultats pouvent être exploités et utilisés dans la pratique pour des techniques d'identifications.
On a déjà vu dans le chapitre précédent que l'estimateur M.V de θ se définit comme solution mesurable $\widehat{\theta}_n$ de :

$$\widehat{\theta}_n = \arg\max_{\theta\in\Theta} L_n(\theta) = \arg\min_{\theta\in\Theta} \widetilde{I}_n(\theta), \tag{3.1}$$

où $\widetilde{I}_n(\theta)$ donnée dans (2.15) et $L_n(\theta)$ est la vraisemblance conditionnelle gaussienne donnée dans (2.10).

Sous la condition $\gamma(A_0) < 0$ et $\sum_{j=1}^{q} \beta_j < 1$ nous désignons $\sigma_t^2 = \{\sigma_t^2(\theta)\}$ la solution strictement stationnaire, ergodique et non anticipative de :

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1} \beta_j \sigma_{t-j}^2, \quad \forall\ t \in \mathbb{Z}. \tag{3.2}$$

Soit

$$I_n(\theta) = n^{-1} \sum_{t=1}^{n} \ell_t, \quad et \quad \ell_t = \ell_t(\theta) = \ell_t(\theta, \varepsilon_t, \dots) = \frac{\varepsilon_t^2}{\sigma_t^2} + \log \sigma_t^2.$$

et $\mathcal{A}_\theta(z) = \sum_{i=1}^{q} \alpha_i z^i$ et $\mathcal{B}_\theta(z) = 1 - \sum_{j=1}^{p} \beta_j z^j$. Par convention, $\mathcal{A}_\theta(z) = 0$ si $q = 0$ et $\mathcal{B}_\theta(z) = 1$ si $p = 0$.
Pour obtenir les propriétés asymptotiques du l' E.M.V, il y'a lieu d'imposer les hypothèses suivantes :
H1 :$\theta_0 \in \Theta$, et Θ est compacte.

H2 :$\gamma(A_0) < 0$ et $\forall \theta \in \Theta,\ \sum_{j=1}^{q} \beta_j < 1$.
H3 :η_t^2 a une loi non dégénérée et $\mathbb{E}(\eta_t^2) = 1$.
H4 :Si $p > 0$, $\mathcal{A}_0(z)$ et $\mathcal{B}_0(z)$ n'ont pas de racine commune, $\mathbf{A}_0(1) \neq 0$, et $\alpha_{0q} + \beta_{0p} \neq 0$.
H5 :$\kappa_\eta := \mathbb{E}[\eta_t^4] < \infty$.
Il est à noter que, dans **H2**, la condition de stationnarité stricte est imposée uniquement au modèle à la vraie valeur θ_0. Pour $\theta \neq \theta_0$ nous avons seulement besoin de la condition de la faible stationnarité $\sum_{j=1}^{p} \beta_j < 1$. Une conséquence importante de $\gamma(A_0) < 0$ est que $\mathbb{E}[\varepsilon_t^{2s}] < \infty$ pour un certain $s \in (0, 1)$.

Théorème 3.1.1. *Soit $(\widehat{\theta}_n)$ une suite d'estimateur de M.V satisfaisant (3.3). On a :*
- *(i) Si **H2-H4** sont vérifiées alors, $\widehat{\theta}_n \longrightarrow \theta_0$ lorsque $n \to \infty$ presque sûrement.*
- *(ii) Si **H1-H5** sont vérifiées, $\sqrt{n}(\widehat{\theta}_n - \theta_0) \xrightarrow{\mathcal{L}} \mathcal{N}(0, (\kappa_\eta - 1)J^{-1}\eta)$, où*

$$J := \mathbb{E}_{\theta_0}\left(\frac{1}{\sigma_t^4(\theta_0)} \frac{\partial \sigma_t^2(\theta_0)}{\partial \theta} \frac{\partial \sigma_t^2(\theta_0)}{\partial \theta'}\right). \tag{3.3}$$

La question à laquelle nous nous interéssons dans cette partie est : Quelles sont les conditions à imposer lorsque certains coefficients GARCH sont nuls pour avoir la consistence et la normalité asymptotique.

Dans ce qui suit, nous supposons que les vraies valeurs des paramètres appartienent à $\Theta := \{\theta_0 \in \Theta\ : \theta_{0i} = 0,$ pour certains $i > 0\}$. Pour mieux caractériser Θ nous définissons $\theta_0(\varepsilon)$ comme vecteur obtenu en remplaçant tous les coefficients nuls de θ_0 par $\theta_0(\varepsilon)$ et nous faisons l'hypothèse suivante :
H6 :$\theta_0(\epsilon) \in \overset{\circ}{\Theta}$ pour $\epsilon > 0$.
Par exemple, si l'espace des paramètres est spécifié comme $\Theta = [\underline{\omega}, \overline{\omega}] \times [0, \overline{\alpha}_1] \times \cdots \times [0, \overline{\alpha}_q] \times [0, \overline{\beta}_1] \times \cdots \times [0, \overline{\beta}_p]$, l'hypothèse H6 est satisfaite lorsque $\overline{\omega} > \omega_0 > \underline{\omega} > 0$ et $0 \leq \theta_0 < \theta := (\overline{\omega}, \overline{\alpha}_1, \ldots, \overline{\beta}_p)'$.

3.2 Distribution asymptotique de $\widehat{\theta}_n$ lorsque certains coefficients de θ_n sont nuls

La distribution asymptotique gaussienne pour $\sqrt{n}(\widehat{\theta}_n - \theta_0)$ est non vérifiée lorsque les composents $\widehat{\theta}_{in}$ de $\widehat{\theta}_n$ sont obligés d'être non négative et certains coefficient sont nuls.
Si par exemple , $\theta_{0i} = 0$ puis

$$\sqrt{n}(\widehat{\theta}_{in} - \theta_{0i}) = \sqrt{n}\widehat{\theta}_{in} \geq 0 \quad \forall n,$$

et la distribution asymptotique de cette variable n'est pas gaussienne.

3.2.1 Non existence possible de J sous H2-H5

En éliminant **H1**, considérons le modèle ARCH(2)

$$\varepsilon_t = \sigma_t \eta_t$$
$$\sigma_t^2 = \omega_0 + \alpha_{01}\varepsilon_{t-1}^2 + \alpha_{02}\varepsilon_{t-2}^2,$$

où $\omega_0 > 0$, $\alpha_{01} \geq 0$, $\alpha_{02} = 0$, et la distribution de la suite (η_t) *iid* est défini, pour $a > 1$, par :

$$P(\eta_t = a) = P(\eta_t = -a) = \frac{1}{2a^2}, \quad P(\eta_t = 0) = 1 - \frac{1}{a^2}.$$

Ce modèle ARCH(2) est utilisé pour générer la fonction de quasi-vraisemblance mais ε_t est en fait un ARCH(1). La condition de stationnarité stricte $\gamma(A_0) < 0$ prend la forme $\alpha_{01} < \exp[-\mathbb{E}(\log \eta_t^2)]$ pour un ARCH(1). Le processus $\{\varepsilon_t;\ t \in \mathbb{Z}\}$ est donc strictement stationnaire pour toute valeur de α_{01} puisque $\exp[-\mathbb{E}(\log \eta_t^2)] = +\infty$. Toutefois $\{\varepsilon_t;\ t \in \mathbb{Z}\}$ n'est pas de deuxième ordre lorsque $\alpha_{01} \geq 1$.
On a :

$$\frac{1}{\sigma_t^2}\frac{\partial \sigma_t}{\partial \alpha_2}(\theta_0) = \frac{\varepsilon_{t-2}^2}{\omega_0 + \alpha_{01}\varepsilon_{t-1}^2},$$

d'où

$$\mathbb{E}_{\theta_0}\left\{\frac{1}{\sigma_t^2}\frac{\partial \sigma_t^2}{\partial \alpha_2}(\theta_0)\right\}^2 \geq \mathbb{E}_{\theta_0}\left[\left\{\frac{\varepsilon_{t-2}^2}{\omega_0 + \alpha_{01}\varepsilon_{t-1}^2}\right\}^2 \middle| \eta_{t-1} = 0\right] P(\eta_{t-1} = 0)$$
$$= \frac{1}{\omega_0^2}\left(1 - \frac{1}{a^2}\right)\mathbb{E}_{\theta_0}(\varepsilon_{t-2}^4)$$

d'abord parce que $\eta_{t-1} = 0$ entraîne $\varepsilon_{t-1} = 0$ et d'autre part parce que η_{t-1} et ε_{t-2} sont indépendants. Alors, J n'existe pas si $\mathbb{E}_{\theta_0}(\varepsilon_t^4) = \infty$.

3.2.2 Résultat principal et la normalité asymptotique

Il est donc clair que les hypothèses **H2-H5** ne suffisent pas pour garantir l'existence de J.
Nous introduisons deux hypothèses alternatives. La première est une condition sur les moments.
H7 :$\mathbb{E}_{\theta_0}\varepsilon_t^6 < \infty$.
Dans de nombreux cas intéressants, sauf les modèles ARCH(q), aucune hypothèse de moment exigé sur ε_t^2. En effet, il suffira d'assurer l'existence de moments pour σ_t^2.
Notez que sous la condition $\gamma(A_0) < 0$, la solution strictement stationnaire $\sigma_t^2(\theta_0)$ a une extension de la forme :

$$\sigma_t^2(\theta_0) = c_0 + \sum_{j=1}^{\infty} b_{0j}\varepsilon_{t-j}^2,$$

avec $c_0 > 0$, $b_{0j} \geq 0$. Des extensions similaires sont valables pour les dérivés par rapport à θ. Le contrôle des moments décomptera sur le fait que chaque terme apparaissant dans le numérateur de ce rapport est également présent dans le dénominateur. Nous considérons donc l'hypothèse suivante :
H8 :$b_{0j} > 0$ pour tout $j \geq 1$, où $\sigma_t^2(\theta_0) = c_0 + \sum_{j=1}^{\infty} b_{0j}\varepsilon_{t-j}^2$.
Il convient de noter qu'une simple condition suffisante pour **H8** est $\alpha_{01} > 0$ et $\beta_{01} > 0$

(parce que $b_{0j} \geq \alpha_{01}\beta_{01}^{j-1}$). Une condition nécessaire est évidemment que $\alpha_{01} > 0$ (parce que $b_{01} = \alpha_{01}$). Plus généralement, une condition nécessaire et suffisante pour que **H8** soit vérifié

$$\{j \mid \beta_{0,j} > 0\} \neq \varnothing \quad et \quad \prod_{i=1}^{j_0} \alpha_{01} > 0 \quad pour \quad j_0 = \min\{j|\beta_{0,j} > 0\}, \tag{3.4}$$

ce qui signifie que les coefficients ARCH α ne peut pas annuler jusqu'à la commande j_0 du premier coefficient de GARCH β égal à zéro. L'hypothèse **H8** ne s'applique pas à ARCH(q).

Théorème 3.2.1. ***[C. Francq et J. M. Zakoïan (2007)]*** *Soit $(\widehat{\theta}_n)$ une suite d'estimateurs M.V de modèle GARCH satisfaisant (3.1) . Ensuite, si* ***H2*** *-* ***H6*** *et* ***H7*** *ou* ***H8*** *sont verifiées, alors*

$$\sqrt{n}(\widehat{\theta}_n - \theta_0) \xrightarrow{\mathcal{L}} \lambda^\Lambda := \arg\inf_{\lambda \in \Lambda} \{\lambda - Z\}' J \{\lambda - Z\},$$

avec $Z \sim \mathcal{N}(0, (\kappa_\eta - 1)J^{-1})$, $\Lambda = \Lambda(\theta_0) = \Lambda_1 \times \Lambda_{p+q+1}$, *où* $\Lambda_1 = \mathbb{R}$ *et pour* $i = 2, \ldots, p+q+1$, $\Lambda_i = \mathbb{R}$ *si* $\theta_{0i} \neq 0$ *et* $\Lambda_i = [0, \infty)$ *si* $\theta_{0i} = 0$.

Dans le cas ARCH le résultat peut être énoncé comme suit :

Corollaire 3.2.1. ***[C. Francq et J. M. Zakoïan (2007)]*** *Soit $p = 0$ et $(\widehat{\theta}_n)$ une suite d'estimateurs M.V satisfaisant (3.3) . Si $\gamma(A_0) < 0$, sous les hypothèses* ***H3*** *et* ***H5-H7****, alor*

$$\sqrt{n}(\widehat{\theta}_n - \theta_0) \xrightarrow{\mathcal{L}} \lambda^\Lambda := \arg\inf_{\lambda \in \Lambda} \{\lambda - Z\}' J \{\lambda - Z\},$$

avec $Z \sim \mathcal{N}(0, (\kappa_\eta - 1)J^{-1})$, $\Lambda = \Lambda(\theta_0) = \Lambda_1 \times \Lambda_{q+1}$, *où* $\Lambda_1 = \mathbb{R}$ *et pour* $i = 2, \ldots, q+1$, $\Lambda_i = \mathbb{R}$ *si* $\theta_{0i} \neq 0$ *et* $\Lambda_i = [0, \infty)$ *si* $\theta_{0i} = 0$.

Remarque 3.2.1. *Pour $\theta_0 \in \overset{o}{\Theta}$, le résultat de ce théorème se réduit à celui du théorème (3.1.1) . En effet, dans ce cas $\Lambda = \mathbb{R}^{p+q+1}$ et $\lambda^\Lambda = Z \sim \mathcal{N}(0, (\kappa_n - 1)J^{-1})$. Par conséquent, le théorème (3.2.1) n'a d'intérêt que lorsque certains coefficients sont nuls.*

λ^Λ est la projection orthogonale de Z sur Λ, où l'orthogonalité est définie dans la métrique associée à la structure de covariance J, à savoir $x \perp y$ si $x'Jy = 0$. Il est uniquement déterminé parce que Λ est convexe. De plus, le fait que Λ est un cône convexe dont les faces sont des sous-espaces permet d'obtenir cette projection de manière plus explicite. Supposons, sans perte de généralité, que les premiers composants d_1 de θ_0 sont positifs, et que les derniers composants d_2 sont nuls, avec $d_1 + d_2 = p+q+1$. On a $\Lambda = \mathbb{R}^{d_1} \times [0, \infty)^{d2} = \{\lambda \in \mathbb{R}^{d_1+d_2} | K\lambda \geq 0\}$, où $K = (0_{d_2 \times d_1}, I_{d_2})$. Soit $\mathcal{K} = \{K_1, \ldots, K_{2^{d_2}-1}\}$, où K_i sont des matrices obtenues en annulant 0, 1 ou plusieurs (jusqu'à $d_2 + 1$) lignes de K. Soit $M_i = K_i'(K_i J^{-1} K_i')^{-1} K_i$, soit $P_i = I_{d_1+d_2} - J^{-1}M_i$ et notons par $\lambda_{K_i} = P_i Z$ la projection de Z sur le sous-espace linéaire de $\mathbb{R}^{d_1+d_2}$ passé par l'un des visages $2^{d_2} - 1$ de Λ (y compris le "visage" $\mathbb{R}^{d_1} \times \{0\}^{d_2}$), Défini par $K_i\lambda = 0$. Ensuite nous avons, avec $\mathcal{C} = \{\lambda_{K_i} = K_i \in \mathcal{K} \quad et \quad K\lambda_{k_i} \geq 0\}$,

$$\lambda^\Lambda = Z 1_\Lambda(Z) + 1_{\Lambda^c}(Z) \times \arg\min_{\lambda \in \mathcal{C}} \|\lambda - Z\|_J$$

$$= Z1_\Lambda(Z) + \sum_{i=1}^{2^{d_2-1}} P_i Z 1_{\mathcal{D}_i}(Z),$$

pour une partition $(\mathcal{D}_i)$ de $\mathbb{R}^d - \Lambda$. Pour illustrer cette proposition, considérons l'exemple suivant.

Exemples 2. ***[Un coefficient zéro]*** *Supposons qu'un seul composant de θ_0 est zéro, les autres composants sont positifs. Ainsi $d_2 = 1$, $\Lambda = \mathbb{R}^{d_1} \times [0, \infty)$, $K = (0, \dots, 0, 1)$, $\mathcal{K} = \{K\}$ et soit $Z = (Z_1, \dots Z_d)'$,*

$$\lambda^\Lambda = Z1_{Z_d<0} + PZ1_{Z_d<0}, \qquad P = I_d - J^{-1}K'(KJ^{-1}K')^{-1}K.$$

Il s'ensuit que

$$\lambda^\Lambda = Z - Z_d^- c \tag{3.5}$$

où $Z_d^- = Z_d 1_{Z_d<0}$, et $c = \mathbb{E}(Z_d Z)/var(Z_d)$ est la dernière colonne de J^{-1} divisé par le (d, d) -élément de cette matrice. Notez que le dernier composant de λ^Λ est $Z_d^+ := Z_d 1_{Z_d>0}$. soit $\lambda^\Lambda = (\lambda_1^\Lambda, \dots, \lambda_d^\Lambda)'$, on voit aussi que $\lambda_i^\Lambda = Z_i$ si et seulement si $Cov(Z_i, Z_d) = 0$.

Avant de prouver le théorème (3.2.1), nous donnons plusieurs lemmes. Notez que lorsque certains coefficients sont nuls, la fonction $\sigma_t^2(\theta)$ peut être négative dans un voisinage de θ_0 et $\ell_t(\theta)$ peut être non défini dans ce voisinage. Au lieu d'un développement de Taylor, nous utiliserons un développement unilatéral, basé sur des dérivées à droits de $I_n(\theta) = n^{-1}\sum_{t=1}^n \ell_t(\theta)$ sur θ_0. Nous notons par :

$$\frac{\partial \sigma_t^2(\theta_0)}{\partial \theta} := \left(\frac{\partial \sigma_t^2(\theta_0)}{\partial \theta_i}\right)_{i=1,\dots,p+q+1},$$

et

$$\frac{\partial \ell_t(\theta_0)}{\partial \theta} := \left(\frac{\partial \ell_t(\theta_0)}{\partial \theta_i}\right)_{i=1,\dots,p+q+1}.$$

Les vecteurs de dérivées partielles de σ_t et ℓ_t à θ_0 avec le i-ième dérivé remplacé par la dérivée droit lorsque $\theta_{0i} = 0$. Nous utilisons la même convention pour les dérivés de I_n, $\widetilde{\ell}_t$ et $\widetilde{I}_n$ à θ_0 et pour les deuxièmes dérivées partielles. Par cette convention, les dérivés de $\ell_t(\theta) = \varepsilon_t^2/\sigma_t^2 + \log \sigma_t^2$ sont données par (2.37) et (2.38),
nous avons :

$$\sigma_t^2 = \sum_{k=0}^{\infty} B^k(1,1)\left(\omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-k-i}^2\right), \tag{3.6}$$

$$\frac{\partial \sigma_t^2}{\partial \omega} = \sum_{k=0}^{\infty} B^k(1,1), \qquad \frac{\partial \sigma_t^2}{\partial \alpha_i} = \sum_{k=0}^{\infty} B^k(1,1)\varepsilon_{t-k-i}^2, \tag{3.7}$$

$$\frac{\partial \sigma_t^2}{\partial \beta_j} = \sum_{k=1}^{\infty} B_{k,j}(1,1)\left(\omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-k-i}^2\right), \tag{3.8}$$

où

$$B_{k,j} = \frac{\partial B^k}{\partial \beta_j} = \sum_{m=1}^{k} B^{m-1}B^{(j)}B^{k-m}, \quad B = \begin{pmatrix} \beta_1 & \beta_2 & \dots & \beta_p \\ 1 & 0 & \dots & 0 \\ \vdots & & & \\ 0 & \dots & 1 & 0 \end{pmatrix}, \tag{3.9}$$

et B^j est une matrice $(p \times p)$ avec $(1, j)$ est l'élément 1 et tous les autres éléments sont nuls. Propriétés élémentaires de la matrice B sont établis dans le lemme suivant.

Lemme 3.2.1. ***[C. Francq et J. M. Zakoïan (2007)]*** *pour* $j = 1, \dots, p$

$$B^{(j)}B^k \leq B^{k-j+1}, \quad \textit{pour tout} \quad k \geq j-1, \tag{3.10}$$

$$B^{(j)}B^k = B^{j-k}, \quad \textit{pour tout} \quad 0 \leq k < j, \tag{3.11}$$

$$AB^{(1)} \leq A, \quad \textit{et} \quad \textit{pour} \quad j \neq \ell_2, \quad \{AB^{(j)}\}(\ell_1, \ell_2) = 0, \quad \forall\ A \geq 0, \tag{3.12}$$

$$AB^{(1)}B^{(j)} = B^{(j)}, \quad \textit{et} \quad B^{(i)}B^{(j)} = 0, \quad \textit{pour} \quad i > 1, \tag{3.13}$$

$$B^k(1,1) \geq \beta_j B^{k-j}(1,1), \quad \textit{pour} \quad \textit{tout} \quad k \geq j. \tag{3.14}$$

Preuve. Nous notons que, lorsque $j \leq p$, le j-ième ligne de $B^{(j)}$ est la première ligne de B. La première ligne de $B^{(j)}A$ est le j-ième ligne de A et les autres éléments de $B^{(j)}A$ sont des zéros. Ainsi $B^{(j)}B^j \leq B$. Multiplions les deux côtés de l'inégalité précédente par B^{k-j}, dont les éléments sont non négatifs, donne (3.12). Pour $k < j \leq p$, le j-ième ligne de B^k est nul, sauf un "1" dans le j-ième position, qui montre (3.13). Le j-ième colonne de $AB^{(j)}$ est la première colonne de A et les autres éléments de $AB^{(j)}$ sont des zéros. Ainsi (3.14), (3.15) sont évidents. L'inégalité (3.16)vient de $B^k(1,1) = \sum_{j=1}^{p} \beta_j B^{k-1}(j,1) \geq \beta_j B^{k-1}(j,1) = \beta_j B^{k-j}(1,1)$. $\square$

Lemme 3.2.2. ***[C. Francq et J. M. Zakoïan (2007)]*** *Sous les hypothèses du théorème(3.2.1), nous avons :*

$$\mathbb{E}_{\theta_0}\left\|\frac{1}{\sigma_t^2}\frac{\partial \sigma_t^2}{\partial \theta}(\theta_0)\right\| < \infty, \qquad \mathbb{E}_{\theta_0}\left\|\frac{1}{\sigma_t^4}\frac{\partial \sigma_t^2}{\partial \theta}\frac{\partial \sigma_t^2}{\partial \theta'}(\theta_0)\right\| < \infty,$$

$$\mathbb{E}_{\theta_0}\left\|\frac{1}{\sigma_t^2}\frac{\partial^2 \sigma_t^2}{\partial \theta \partial \theta'}\right\| < \infty.$$

Preuve. Soient K et ρ des constantes, dont les valeurs peuvent changer de ligne à ligne mais toujours satisfaire $K > 0$ et $0 < \rho < 1$. Puisque σ_t^{-2} est borné par $1/\underline{\omega}$, la preuve du ce lemme est simple sous **H7**.
Supposons maintenant que **H8**, au lieu de **H7**. Par (3.9), $\partial \sigma_t^2(\theta_0)/\partial \theta$ est borné, $\sum_{k=0}^{\infty} B^k$ est fini sous **H2**. Puisque $\sigma_t^2(\theta_0 \geq \omega_0 > 0$, $\{\partial \sigma_t^2(\theta_0)/\partial \omega\}/\sigma_t^2(\theta_0)$ possède donc des moments de tout ordre. Considérons les dérivés en ce qui concerne α_i. Soit B_0 être la matrice B pour $\theta = \theta_0$. Nous avons, d'après (3.8)

$$\sigma_t^2(\theta_0) = \omega_0 \sum_{k=0}^{\infty} B_0^k(1,1) + \sum_{k=1}^{\infty}\sum_{\ell=1}^{k} \alpha_{0\ell} B_0^{k-\ell}(1,1)\varepsilon_{t-k}^2,$$

par convention $\alpha_{0\ell} = 0$ quand ℓ n'appartient pas à $\{1, \dots, q\}$. Par **H8**, pour tout $k > 0$ il existe un entier $i_k \in \{1, \dots, \min(q,k)\}$ tel que

$$\sum_{\ell=1}^{k} \alpha_{0\ell} B_0^{k-\ell}(1,1) \geq \alpha_{0i_k} B_0^{k-i_k}(1,1) \geq \underline{\alpha} B_0^{k-i_k}(1,1) > 0, \tag{3.15}$$

pour une constante positive $\underline{\alpha}$ (on peut prendre $\underline{\alpha} = \min\{\alpha_{0i} \ : \alpha_{0i} \neq 0\}$). Pour tout $s \in [0,1]$, on aura (3.8) , (3.9),

$$\frac{1}{\sigma_t^2}\frac{\partial \sigma_t^2}{\partial \alpha_i}(\theta_0) = \frac{\sum_{k=0}^{\infty} B_0^k(1,1)\varepsilon_{t-i-k}^2}{\sum_{k=0}^{\infty}(1,1)\left(\omega_0 + \sum_{j=1}^{q}\alpha_{0j}\varepsilon_{t-j-k}^2\right)}$$

$$\leq \sum_{k=i}^{\infty} \frac{B_0^{k-i}(1,1)\varepsilon_{t-k}^2}{\omega_0 + \underline{\alpha} B_0^{k-i_k}(1,1)\varepsilon_{t-k}^2} \leq \sum_{k=i}^{\infty} \omega_0^s \underline{\alpha}^{1-s}\{B_0^{k-i_k}(1,1)\}^{1-s}, \tag{3.16}$$

où la dernière inégalité découle de $ax/(b+cx) \leq ax^s/(b^s c^{1-s})$ pour tous $a, x \geq 0$, $b, c > 0$ et $s \in [0,1]$. Cette dernière inégalité provient de l'inégalité élémentaire $x/(1+x) \leq x^s$ pour tous $x \geq 0$ et tout $s \in [0,1]$. pour tout fixe $s \in [0,1]$, nous montrons pour tout $s \in [0,1]$ que

$$B_0^{k-i}(1,1)/\{B_0^{k-i_k}(1,1)\}^{1-s} \leq K\rho^k \qquad \textit{pour tout} \quad k. \tag{3.17}$$

Par **H2** et la compacité de Θ, on a $\sup_{\theta\in\Theta}\rho(B) < 1$. Ainsi $\|B_0^k\| \leq K\rho^k$ pour tous k, et i_k appartient à l'ensemble fini $\{1\ldots,q\}$, on a $\left\{B_0^{k-i_k}(1,1)\right\}^s \leq K\rho^k$, et il suffit de montrer que $B_0^{k-i}(1,1)/B_0 k - i_k(1,1)$ est borné par une constante indépendante de k. Il suffit de considérer k tel que $B_0^{k-i}(1,1) \neq 0$. Soit j_0 être défini par (3.6) et soit $r_i \in \{1,\ldots,j_0\}$ être tel que $i-1 \equiv r_i - 1(mod\ j_0)$, C'est $i = q_i j_0 + r_i$ avec $q_i \geq 0$. D'après (3.16) , nous avons $B_0^{k-r_i}(1,1) = B_0^{k-i+q_i j_0}(1,1) \geq \beta_{0j_0}^{q_i} B_0^{k-i}(1,1) > 0$. En outre, $\alpha_{0r_i} \neq 0$ par (3.6) . Ainsi on peut prendre $i_k = r_i$ dans (3.17) , de sorte que nous avons

$$B_0^{k-i}(1,1)/B_0^{k-i_k}(1,1) \leq 1/\beta_{0j_0}^{q_i}, \tag{3.18}$$

et ainsi (3.19) est valide . Puis (3.18) donne

$$\mathbb{E}_{\theta_0}\frac{1}{\sigma_t^2}\frac{\partial \sigma_t^2}{\partial \alpha_i}(\theta_0) \leq K\left\{\sum_{k=1}^{\infty}\rho^k\right\}\mathbb{E}_{\theta_0}\varepsilon_t^{2s}. \tag{3.19}$$

Puisque ε_t^2 a un moment d'ordre s, pour certains $s \in [0,1]$, le membre de droite dans la dernière inégalité est fini. Par conséquent $\sigma_t^{-2}(\sigma_t^2/\partial\alpha_i)$ a un moment d'ordre 1 à $\theta = \theta_0$. Passons maintenant aux dérivés en ce qui concerne β_j. Compte tenu de (3.6) nous avons

$$\frac{\partial \sigma_t^2(\theta_0)}{\partial \beta_j} = \omega_0 \sum_{k=0}^{\infty} B_{k,j}(1,1) + \sum_{k=2}^{\infty}\sum_{\ell=1}^{k} \alpha_{0\ell} B_{k-\ell,j}(1,1)\varepsilon_{t-k}^2, \tag{3.20}$$

où $B_{0,j} = 0$ et pour $k > 0$, les matrices $B_{k,j}$ définis en (3.11) sont pris à θ_0. Pour tout $0 \leq \ell \leq k-j$ et par (3.12) et (3.13), nous obtenons :

$$B_{k-\ell,j} \leq \sum_{m=1}^{k-\ell-j} B_0^{m-1}B_0^{k-\ell-m-j+1} + \sum_{m=k-\ell-j+1}^{k-\ell} B_0^{m-1}B^{(j-k+\ell+m)}$$

$$= (k-\ell-j)B_0^{k-\ell-j} + \sum_{m=k-\ell-j+1}^{k-\ell} B_0^{m-1}B^{(j-k+\ell+m)},$$

Avec (3.14), on aura

$$B_{k-\ell,j}(1,1) \leq (k-\ell-j)B_0^{k-\ell-j}(1,1) + B_0^{k-\ell-j}(1,1)$$
$$\leq kB_0^{k-\ell-j}(1,1). \tag{3.21}$$

Pour $k-j < \ell \leq k$ nous obtenons aussi,

$$B_{k-\ell,j} \leq \sum_{m=1}^{k-\ell} B_0^{m-1}B^{(j-k+\ell+m)}, \qquad et \;\; ainsi \quad B_{k-\ell,j}(1,1) = 0. \tag{3.22}$$

Par conséquent, de (3.22) on en déduit

$$\frac{\partial \sigma_t^2(\theta_0)}{\partial \beta_j} \leq \omega_0 \sum_{k=j}^{\infty} kB_0^{k-j}(1,1) + \sum_{k=j+1}^{\infty}\sum_{\ell=1}^{k-j} \alpha_{0\ell} k B_0^{k-\ell-j}(1,1)\varepsilon_{t-k}^2.$$

Par (3.18),on aura

$$\frac{1}{\sigma_t^2}\frac{\partial \sigma_t^2}{\partial \beta_j}(\theta_0) \leq K + \sum_{k=j+1}^{\infty}\sum_{\ell=1}^{k-j} \alpha_{0\ell} k \frac{B_0^{k-\ell-j}(1,1)\varepsilon_{t-k}^2}{\omega_0^s \underline{\alpha}^{1-s}\{B_0^{k-i_k}(1,1)\}^{1-s}}. \tag{3.23}$$

Ainsi

$$\mathbb{E}_{\theta_0}\frac{1}{\sigma_t^2}\frac{\partial \sigma_t^2}{\partial \beta_j}(\theta_0) \leq K + K\left\{\sum_{k=1}^{\infty} k\rho^{ks}\right\}\mathbb{E}_{\theta_0}[\varepsilon_t^{2s}] < \infty, \tag{3.24}$$

par des arguments déjà utilisés dans (3.21). Cela permet de conclure que la première formule du lemme (3.4.2) existe. Appliquons l'inégalité de Hölder dans (3.18) , (3.25) avec s tel que $\mathbb{E}\varepsilon_t^{4s} < \infty$, on peut montrer que $\|\sigma_t^{-2}\partial\sigma_t^2(\theta_0)/\partial\theta\|_2 < \infty$. La seconde formule du lemme (3.4.2) donc existe . Passons maintenant aux dérivés de second ordre de σ_t^2. Il résulte de (3.9) que

$$\frac{\partial^2\sigma_t^2}{\partial\omega^2} = \frac{\partial^2\sigma_t^2}{\partial\omega\partial\alpha_i} = 0 \quad et \quad \frac{\partial^2\sigma_t^2}{\partial\omega\partial\beta_j} = \sum_{k=1}^{\infty} B_{k,j}(1,1).$$

Ainsi, $\partial^2\sigma_t^2/\partial\omega\partial\beta_j \leq \sum_{k=j}^{\infty} kB_0^{k-j}(1,1) < \infty$ par (3.21) , (3.22) avec $\ell = 0$, ce qui prouve que $\partial^2\sigma_t^2(\theta_0)/\partial\omega\partial\theta_i$ est borné et admet des moments dans n'importe quel ordre. La même conclusion vaut pour $\{\partial^2\sigma_t^2(\theta_0)/\partial\omega\partial\theta_i\}/\sigma_t^2(\theta_0)$. Par (3.9) , (3.10) nous trouvons

$$\frac{\partial^2\sigma_t^2}{\partial\alpha_i\partial\alpha_j} = 0 \quad et \quad \frac{\partial^2\sigma_t^2}{\partial\alpha_i\partial\beta_j} = \sum_{k=2}^{\infty} B_{k-i,j}(1,1)\varepsilon_{t-k}^2,$$

et les arguments utilisés pour la dérivée du premier ordre en ce qui concerne β_j prouve que $\{\partial^2\sigma_t^2(\theta_0)/\partial\alpha_i\partial\theta\}$ est intégrable. Différenciation (3.22) en ce qui concerne $\beta_{j'}$ donne

$$\frac{\partial^2\sigma_t^2}{\partial\beta_j\partial\beta_{j'}} = \omega_0\sum_{k=0}^{\infty} B_{k,j,j'}(1,1) + \sum_{k=2}^{\infty}\sum_{\ell=1}^{k}\alpha_{0\ell}B_{k-\ell,j,j'}(1,1)\varepsilon_{t-k}^2 \tag{3.25}$$

où

$$B_{k,j,j'} = \frac{\partial B_{k,j}}{\partial \beta_{j'}} = \sum_{m=1}^{k} B_{m-1,j'} B^{(j)} B_0^{k-m} + \sum_{m=1}^{k} B_0^{m-1} B^{(j)} B_{k-m,j'}$$
$$:= B_{k,j,j'}^{(1)} + B_{k,j,j'}^{(2)}.$$

Nous donnons d'abord une limite aux termes du formulaire $B^{(j)} B_{k,j'}$ impliqué dans $B_{k,j,j'}^{(2)}$. Tout d'abord, notez que lorsque $k \leq p$, seul le premier k lignes de B_0^k contiennent des termes dépendant de la β_j. Ainsi le dernier $p-k+1$ lignes de $B_{k,j'}$ sont égaux à zéro, et il en résulte que

$$B^{(j)} B_{k,j'} = 0 \quad \textit{pour} \quad k < j. \tag{3.26}$$

En utilisant successivement (3.13) , (3.12) , (3.15) , on obtient, pour $j, j' = 1, \ldots, p$ et $k > 0$,

$$\begin{aligned} B^{(j)} B_{k,j'} &= \sum_{n=1}^{k} B^{(j)} B_0^{n-1} B^{(j')} B_0^{k-n} \\ &\leq \sum_{n=1}^{j} B^{(j-n+1)} B^{(j')} B_0^{k-n} + \sum_{n=j+1}^{k} B_0^{n-j} B^{(j')} B_0^{k-n} \\ &= B^{(j')} B_0^{k-j} + \sum_{n=j+1}^{k} B_0^{n-j} B^{(j')} B_0^{k-n}, \end{aligned}$$

où par convention $B_0^k = B^{(k+1)} = 0$ pour $k < 0$ et $\sum_{n=k}^{k'} x_n = 0$ pour $k > k'$. En utilisant à nouveau (3.13), (3.12), nous obtenons

$$B^{(j)} B_{k,j'} \leq B^{(j+j'-k)} + \sum_{n=j+1}^{k} B_0^{n-j} B^{(j'-k+n)} \quad \textit{pour} \quad j \leq k < j+j', \tag{3.27}$$

$$\begin{aligned} B^{(j)} B_{k,j'} &= B^{(j')} B_0^{k-j} + \sum_{n=j+1}^{k-j'} B_0^{n-j} B^{j'} B_0^{k-n} + \sum_{n=k-j'+1}^{k} B_0^{n-j} B^{j'} B_0^{k-n} \\ &\leq (k-j'-j+1) B_0^{k-j-j'+1} \\ &\qquad + \sum_{n=k-j'+1}^{k} B_0^{n-j} B^{(j'-k+n)}, \quad k \geq j+j'. \end{aligned} \tag{3.28}$$

De (3.28) on obtient $B_{k,j,j'}^{(2)} := \sum_{m=1}^{k} B_0^{m-1} B^{(j)} B_{k-m,j'} = 0$ pour $k \leq j$. En utilisant le fait que la première colonne de $B^{(j)}$ est nul pour $j > 1$, (3.28), (3.29) impliquent $B_{k,j,j'}^{(2)}(1,1) = 0$ pour $j \leq k < j+j'$. Avec le même argument, (3.28), (3.29) et (3.30) montrent que pour $k \geq j+j'$

$$B_{k,j,j'}^{(2)}(1,1) = \sum_{m=1}^{k-j-j'} B_0^{m-1} B^{(j)} B_{k-m,j'}(1,1)$$

$$\leq \sum_{m=1}^{k-j-j'} (k-m-j'-j+1) B_0^{k-j-j'}(1,1)$$
$$\leq \frac{(k-j-j')(k-j-j'+1)}{2} B_0^{k-j-j'}(1,1) \leq k^2 B_0^{k-j-j'}(1,1).$$

De même nous avons $B^{(1)}_{k,j,j'}(1,1) \leq k^2 B_0^{k-j-j'}(1,1)$. Donc, de (3.27) on en déduit

$$\frac{\partial \sigma_t^2}{\partial \beta_j \partial \beta_j'}(\theta_0) \leq 2\omega_0 \sum_{k=j+j'}^{\infty} k^2 B_0^{k-j-j'}(1,1) + 2 \sum_{k=j+j'+1}^{\infty} \sum_{\ell=1}^{k-j-j'} \alpha_{0\ell} k^2 B_0^{k-\ell-j-j'}(1,1)\varepsilon_{t-k}^2,$$

Par les arguments utilisés pour montrer (3.26), nous concluons que

$$\mathbb{E}_{\theta_0} \frac{1}{\sigma_t^2} \frac{\partial^2 \sigma_t^2}{\partial \beta_j \partial \beta_j'}(\theta_0) < \infty,$$

ce qui montre l'existence de la dernière formule du lemme (3.4.2) . □

Lemme 3.2.3. ***[C. Francq et J. M. Zakoïan (2007)]*** *Sous les hypothèses du théorème (3.2.1),*

$$\mathbb{E}_{\theta_0} \left\| \frac{\partial \ell_t(\theta_0)}{\partial \theta} \frac{\partial \ell_t(\theta_0)}{\partial \theta'} \right\| < \infty \qquad et \qquad \mathbb{E}_{\theta_0} \left\| \frac{\partial^2 \ell_t(\theta_0)}{\partial \theta \partial \theta'} \right\| < \infty.$$

Preuve. Au vu du lemme (3.4.2), les dérivées de σ_t^2 divisé par σ_t^2 possèdent des moments de second ordre. Pour $\theta = \theta_0$, la variable $\varepsilon_t^2/\sigma_t^2 = \eta_t^2$ possède un moment de premier ordre et est indépendant des termes impliquant σ_t^2 et ses dérivés. Les résultats découlent ensuite de (2.38) , en utilisant l'inégalité de Hölder. □

Lemme 3.2.4. ***[C. Francq et J. M. Zakoïan (2007)]*** *Sous les hypothèse de théorème (3.2.1),*

$$J \textit{ est non-singulière} \qquad et \qquad Var_{\theta_0}\left\{ \frac{\partial \ell_t(\theta_0)}{\partial \theta} \right\} = \{k_\eta - 1\} J.$$

Preuve. La preuve découle du lemme (3.4.3) et des hypothèses d'identifiabilité **H3, H4** (et la preuve du théorème (2.2.2) (ii)). □

Le lemme suivant, avec le théorème (3.1.1) (i), montre aisément que J peut être systématiquement estimé par $\widehat{J} := \partial^2 I_n(\widehat{\theta}_n)/\partial\theta\partial\theta'$.

Lemme 3.2.5. ***[C. Francq et J. M. Zakoïan (2007)]*** *Sous les hypothèses du théorème (3.2.1) , pour tout $\varepsilon > 0$, il existe un voisinage $\mathcal{V}(\theta_0)$ de θ_0 tel que, presque sûrement,*

$$\mathbb{E}_{\theta_0} \sup_{\theta \in \mathcal{V}(\theta_0)\cap\Theta} \left\| \frac{\partial^2 \ell_t(\theta)}{\partial \theta \partial \theta'} \right\| < \infty \tag{3.29}$$

et

$$\lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} \sup_{\theta \in \mathcal{V}(\theta_0)\cap\Theta} \left\| \frac{\partial^2 \ell_t(\theta)}{\partial \theta \partial \theta'} - \frac{\partial^2 \ell_t(\theta_0)}{\partial \theta \partial \theta'} \right\| \leq \varepsilon. \tag{3.30}$$

Preuve. Lorsque **H7** est supposé, (3.31) est une conséquence de (2.38) . Supposons maintenant que **H8**, au lieu de **H7**, soit valable. Nous montrerons que le lemme (3.4.2) reste vrai dans certains voisinage de θ_0. Soit $j_0 = j_0(\theta_0)$ être le nombre entier défini dans (3.6). Soit $\mathcal{V}(\theta_0)$ être un voisinage de θ_0 tel que

$$\inf_{\theta\in\mathcal{V}(\theta_0)} \prod_{i=1}^{j_0} \alpha_i > 0 \quad et \quad \inf_{\theta\in\mathcal{V}_(\theta_0)} \beta_{j_0} > 0.$$

Pour la suite $(i_k) = (i_k(\theta_0))$ satisfaisant (3.17) et certains $\underline{\alpha} > 0$ (par exemple, on peut prendre $\underline{\alpha} = \inf_{\theta\in\mathcal{V}(\theta_0)} \min_{\{i:1\le i\le j_0\}} \alpha_i$), on a

$$\inf_{\theta\in\mathcal{V}(\theta_0)} \alpha_{ik} B^{k-i_k}(1,1) \geq \underline{\alpha} \inf_{\theta\in\mathcal{V}(\theta_0)} B^{k-i_k}(1,1) > 0.$$

Comme pour (3.18) nous avons alors :

$$\sup_{\theta\in\mathcal{V}(\theta_0)} \frac{1}{\sigma_t^2}\frac{\partial\sigma_t^2(\theta_0)}{\partial\alpha_i} \leq K\sum_{k=i}^{\infty} \sup_{\theta\in\mathcal{V}(\theta_0)} \left\{\frac{B^{k-i}(1,1)}{B^{k-i_k}(1,1)}\right\}\rho^k\varepsilon_{t-k}^{2s}, \tag{3.31}$$

en utilisant $\sup_{\theta\in\mathcal{V}(\theta_0)}\|B^k\| \leq K\rho^k$, qui est une conséquence de $\sup_{\theta\in\Theta}\rho(B) < 1$. Notez que $B_0^{k-i_k}(1,1) \neq 0$ implique $B^{k-i_k}(1,1) \neq 0$ dans $\mathcal{V}(\theta_0)$, mais ça $B_0^{k-i}(1,1) = 0$ n'implique pas $B^{k-i}(1,1) = 0$ dans $\mathcal{V}(\theta_0)$. Cependant, dans tous les cas, nous avons

$$\frac{B^{k-i}(1,1)}{B^{k-i_k}(1,1)} \leq \frac{1}{\beta_{j_0}^{q_i}}.$$

En effet, la dernière égalité est simple lorsque $B^{k-i}(1,1) = 0$ et découle de (3.20) lorsque $B^{k-i}(1,1) \neq 0$. Il en résulte que le sup à l'intérieur de la somme en (3.33) est borné. Donc.

$$\left\|\sup_{\theta\in\mathcal{V}(\theta_0)} \frac{1}{\sigma_t^2}\frac{\partial\sigma_t^2(\theta_0)}{\partial\alpha_i}\right\|_3 < \infty.$$

Une existence similaire de moments peut être montrer pour les autres dérivés impliqués dans la dérivée seconde de $\ell_t(\theta)$, résulte (3.31).
Maintenant, sous **H7** ou **H8**, le théorème ergodique montre que

$$\lim_{n\to\infty}\frac{1}{n}\sum_{t=1}^{n}\sup_{\theta\in\mathcal{V}(\theta_0)\cap\Theta}\left\|\frac{\partial^2\ell_t(\theta)}{\partial\theta\partial\theta'} - \frac{\partial^2\ell_t(\theta_0)}{\partial\theta\partial\theta'}\right\| = \mathbb{E}_{\theta_0}\sup_{\theta\in\mathcal{V}(\theta_0)\cap\Theta}\left\|\frac{\partial^2\ell_t(\theta)}{\partial\theta\partial\theta'} - \frac{\partial^2\ell_t(\theta_0)}{\partial\theta\partial\theta'}\right\|.$$

Cette attente diminue à 0 lorsque le voisinage $V(\theta_0)$ diminue jusqu'au singleton θ_0. Ainsi (3.32) est également prouvé. □

Le lemme suivant montre que les valeurs initiales sont asymptotiquement négligeables.

Lemme 3.2.6. ***[C. Francq et J. M. Zakoïan (2007)]*** *Sous les hypothèses du théorème (3.2.1)*

$$\left\|n^{-1/2}\sum_{t=1}^{n}\left\{\frac{\partial\ell_t(\theta_0)}{\partial\theta} - \frac{\partial\widetilde{\ell_t}(\theta_0)}{\partial\theta}\right\}\right\| \to 0 \tag{3.32}$$

et

$$\sup_{\theta\in\mathcal{V}(\theta_0)\cap\Theta}\left|n^{-1}\sum_{t=1}^{n}\left\{\frac{\partial^2\ell_t(\theta_0)}{\partial\theta\partial\theta'}-\frac{\partial^2\widetilde{\ell}_t(\theta)}{\partial\theta\partial\theta'}\right\}\right|\xrightarrow{P}0. \tag{3.33}$$

Preuve. Nous avons :

$$\sup_{\theta\in\mathcal{V}(\theta_0)\cap\Theta}\left|n^{-1}\sum_{t=1}^{n}\left\{\frac{\partial^2\ell_t(\theta)}{\partial\theta_i\partial\theta_j}-\frac{\partial^2\widetilde{\ell}_t(\theta)}{\partial\theta_i\partial\theta_j}\right\}\right|\leq Kn^{-1}\sum_{t=1}^{n}\rho^t\Upsilon_t,$$

où $K>0$, $\rho\in[0,1]$ et

$$\Upsilon_t=\sup_{\theta\in\mathcal{V}(\theta_0)\cap\Theta}\left\{1+\frac{\varepsilon_t^2}{\sigma_t^2}\right\}\left\{1+\frac{1}{\sigma_t^2}\frac{\partial^2\sigma_t^2}{\partial\theta_i\partial\theta_j}+\frac{1}{\sigma_t^2}\frac{\partial\sigma_t^2}{\partial\theta_i}\frac{1}{\sigma_t^2}\frac{\partial\sigma_t^2}{\partial\theta_j}\right\}.$$

On sait que, dans l'hypothèse de stationnarité stricte **H2**, ε_t admet un moment d'ordre $6s$ pour certains $s>0$.
En utilisant l'inégalité de Hölder, il s'ensuit que $\mathbb{E}\Upsilon_t^s<\infty$. L'inégalité de Markov et l'inégalité élémentaire $(a+b)^s\leq a^s+b^s$ pour tous $a,\ b\geq 0,\ s\in[0,1]$ entraîner

$$\forall\,\varepsilon>0,\quad \mathbb{P}\left(Kn^{-1}\sum_{t=1}^{n}\rho^t\Upsilon_t>\varepsilon\right)\leq K\mathbb{E}(\Upsilon_t^s)\varepsilon^{-s}n^{-s}\sum_{t=1}^{n}\rho^{st}\to 0,\quad \textit{quand}\ \ n\to\infty,$$

ce qui montre (3.35). La convergence (3.34) est représentée par des arguments similaires. □

Le lemme suivant établit la normalité asymptotique.

Lemme 3.2.7. ***[C. Francq et J. M. Zakoïan (2007)]*** *Sous les hypothèses du théorème (3.2.1) $J_n:=\frac{\partial^2 I_n(\theta_0)}{\partial\theta\partial\theta'}$ est p.s. une matrice définie positive pour n suffisamment grand et*

$$Z_n:=-J_n^{-1}\sqrt{n}\frac{\partial I_n(\theta_0)}{\partial\theta}\xrightarrow{\mathcal{L}}Z,\qquad avec\qquad Z\sim\mathcal{N}\{0,(k_\eta-1)J^{-1}\}.$$

Preuve. Le théorème central limite pour les différences de martingale stationnaire de carré intégrables implique que

$$\sqrt{n}\frac{\partial I_n(\theta_0)}{\partial\theta}=n^{-1/2}\sum_{t=1}^{n}(1-\eta_t^2)\frac{1}{\sigma_t^2}\frac{\partial\sigma_t^2}{\partial\theta}(\theta_0)\xrightarrow{\mathcal{L}}\mathcal{N}\{0,(k_\eta-1)J\}.$$

Le théorème ergodique et le lemme (3.4.4) montrent que $J_n\to J$ presque sûrement lorsque $n\to\infty$. La conclusion découle du lemme Slutsky. □

Retour à la preuve du théorème (3.2.1) :
La notation $a_n\overset{o_p(1)}{=}b_n$ est synonyme de suites (a_n) et (b_n) tel que a_n-b_n converge à zéro en probabilité. Quand $\theta_0\in\overset{\circ}{\theta}$, (preuve du théorème(2.2.2) a montré que, sous H1 - H5,

$$\sqrt{n}(\widehat{\theta}_n-\theta_0)\overset{o_p(1)}{=}Z_n:=-J^{-1}\sqrt{n}\frac{\partial I_n(\theta_0)}{\partial\theta}. \tag{3.34}$$

Cette relation n' est pas vérifiée quand certains coefficients égaux à zéro parceque, au moins une composante du vecteur de gauche est une variable aléatoire positive. Au lieu de cela, nous établirons que, pour tous $\theta_0 \in \Theta$,

$$\sqrt{n}(\widehat{\theta}_n - \theta_0) \overset{o_p(1)}{=} \lambda_n^\Lambda \tag{3.35}$$

où $\lambda_n^\Lambda = \arg\inf_{\lambda\in\Lambda}\{\lambda - Z_n\}' J_n \{\lambda - Z_n\}$. Quand $\theta_0 \in \overset{o}{\Theta}$ on a $\lambda_n^\Lambda = Z_n$ parce que $\Lambda = \mathbb{R}^{p+q+1}$, alors (3.37) se réduit à (3.38) dans ce cas. Dans le cas général, λ_n^Λ peut être interprété comme la projection orthogonale de Z_n sur Λ pour le produit intérieur $< x, y > J_n = x' J_n y$. Il sera commode d'approcher cette projection de celle de Z_n sur l'espace $\sqrt{n}(\Theta - \theta_0)$ qui, en supposant que Θ contient un hypercube, augmente à Λ. Cette projection peut être écrite comme : $\sqrt{n}(\theta_{J_n}(Z_n) - \theta_0)$ avec

$$\theta_{J_n}(Z_n) = \arg\inf_{\theta\in\Theta} \|Z_n - \sqrt{n}(\theta - \theta_0)\|_{J_n}, \quad \textit{tandis que} \quad \lambda_n^\Lambda = \arg\inf_{\lambda\in\Lambda} \|Z_n - \lambda\|_{J_n}.$$

La preuve du théorème (3.2.1) repose sur une forme quadratique environ θ_0 de la fonction de quasi-vraisemblance. En utilisant développement de Taylor pour une fonction avec des dérivées partielles droites, on obtient, pour tous θ et θ_0 dans Θ,

$$\widetilde{I}_n(\theta) = \widetilde{I}_n(\theta_0) + \frac{\partial \widetilde{I}_n(\theta_0)}{\partial\theta'}(\theta - \theta_0) + \frac{1}{2}(\theta - \theta_0)' \frac{\partial^2 \widetilde{I}_n(\theta_0)}{\partial\theta\partial\theta'}(\theta - \theta_0) + R_n(\theta) \tag{3.36}$$

$$= \widetilde{I}_n(\theta_0) - \frac{1}{2n} Z_n' J_n \sqrt{n}(\theta - \theta_0) - \frac{1}{2n}\sqrt{n}(\theta - \theta_0)' J_n Z_n$$
$$+ \frac{1}{2}(\theta - \theta_0)' J_n (\theta - \theta_0) + R_n(\theta) + R_n^*(\theta)$$

$$= \widetilde{I}_n(\theta_0) + \frac{1}{2n}\|Z_n - \sqrt{n}(\theta - \theta_0)\|^2_{J_n} - \frac{1}{2n} Z_n' J_n Z_n + R_n(\theta) + R_n^*(\theta), \tag{3.37}$$

où $R_n(\theta)$ et $R_n^*(\theta)$ sont les termes restants qui seront discutés ci-dessous.
Nous établirons les résultats intermédiaires suivants. Pour tous $\theta_0 \in \Theta$,

1. $\sqrt{n}(\theta_{J_n}(Z_n) - \theta_0) = O_p(1)$,
2. $\sqrt{n}(\widehat{\theta}_n - \theta_0) = O_p(1)$
3. Pour toute suite (θ_n) tel que $\sqrt{n}(\theta_n - \theta_0) = O_p(1)$,

$$R_n(\theta_n) = o_p(n^{-1}), \quad R_n^*(\theta_n) = o_p(1)(n^{-1}),$$

4. $\|Z_n - \sqrt{n}(\widehat{\theta}_n - \theta)\|^2_{J_n} \overset{o_p(1)}{=} \|Z_n - \lambda_n^\Lambda\|^2_{J_n}$,
5. $\sqrt{n}(\widehat{\theta} - \theta_0) \overset{o_p(1)}{=} \lambda_n^\Lambda$,
6. $\lambda_n^\Lambda \xrightarrow{\mathcal{L}} \lambda^\Lambda$.

Pour prouver (1), nous remarquons d'abord que, eu égard au lemme (3.4.4) , l'affirmation selon laquelle $\|x\|_{J_n}$ est une norme,p.s. pour n grand, est justifié. L'inégalité du triangle donne

$$\begin{aligned}\|\sqrt{n}(\theta_{J_n}(Z_n)-\theta_0)\|_{J_n} &\leq \|Z_n-\sqrt{n}(\theta_{J_n}(Z_n)-\theta_0)\|_{J_n}+\|Z_n\|_{J_n}\\ &\leq \|Z_n\|_{J_n}+\|Z_n\|_{J_n}=O_p(1),\end{aligned}$$

où la deuxième inégalité est vérifiée parce que $\theta_0 \in \Theta$ et $\theta_{Jn}(Z_n)$ minimise $\|Z_n-\sqrt{n}(\theta-\theta_0)\|_{J_n}$ plus de Θ et l'égalité découle du lemme (3.4.7) . Ainsi (1) est prouvé.

Par dévloppement de Taylor :

$$\widetilde{I}_n(\theta)=\widetilde{I}_n(\theta_0)+\frac{\partial \widetilde{I}_n(\theta_0)}{\partial \theta'}(\theta-\theta_0)+\frac{1}{2}(\theta-\theta_0)'\left[\frac{\partial^2 \widetilde{I}_n(\theta^*_{ij})}{\partial\theta\partial\theta'}\right](\theta-\theta_0),$$

où le θ^*_{ij} se situer entre θ et θ_0, le premier terme restant de (3.38) satisfait

$$R_n(\theta)=\frac{1}{2}(\theta-\theta_0)'\left\{\left[\frac{\partial^2 \widetilde{I}_n(\theta^*_{ij})}{\partial\theta\partial\theta'}\right]-\frac{\partial^2 \widetilde{I}_n(\theta_0)}{\partial\theta\partial\theta'}\right\}(\theta-\theta_0). \tag{3.38}$$

Quand $\theta=\widehat{\theta}_n$, selon le théorème 1 (i), (3.35) , (3.32) , la différence entre les dérivés de second ordre tend vers zéro en probabilité lorsque n tends vers l'infini. Donc

$$R_n(\widehat{\theta}_n)=o_p(\|\widehat{\theta}_n-\theta_0\|^2)=o_p(\|\widehat{\theta}_n-\theta_0\|^2_{J_n}).$$

Le second terme restant de (3.39) est donné par

$$R^*_n(\theta)=\left\{\frac{\partial \widetilde{I}_n(\theta_0)}{\partial\theta}-\frac{\partial I_n(\theta_0)}{\partial\theta}\right\}(\theta-\theta_0)+\frac{1}{2}(\theta-\theta_0)'\left\{\frac{\partial^2 \widetilde{I}_n(\theta_0)}{\partial\theta\partial\theta'}-J_n\right\}(\theta-\theta_0). \tag{3.39}$$

Par conséquent, compte tenu de (3.34) , (3.35), nous avons

$$R^*_n(\widehat{\theta}_n)=o_p(n^{-1/2}\|\widehat{\theta}-\theta_0\|_{J_n})+o_p(\|\widehat{\theta}_n-\theta_0\|^2_{J_n}).$$

Nous avons alors

$$\begin{aligned}\widetilde{I}_n(\widehat{\theta}_n)-\widetilde{I}_n(\theta_0) &= \frac{1}{2n}\|Z_n-\sqrt{n}(\widehat{\theta}_n-\theta_0)\|^2_{J_n}-\frac{1}{2n}\|Z_n\|^2_{J_n}+R_n(\widehat{\theta}_n)+R^*_n(\widehat{\theta}_n)\\ &= \frac{1}{2n}\Big\{\|Z_n-\sqrt{n}(\widehat{\theta}_n-\theta_0)\|^2_{J_n}-\|Z_n\|^2_{J_n}\\ &\quad +o_p(\|\sqrt{n}(\widehat{\theta}_n-\theta_0)\|_{J_n})+o_p(\|\sqrt{n}(\widehat{\theta}_n-\theta_0)\|^2_{J_n})\Big\}\leq 0,\end{aligned}$$

parce que $\widehat{\theta}_n$ minimise $\widetilde{l}_n(.)$ plus de Θ. Il s'ensuit que

$$\begin{aligned}\|Z_n-\sqrt{n}(\widehat{\theta}_n-\theta_0)\|^2_{J_n} &\leq \|Z_n\|^2_{J_n}+o_p(\|\sqrt{n}(\widehat{\theta}_n-\theta_0)\|_{J_n})+o_p(\|\sqrt{n}(\widehat{\theta}_n-\theta_0)\|^2_{J_n})\\ &\leq \{\|Z_n\|_{J_n}+o_p(\|\sqrt{n}(\widehat{\theta}_n-\theta_0)\|_{J_n})\}^2,\end{aligned}$$

où la dernière inégalité tient parce que $\|Z_n\|_{J_n} = O_P(1)$. On en déduit par l'inégalité triangulaire que :

$$\begin{aligned}\|\sqrt{n}(\widehat{\theta}_n - \theta_0)\|_{J_n} &\leq \|\sqrt{n}(\widehat{\theta}_n - \theta_0) - Z_n\|_{J_n} + \|Z_n\|_{J_n} \\ &\leq 2\|Z_n\|_{J_n} + o_p(\|\sqrt{n}(\widehat{\theta}_n - \theta_0)\|_{J_n}).\end{aligned}$$

Ainsi $\|\sqrt{n}(\widehat{\theta}_n - \theta_0)\|_{J_n}\{1 + o_p(1)\} \leq 2\|Z_n\|_{J_n} = O_p(1)$, et (2)suit facilement.
Compte tenu de (3.40) , (3.35) , (3.32) , nous avons $R_n(\theta_n) = o_P(\|\theta_n - \theta_0\|^2) = o_P(n^{-1})$, ce qui prouve la première partie de (3). La deuxième égalité découle également de (3.41) et $R_n^*(\theta_n) = o_p(n^{-1/2}\|\theta_n - \theta_0\|) + o_p(n^{-1})$. Par (A.33) , par le fait que $\widehat{\theta}_n$ minimise $\widetilde{l}_n(.)$ et cela $\theta_{Jn}(Zn)$ minimise $\|Z_n - \sqrt{n}(\theta - \theta_0)\|_{J_n}$, et par (1)-(3) nous avons

$$\begin{aligned}0 &\leq \|Z_n - \sqrt{n}(\widehat{\theta}_n - \theta_0)\|^2_{J_n} - \|Z_n - \sqrt{n}(\theta_{J_n}(Z_n) - \theta_0)\|^2_{J_n} \\ &= 2n\{\widetilde{I}_n(\widehat{\theta}_n - \widetilde{I}_n(\theta_{J_n}(Z_n))\} - 2n\{(R_n + R_n^*)(\widehat{\theta}_n) - (R_n + R_n^*)(\theta_{J_n}(Z_n))\} \\ &\leq -2n\{(R_n + R_n^*)(\widehat{\theta}_n) - (R_n + R_n^*)(\theta_{J_n}(Z_n))\} = o_p(1)\end{aligned}$$

Maintenant depuis $\sqrt{n}(\theta_{J_n}(Z_n) - \theta_0) = \lambda_n^\Lambda$ pour n suffisamment grand, (4) prouvé.
Le vecteur λ_n^Λ être la projection de Z_n sur le l'ensemble convexe Λ pour le produit scalaire $< x, y >_{J_n}$, il est caractérisé par $\lambda_n^\Lambda \in \Lambda$, $< Z_n - \lambda_n^\Lambda, \lambda_n^\Lambda - \lambda >_{J_n} \ \geq 0, \ \forall \ \lambda \in \Lambda$. Ainsi

$$\begin{aligned}\|\sqrt{n}(\widehat{\theta}_n - \theta_0) - Z_n\|^2_{J_n} &= \|\sqrt{n}(\widehat{\theta}_n - \theta_0) - \lambda_n^\Lambda\|^2_{J_n} + \|\lambda_n^\Lambda - Z_n\|^2_{J_n} \\ &\quad +2 < \sqrt{n}(\widehat{\theta}_n - \theta_0) - \lambda_n^\Lambda, \lambda_n^\Lambda - Z_n >_{J_n} \\ &\geq \|\sqrt{n}(\widehat{\theta}_n - \theta_0) - \lambda_n^\Lambda\|^2_{J_n} + \|\lambda_n^\Lambda - Z_n\|^2_{J_n}.\end{aligned}$$

Donc, par (4)

$$\|\sqrt{n}(\widehat{\theta}_n - \theta_0) - \lambda_n^\Lambda\|^2_{J_n} \leq \|Z_n - \sqrt{n}(\widehat{\theta}_n - \theta_0)\|^2_{J_n} - \|Z_n - \lambda_n^\Lambda\|^2_{J_n} = o_p(1),$$

et (v) est prouvé.

Comme $(Z_n, J_n) \xrightarrow{\mathcal{L}} (Z, J)$ par le lemme(3.4.7), alors (6) est vérifiée.
La preuve du théorème(3.4.1) découle de (5) et de (6).

Nous avons étudié dans ce chapitre les résultats de convergence et de normalité asymptotique, publié par C. Francq et J. M. Zakoïan en 2007, lorsque certains coefficients sont nuls. Indépendament des difficultés mathématiques encontrées lors du déveleppement des calculs des divers lemmes, propositions et théorèmes, les résultats sont très satisfaisant sur le plan statistique. En effet, l'hypothèse restreignante **H1** a été relachée et remplacée par d'autres qui sont plus réalisables en pratique.

Conclusion

Ce travail constitue une introduction aux modèles ARCH/GARCH en donnant quelques outils de base de la théorie des séries chronologiques qui présentent une dynamique non linéaire.

Le concept de la variance conditionnelle joue un grand rôle au depuis les années quatre vingt avec l'article fondateur de Engle (1982). Il caractérise les modèles venus élargir la classe des modèles classiques fondés essentiellement sur une structure de dépendance linéaire entre une variable à un instant t et ses valeurs passées et celles d'un bruit blanc et de ses valeurs passées. D'un autre coté, il est important de noter que les séries financières sont aussi caractérisées par une volatilité non stationnaire et par des phénomènes d'asymétrie qui ne peuvent pas être pris en compte par les modélisations classique.

Les représentations des différentes approches qui s'offrent à nous pour estimer les paramètres du modèle ARCH/GARCH stationnaire, entre autres la méthode de moindres carrés de Bose et Mukherjee (2003) et le maximum de vraisemblance. Nous les avons détaillé, par suite un intérêt particulier a été accordé à l'estimateur (M.V) du modèle GARCH(p, q) lorsque certains coefficient sont nuls. Les résultats de convergence asymptotiques et de normalité asymptotique ont été présentés et celà sous d'autres conditions que les hypothèses classiques.

Ce travail ouvre de perspectives futures dans le domaine de l'inférence statistique. Plus praticulièrement, il serait intéressant de développer la même idée sur un autre modèle de série chronologique, tel que le IGARCH, MGARCH et de trouver des conditions moins fortes en supposant que certain coefficients sont nuls afin d'obtenir la consistence et la normalité asymptotique.

Annexe

Martingale

Définition 3.2.1. *(Martingale) Soit* $\{X_t;\ t \in \mathbb{N}\}$ *suite de variables aléatoires réelles (v.a.r) sur un espace probabilisé*

$(\Omega, \mathcal{F}, \mathbf{p})$*, et* $(\mathcal{F}_t)_{t\in\mathbb{N}}$ *est une suite de tribu. La suite* $\{(X_t, \mathcal{F}_t) : t = 1, 2 \ldots\}$ *est une martingale si et seulement si :*

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

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

Définition 3.2.2. *(Différence de martingale) Soient* $\{X_t;\ t \in \mathbb{N}\}$ *une suite de variables aléatoires réelles (v.a.r), et* $(\mathcal{F})_{t\in\mathbb{N}}$ *est une suite de tribu. La suite* $\{(X_t, \mathcal{F}_t) : t = 1, 2, \ldots\}$ *est une différence de martingale (ou une suite d'accroissements de martingale) si et seulement si :*

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

Ergodicité

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

Théorème 3.2.2. *Si* $\{Z_t;\ t \in \mathbb{Z}\}$ *est une suite fortement stationnaire et ergodique.*
et si $f : \mathbb{R}^\infty \to \mathbb{R}$ *est une fonction mesurable.*
et soit $Y_t = f(\cdots, Z_{t-1}, Z_t, Z_{t-1}, \cdots)$*, alors* $\{Y_t;\ t \in \mathbb{Z}\}$ *reste un suite fortement stationnaire et ergodique.*

Théorème 3.2.3. *d'ergodicité) Si $\{Z_t;\ t \in \mathbb{Z}\}$ est strictement stationnaire et ergodique, si f est mesurable et si $\mathbb{E}\{|f(\ldots, Z_{t-1}, Z_t, Z_{t+1}, \ldots)|\} < \infty$, alors :*

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

Théorème de chung-fuchs

Théorème 3.2.4. *(Sur les marches aléatoires) Si $X_, \ldots, X_n$ est une suite* iid *telle que $\mathbb{E}(X) = 0$ et $\mathbb{E}|X_1| > 0$, alors p.s.*

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

et

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

Critère de Cauchy

Définition 3.2.3. *(Critère de Cauchy pour la convergence d'une suite de terme $a_n \geq 0$)*
Soit $\lambda = \limsup a_n^{\frac{1}{n}}$.
Si $\lambda < 1 \Longrightarrow \sum_{n=1}^{\infty} a_n < \infty$,
Si $\lambda > 1 \Longrightarrow \sum_{n=1}^{\infty} a_n = +\infty$.

La régression Linéaire Multiple

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

Définition 3.2.4. *De manière générale, le modèle linéaire peut s'écrire de la manière suivante :*

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

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

Notations

On rencontre principalement trois types de notations :

Notation simple

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

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

Notation vectorielle

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

$$y_i = x'_i\beta + \mu_i.$$

Notation matricielle

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

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

avec : $y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, X = \begin{pmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_n \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & \cdots & x_{1k} \\ 1 & x_{21} & \cdots & x_{2k} \\ 1 & \ddots & \vdots & \\ 1 & x_n1 & \cdots & x_{nk} \end{pmatrix}, \ \beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{pmatrix}, \ \mu = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{pmatrix}.$

Théorème de Wold-Cramer

Théorème 3.2.5. *Pour une suite (Z_n) de vecteurs aléatoires de dimension d, $Z_n \xrightarrow{\mathcal{L}} Z$ si et seulement si pour tout $\lambda \in \mathbb{R}^d$, on a $\lambda' Z_n \xrightarrow{\mathcal{L}} \lambda' Z$.*

Théorème central limite T.C.L pour différence de martingale stationnaire

Théorème 3.2.6. *Si $(\nu_t, \mathcal{F})$ est une différence de martingale ν_t est $\mathcal{F}_t$-mesurable et $\mathbb{E}(\nu_t|\mathcal{F}_{t-1}) = 0$, stationnaire ergodique, de carré intégrable, telle que $V(\nu_t) = \sigma_\nu^2 \neq 0$, alors*

$$\frac{1}{\sqrt{n}} \sum_{t=1}^{n} \nu \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma_\nu^2).$$

T.C.L de Lindeberg

Théorème 3.2.7. *On suppose que, pour chaque $n > 0$, $(\eta_{nk}^2, \mathcal{F}_{nk})_{k\in\mathbb{N}}$ est une différence de martingale de carré intégrable. Soit $\sigma_{nk}^2 = \mathbb{E}(\eta_{nk}^2|\mathcal{F}_{n(k-1)})$. Si*

$$\sum_{k=1}^{n} \sigma_{nk}^2 \xrightarrow{P} \sigma_0^2 \quad \textit{quand} \quad n \to \infty,$$

où σ_0^2 est une constante strictement positive, et

$$\sum_{k=1}^{n} \mathbb{E}(\eta_{nk}^2 \mathbb{I}_{|\eta_{nk}|\geq\epsilon}) \to 0 \quad \textit{quand} \quad n \to \infty,$$

pour chaque réel positif ϵ, alors $\sum_{k=1}^{n} \eta_{nk} \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma_0^2)$.

Lemme de Slutsky

Lemme 3.2.8. *Soit* $\{X_n\}_{n=0,\ldots,\infty}$ *une suite de variables aléatoires qui tend en loi vers* X *et soit* $\{Y_n\}_{n=0,\ldots,\infty}$ *une suite de variables aléatoires qui tend en probabilité vers une constante* $c \in \mathbb{R}$, *alors* $X_n + Y_n$ *tend en loi vers* $X + c$ *et* $X_n Y_n$ *tend en loi vers* cX.

Lemme de césaro

Lemme 3.2.9. *Soit* $\{a_n;\ n \in \mathbb{N}^*\}$ *une suite de nombre réels ou complexes, si elle converge vers* ℓ *alors la suite des moyennes de Césàro converge également vers* ℓ *:*

$$C_n = \sum_{k=1}^{n} a_k \longrightarrow \ell, \quad quand \quad n \to \infty.$$

Lemme Borel-cantelli

Lemme 3.2.10. *Soit* $(A_n)(n \geq 1)$ *une suite d'évènement, posons* $A^* = \limsup_{n\to\infty} A_n$.
a Si $\sum_{n\geq 1} P(A_n) < \infty$, *alors* $P(A^*) = 0$. *Autrement dit, avec une probabilité égale à 1 au plus un nombre fini d'évènement* A_n *se réalisent .*
b Supposons les évènement A_n *indépendant deux à deux. Si* $\sum_{n\geq 1} P(A_n) = +\infty$, *alors* $P(A^*) = 1$. *Autrement dit, avec une probabilité égale à 1, une infinité d'évènement* A_n *se réalisent.*

Inégalité de HÖlder

Définition 3.2.5. *Soit* $p, q \in [1, \infty[$, $\frac{1}{p} + \frac{1}{q} = 1$, *et soit* X *et* Y *deux variables aléatoires telles que* $\mathbb{E}(|X|^p) < \infty$ *et* $\mathbb{E}(|Y|^q) < \infty$. *Alors* $\mathbb{E}(|XY|) < \infty$ *et*

$$\mathbb{E}(|XY|) \leq (\mathbb{E}|X|^p)^{1/p} (\mathbb{E}|Y|^q)^{1/q}.$$

Inégalité de Minkowski

Définition 3.2.6. *Soit* $p, q \in [1, \infty[$, *et soit* X *et* Y *deux variables aléatoires alors :*

$$\mathbb{E}(|X + Y|) \leq (\mathbb{E}|X|^p)^{1/p} + (\mathbb{E}|Y|^q)^{1/q}.$$

Cauchy-schwartz

Définition 3.2.7. *Si les variables* X *et* Y *sont de carré intégrable, alors :*

$$|\mathbb{E}(XY)| \leq \sqrt{\mathbb{E}(X^2)}\sqrt{\mathbb{E}(Y^2)}.$$

Inégalité de Markov

Définition 3.2.8. *Soit Z une variable aléatoire réelle définie sur un espace probabilisé $(\Omega, \mathcal{A}, P)$ et supposée p.s. positive ou nulle alors : $\forall\ a > 0$, $P(Z \geq a) \leq \mathbb{E}(Z)/a$.*

Inégalité de Jensen

Définition 3.2.9. *Soit f une fonction convexe sur un intervelle réel I et X une variable aléatoire à valeurs dans I, dont $\mathbb{E}(X)$ existe. Alors*

$$f(\mathbb{E}(X)) \leq \mathbb{E}(f(X)).$$

Théorème de beppo-levi

Théorème 3.2.8. *Soit $(Z_n)_{n\in\mathbb{N}}$ une suite* p.s. *croissante dans $\mathcal{L}(\Omega, \mathcal{F}, \mathcal{P})$ de limite Z*

$$\mathbb{E}(Z_n)_{n\in\mathbb{N}} \textit{ est } \text{p.s.} \textit{ croissante et convergente vers } \mathbb{E}(Z), \quad \textit{lorsque} \quad n \to \infty.$$

Convergence en probabilité

Soit $\{a_n,\ \ n = 1, 2, \ldots\}$ une suite de nombres réelles strictement positifs et soit $\{X_n,\ \ n = 1, 2, \ldots\}$ une suite de variables aléatoires dans le même espace probabilisé.

Définition 3.2.10. ***(Convergence en probabilité vers zéro)*** *On dit que X_n converge vers zéro en probabilité, on écrit $X_n = o_p(1)$ ou $X_n \xrightarrow{P} 0$ si pour tout $\varepsilon > 0$,*

$$\mathbb{P}(|X_n| > \varepsilon) \to 0 \quad \textit{quand} \quad n \to \infty.$$

Définition 3.2.11. ***(Bornitude en probabilité)*** *On dit que $\{X_n\}$ est borné en probabilité, on écrit $X_n = O_p(1)$, si pour tout $\varepsilon > 0$, $\exists\delta(\varepsilon) \in [0, \infty]$,*

$$\mathbb{P}(|X_n| > \delta(\varepsilon)) < \varepsilon \quad \forall\ \ n.$$

Définition 3.2.12. ***(Convergence en probabilité et l'ordre en probabilité)*** *(i) X_n converge en probabilité vers une variable aléatoire X, on écrit $X_n \xrightarrow{P} X$, si et seulement si $X_n - X = o_p(1)$.*
(ii)$X_n = o_p(a_n)$ si et seulement si $a_n^{-1}X_n = o_p(1)$.
(iii)$X_n = O_p(a_n)$ si et seulement si $a_n^{-1}X_n = O_p(1)$.

Proposition 3.2.9. *Si X_n et Y_n, $n = 1, 2, \ldots$, des variables aléatoires dans un même espace probabilisé et $a_n > 0$, $b_n > 0$, $n = 1, 2, \ldots$, alors*
(i) si $X_n = o_p(a_n)$ et $Y_n = o_p(b_n)$, on aura
$X_nY_n = o_p(a_nb_n)$, $\qquad X_n + Y_n = o_p(\max(a_n, b_n))$,
$|X_n|^r = o_p(a_n^r)$ *pour* $r > 0$.
(ii) si $X_n = o_p(a_n)$ et $Y_n = O_p(b_n)$ on aura :
$X_nY_n = o_p(a_nb_n)$.

(i) Reste vraie même si on remplace o_p par O_p.

Bibliographie

[1] C. Franq et J. M. Zakoïn., testing the nullity of GARCH coeficient : correction of the standart tests and relative efficiency comparosons., vol.104, No. 485(March 2009),p. 313-324.

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

[3] C. Franq. and Zakoïan. J.M, Modèles GARCH et à la volatilité stockastique,(2012), édition WILEY.

[4] D. W. K. Andrews, Estimation when a parameter is on a boundary of the parameter space : Part II, Yale University, 1997 (Unpublished manuscript).

[5] Jensen, S. T. and Rahbek, A., (2004) Asymptotic normality of the QMLE estimator og ARCH in the nonstationnary case. Econometrica 72, 641-646.

[6] P. Zaffaroni., stationarity and memory of ARCH(∞) models. Econometric theory, vol.20, No. 1(Feb.,2004), p.147-160.

[7] P. Bougerol, N. Picard, Stationarity of GARCH processes and of some nonnegative time séries, Journal of Econometrics 52 (1992) 115-127.

[8] T. Bollerslev, Generalized autoregressive conditional heteroskedasticity, Journal of econometrics 31 (1986), p.307-327.