REPUBLIQUE ALGERIENNE DEMOCRATIQUE et POPULAIRE.
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique.

UNIVERSITE MOULOUD MAMMERI, TIZI-OUZOU
Faculté des Sciences
Département de Mathématiques

Mémoire de Master
en
MATHEMATIQUES

Option
Modélisation Mathématique

Thème

Estimation des modèles ARCH : cas non stationnaire

Présenté par

LANNABI Rachida

Devant le jury d'examen composé de :

Fazia. KHELLAS	Professeur	UMMTO	Présidente
Abdelghani. HAMAZ	Maître de conférence B	UMMTO	Rapporteur
Farida. MERAKEB	Maître de conférence B	UMMTO	Examinatrice

Soutenue le 09 / 10 / 2014

Remerciements

Nous tenons tout d'abord à remercier Dieu le tout puissant et miséricordieux, qui nous a donné la force et la patience d'accomplir ce modeste travail.

En second lieu, nous tenons à remercier notre encadreur** M^r **Abdelghani HAMAZ, son précieux conseil et son aide durant toute la période du travail.

Nos vifs remerciements vont également aux membres du jury pour l'intérêt qu'ils ont porté à notre recherche en acceptant d'examiner notre travail et de l'enrichir par leurs propositions.

Enfin, nous tenons également à remercier tous nos enseignants et toutes les personnes qui ont participé de près ou de loin à la réalisation de ce travail.

Rachida.

Dédicaces

Je dédie ce travail à:

Mes trés chers parents qui sont sans cesse à mes côtés;

Mes très chers frères et mes très cheres soeurs qui m'ont toujours soutenue;

Toute ma famille;

Lydia avec laquel j'ai partagé cette tâche et toute sa famille;

Tous mes amis(es).

Rachida.

Table des matières

Introduction

Les modèles autorégressifs hétéroscédastiques (ARCH) ont été présentés par Engel dans l'article publié *Econometrica* au début des années 80 (Engle, 1982), en 2003, le prix Nobel pour les sciences économiques a été conjointement attribué à Robert F.Engle et Clive W.J. Par suite, des développement sont apportés pour les modèles ARCH, qui sont généralisés en 1986 (GARCH), ces modèles sont devenus extrêmement populaires parmi les académiques et les praticiens.

Les modèles GARCH ont mené à un changement fondamental à la modélisation des séries financières, et sont particulièrement annoncés pour prendre en compte les caractéristiques importantes de ces séries (stationnarité, volatilité, asymétrie, saisonnalité, $\cdots$). 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'insuffisance des modèles de type 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 extensions des modèles ARCH ont été édités, ces extensions ont provoqué des nouvelles directions pour la recherche statistique et probabilité.

L'objectif de ce travail est d'essayer de viser l'essentiel de la littérature statistique des modèles ARCH et GARCH, et de donner des démonstrations de certains des résultats théoriques. L'accent est mis sur l'étude des diverses méthodes d'estimation de ces modèles où la consistence et le comportement asymptotique des estimateurs sont étudiés.

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

L'inférence statistique fait l'objet de la deuxième partie, nous présentons les différentes méthodes d'estimation, commençant par la méthode des moindres carrés, ensuite la méthode de quasi-maximum de vraisemblance, et enfin la méthode des deux phases (Bose et Mukherjee, 2003). Le dernier chapitre est consacré à l'estimation des paramètres du modèle ARCH(1) par la méthode du quasi-maximum de vraisemblance dans le cas où la condition de stationnarité stricte n'est pas vérifiée, et nous étudions les propriétés asymptotiques de ces estimateurs. Ce travail est accompagné d'une étude par simulation à base de langage de programmation **R**, qui confirme les résultats asymptotiques pour α_0, et illustre également l'impossibilité d'estimer le paramètre ω_0 avec une précision raisonnable sous l'hypothèse de non stationnarité.

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

Chapitre 1

Processus conditionnellement hétéroscédastiques

Dans ce chapitre, on s'intéresse à la modélisation d'une série chronologique par un processus de type autorégressif conditionnellement Hétéroscédastique (ARCH), introduit pour la première fois par Engle en (1982), et généralisé par Bollerslev en (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 fructueuse car elle permet une étude complète des propriétés des solutions tout en étant assez générale. Les modèles GARCH sont en effet susceptibles de capter les propriétés caractéristiques des séries financières.
Nous présentons d'abord des définitions et des représentations des modèles ARCH et GARCH, nous établissons la condition de 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 Principales propriétés des séries financières

Les séries financières (rentabilités d'action, taux d'intérêt, taux de change, $\cdots$), généralement sont des séries de prix d'actif et de rendements, ou de façon proche les modifications de logarithme de prix entre $t-1$ et t. L'unité temporelle peut être le jour, la semaine ou le mois.
Soit p_t le prix d'un actif à la date t 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) :

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

où $s_t = \left(p_t - p_{t-1}\right)/p_{t-1}$ désigne la variation relative des prix.
Ces séries présentent des propriétés que nous allons présenter, la plupart sont discutées dans Taylor (1985) et cambbell (1990).

Propriétés 1.1.1. [La volatilité] *C'est la plus importante propriété, comme remarqué Mandelbrot (1963), est le fait que, "les grandes variations de prix tendent à être suivies de grandes variations", de signe quelconque, "et les petites variations tendent à être suivies de petites variations". En d'autres termes la volatilité (variance conditionnelle) évolue avec le temps .*

Propriétés 1.1.2. [La Stationnarité] *Les processus stochastiques 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 1.1.1. [La stationnarité stricte] Le processus $(X_t)_{t\in\mathbb{Z}}$ est dit strictement ou fortement stationnaire si $\forall\ t_1, t_2, \cdots, t_n, h \in \mathbb{Z}$, on a :

$$\mathcal{L}\left(X_{t_1}, X_{t_2}, \cdots, X_{t_n}\right) = \mathcal{L}\left(X_{t_{1+h}}, X_{t_{2+h}}, \cdots, X_{t_{n+h}}\right).$$

Définition 1.1.2. [La stationnarité faible] Un processus $(X_t)_{t\in\mathbb{Z}}$ est dit stationnaire ou faiblement stationnaire ou stationnaire au second ordre si :

1. $\mathbb{E}\left(X_t^2\right) < \infty$,
2. $\mathbb{E}\left(X_t\right) = \mu, \forall t \in \mathbb{Z}$ (ne dépend pas de temps t),
3. $\mathrm{cov}(X_t, X_{t+h}) = \mathbb{E}\left(\left(X_t - \mathbb{E}\left(X_t\right)\right)\left(X_{t+h} - \mathbb{E}\left(X_{t+h}\right)\right)\right) = \gamma(h), \forall t, h \in \mathbb{Z}$ (ne dépend pas de temps t).

Propriétés 1.1.3. [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 de rendements. Les valeurs négatives (baisses du cours) tendent à provoquer une augmentation de la volatilité supérieure 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 passés des cours ou rendements à la volatilité de ces derniers).*

Propriétés 1.1.4. [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 très volatiles à la fermeture (week-end et jours fériés), (la volatilité tend à augmenter lorsque les marchés ferment).*

Propriétés 1.1.5. [**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 supérieur à 3 pour ces séries.*

Définition 1.1.3. [**Kurtosis**] Le Kurtosis d'une variable aléatoire X correspond à son moment centré d'ordre 4, c'est à dire : $\mu_4 = \mathbb{E}[(X-\mu)^4]$.

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 :

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

- Si le Kurtosis > 3 (queues épaisses) la distribution est dite leptokurtique.
- Si le Kurtosis < 3, la distribution est dite platikurtique.

Propriétés 1.1.6. [**L'auto-corrélation**] *Les auto-corrélations de la série* (ε_t) *sont faibles, ce qui signifie que la série* (ε_t) *est proche d'un bruit blanc, par contre la série* (ε_t^2) *a des fortes auto-corrélations. Ce qui est incompatible avec une hypothèse de bruit blanc.*

Définition 1.1.4. [**Bruit blanc**] On dit que (Z_t) est un bruit blanc faible de moyenne nulle et de variance σ^2 noté $Z_t \sim \mathbf{BB}(0, \sigma^2)$ lorsque :

$cov(Z_t, Z_i) = 0$, $\forall t \neq i$, (absence d'autocorrélation des erreurs).

On dit que (Z_t) est un bruit blanc fort de moyenne nulle et de variance σ^2 lorsque :

$Z_t \perp Z_i \ \forall\ t \neq i$, (les variables Z_t et Z_i sont indépendantes).

1.2 Présentation du modèle ARCH et GARCH

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

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

Définition 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 $\omega > 0$, $\alpha_i \geq 0$, $i = 1, \cdots, q$, et B est l'opérateur retard tel que $\alpha(B) = \sum_{i=1}^{q} \alpha_i B^i$ avec $B^i \varepsilon_t^2 = \varepsilon_{t-i}^2$.

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

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

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

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

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

$$\Rightarrow$$

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

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

1.2.1 Propriétés d'un processus ARCH

Le processus ε_t ARCH défini par l'équation (1.1) possède les propriétés statistiques suivantes :

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

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

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

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

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

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

□

Propriétés 1.2.2. [La variance conditionnelle] *C'est la propriété centrale des processus ARCH. Le processus $\left(\varepsilon_t\right)_{t\in\mathbb{Z}}$ défini par l'équation (1.1) a une variance conditionnelle dépend du temps et vérifie :*

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

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

□

Propriétés 1.2.3. [Les auto-covariances conditionnelles] *Les auto-covariances conditionnelles du processus $\left(\varepsilon_t\right)_{t\in\mathbb{Z}}$ sont nulles. C'est à dire que :*

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

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

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

□

Propriétés 1.2.4. [Moment centré d'ordre quatre] *Les modèles ARCH permettent d'avoir des processus avec des queues de distribution plus épaisses. Le moment conditionnel centré d'ordre 4 du processus* $(\varepsilon_t)_{t\in\mathbb{Z}}$ *vérifie :*

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

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

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

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

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

Preuve. On a : $\varepsilon_t = \eta_t\sigma_t$,
et on sait que si une variable X centré et suit une loi normale alors : $\mathbb{E}\left(X^4\right) = 3\mathbb{E}\left(X^2\right)^2$,

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

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

□

Remarque 1.2.1. Le kurtosis d'un processus ARCH est toujours supérieur à 3, la loi non conditionnelle d'un processus ARCH est donc une loi de distribution à queue épaisse, est donc plus aplatie qu'une gaussienne, on dit que cette distribution est leptokurtique .

Remarque 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.

Nous avons vu que le processus ARCH décrit précédemment dans la définition (1.1) admet une représentation autorégressive (AR) de la variance conditionnelle $(\sigma_t^2)_t$, donnée par l'équation (1.2). De manière similaire aux processus ARMA, il semble naturel de généraliser l'expression (1.2) en permettant à $(\sigma_t^2)_t$ de présenter une partie moyenne-mobile. Bollerslev (1986) a donc introduit les processus GARCH(p,q).

Définition 1.2.2. [Processus GARCH(p,q) semi-fort] On dit que $(\varepsilon_t)_{t\in\mathbb{Z}}$ est un processus GARCH(p,q) si ses deux premiers moments conditionnels existent et vérifient :

1. $\mathbb{E}\left(\varepsilon_t|\varepsilon_s, s<t\right)=0, \quad t\in\mathbb{Z}$;
2. Il existe des constantes ω, α_i, $i=1,\cdots,q$ et β_j, $j=1,\cdots,p$ telles que :

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

L'équation (1.3) est équivalente à :

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

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

Le processus d'innovation pour ε_t^2 est par définition $\nu_t=\varepsilon_t^2-\mathbb{E}\left\{\varepsilon_t^2|\varepsilon_j,j\leq t-1\right\}=\varepsilon_t^2-\sigma_t^2$, qui vérifie $\mathbb{E}\left(\nu_t|\mathcal{F}_{t-1}\right)=0$, 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 :

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

$$\Rightarrow$$

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

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

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

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

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

$\Leftrightarrow$

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

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

1.2.2 Étude de la stationnarité

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

1.Cas d'un GARCH(1,1)

Dans le cas où p = q = 1, le modèle (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 :

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

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

Théorème 1.2.1. [La stationnarité stricte du modèle GARCH(1,1)] *Si*

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

la série

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

converge presque sûrement et le processus (ε_t) *défini par* $\varepsilon_t = \sqrt{h_t}\eta_t$ *est l'unique solution strictement stationnaire du modèle* (1.6), *cette solution est non anticipative et ergodique. Si* $\gamma \geq 0$ *et* $\omega > 0$, *il n'existe pas de solution strictement stationnaire.*

Remarque 1.2.3.

1. $\gamma = \mathbb{E}\left\{\log \alpha\left(\eta_t\right)\right\}$ existe toujours dans $[-\infty, +\infty[$, car $\mathbb{E}\left(\log^+\left\{\alpha\left(\eta_t\right)\right\}\right) \leq \mathbb{E}\left\{\alpha\left(\eta_t\right)\right\} = \alpha + \beta$.
2. Si $\alpha + \beta < 1 \Rightarrow \gamma < 0$, inversement si $\gamma < 0 \Rightarrow \beta < 1$,
   car si $\alpha + \beta < 1$, alors $\mathbb{E}\left\{\log\left(\alpha\eta_t^2 + \beta\right)\right\} \leq \log\left\{\mathbb{E}\left(\alpha\eta_t^2 + \beta\right)\right\} = \log\left(\alpha + \beta\right) < 0$.
   Inversement si $\gamma < 0$, alors par absurde supposons que $\beta \geq 1$ donc :
   $\mathbb{E}\left\{\log\left[\alpha\eta_t^2 + \beta\right]\right\} = \mathbb{E}\left\{\log\left[\beta\left(\frac{\alpha}{\beta}\eta_t^2 + 1\right)\right]\right\} = \mathbb{E}\left\{\log\left(\beta\right) + \log\left(\frac{\alpha}{\beta}\eta_t^2 + 1\right)\right\} > 0$, d'où la contradiction.
3. Dans le cas ARCH(1) ($\beta = 0$), la contrainte de stationnarité stricte peut s'écrire comme suit
   $$0 \leq \alpha < \exp\left\{-\mathbb{E}\left(\log \eta_t^2\right)\right\}, \tag{1.10}$$
   En effet,
   $$\begin{aligned} -\infty \leq \gamma &= \left\{\mathbb{E}\left(\log \alpha + \log \eta_t^2\right)\right\} < 0 \\ &\Leftrightarrow -\infty \leq \log \alpha < -\mathbb{E}\left(\log \eta_t^2\right) \\ &\Leftrightarrow 0 \leq \alpha < \exp\left\{-\mathbb{E}\left(\log \eta_t^2\right)\right\}. \end{aligned}$$
   Par exemple dans le cas où $\eta_t \sim \mathcal{N}(0,1)$ la condition est : $\alpha < 3.56$.
4. Dans le cas où $\omega = 0$ et $\gamma < 0$, il est clair d'après (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 $N \geq 1$:

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

Soit $h_t = \lim_{N \to +\infty} h_t\left(N\right) \in [0, +\infty]$, puisque les termes sont positifs.
De plus $h_t\left(N\right) = \omega + \alpha\left(\eta_{t-1}\right)h_{t-1}\left(N-1\right)$, donc quand $N \to +\infty$, on obtient :
$h_t = \omega + \alpha\left(\eta_{t-1}\right)h_{t-1}$.
Maintenant montrons que le processus limite (h_t) est à valeurs finies si et seulement si

$\gamma < 0$.
En utilisant la règle de Cauchy pour les séries à termes positifs, on a :

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

donc

$$\begin{aligned}
\left\{\alpha\left(\eta_{t-1}\right) \cdots \alpha\left(\eta_{t-n}\right)\right\}^{1/n} &= \exp\left\{\frac{1}{n}\log\left(\alpha\left(\eta_{t-1}\right)\cdots\alpha\left(\eta_{t-n}\right)\right)\right\} \\
&= \exp\left\{\frac{1}{n}\sum_{i=1}^{n}\log\left[\alpha\left(\eta_{t-i}\right)\right]\right\} \\
&\to \exp\left\{\mathbb{E}\left[\log\left(\alpha\eta_{t-i}^2\right)\right]\right\} = \exp\left(\gamma\right) \quad p.s. \qquad (1.12)
\end{aligned}$$

par l'application de la loi forte des grands nombres, quand $n \to +\infty$, à la suite *iid* $\left(\log\left\{\alpha\left(\eta_{t-i}\right)\right\}\right)$, la série h_t converge presque sûrement dans $\mathbb{R}$, par application de la règle de Cauchy, et le processus limite, (h_t), est à valeurs réelles positives. Par suite le processus (ε_t) défini par $\varepsilon_t = \sqrt{h_t}\eta_t$, est strictement stationnaire et ergodique (théorème d'ergodicité), et est non anticipatif car il s'écrit comme fonction mesurable des variables η_{t-i}, $i \geq 0$, de plus (ε_t) vérifie le modèle (1.6).
Montrons l'unicité :
Soit $\tilde{\varepsilon}_t = \sigma_t\eta_t$ une autre solution strictement stationnaire. D'après (1.11), on a :

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

Par suite

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

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

Si $\gamma > 0$, d'après (1.12) et la règle de Cauchy, $\sum_{n=1}^{N}\alpha\left(\eta_{t-1}\right)\cdots\alpha\left(\eta_{t-n}\right) \to +\infty$, $p.s$ 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 $n > 0$,

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

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

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

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

□

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 précisément (ε_t) est un bruit blanc.*
Si $\alpha+\beta\geq 1$, il n'existe pas de solution non anticipative stationnaire au second ordre.

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

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

soit

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

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

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

□

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

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

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

où
$Z_t^{'} = \left(\varepsilon_t^2, \varepsilon_{t-1}^2, \cdots, \varepsilon_{t-q+1}^2, \sigma_t^2, \cdots, \sigma_{t-p+1}^2\right) \in \mathbb{R}^{p+q}$,
$b_t^{'} = (\omega\eta_t^2, 0, \cdots, 0, \omega, 0, \cdots, 0) \in \mathbb{R}^{p+q}$,

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

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 déroule le modèle (1.14) on obtient :

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

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,

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

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

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

La stationnarité stricte

L'outil principale pour l'étude de la stationnarité stricte est le concept d'exposant de Lyapounov. Soit A une matrice $(p+q) \times (q+q)$. Son rayon spectral, noté $\rho(A)$, est le plus grand module de ses valeurs propres. Soit $\| \, . \, \|$ une norme quelconque sur l'espace des matrices $(p+q) \times (q+q)$. 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 Lyapounov] *Soit $\{A_t\}_{t\in\mathbb{Z}}$ une suite de matrices aléatoires, strictement stationnaire et ergodique, telle que $\mathbb{E}\log^+ \|A_t\| < \infty$. On a :*

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

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

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

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

Remarque 1.2.4.

1. On a toujours $\gamma \leq \mathbb{E}\left(\log \|A_1\|\right)$, 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 $\mathbb{E}\left[\log^+ \| A_t \|\right] < \infty$, et de plus grand exposant de Lyapounov γ. Alors :*

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

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

Théorème 1.2.4. [La stationnarité stricte du modèle GARCH(p,q)] *Le modèle GARCH (p,q) a une unique solution strictement stationnaire non anticipative et ergodique si et seulement si $\gamma < 0$, où γ est le plus grand exposant de Lyapounov de la suite $\{A_t\}_{t\in\mathbb{Z}}$.*

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

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

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

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

et

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

Par application de la règle de Cauchy, $\gamma < 0$ implique Z_t est bien défini. Soit $Z_{q+1,t}$ la q+1-ème composante de Z_t. En posant $\varepsilon_t = \sqrt{Z_{q+1,t}}\eta_t$ on définit une solution strictement stationnaire du modèle (1.5). D'après (1.16) ε_t s'exprime comme fonction mesurable de η_t, $\eta_{t-1}, \cdots$. 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 $n \geq 0$,

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

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

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

par suite

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

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

Finalement Nous montrons la partie nécessaire du théorème. D'après le lemme (1.2.1), il suffit d'établir (1.22). Nous allons montrer que, pour $1 \leq i \leq p+q$

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

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 $t > 0$

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

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

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

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

$$A_{-k} e_{q+i} = \beta_i \eta^2_{-k} e_1 + \beta_i e_{q+1} + e_{q+i+1}, \quad i = 1, \cdots, p \tag{1.26}$$

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

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

donc (1.24) est vraie pour $i = q+2$, et par récurrence, pour $i = q+j$, $j = 1, \cdots, p$, en utilisant (1.26). Par ailleurs, on remarque que $A_{-k} e_q = \alpha_q \eta^2_{-k} e_1 + \alpha_q e_{q+1}$, 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

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

et une récurrence ascendante. Le théorème (1.2.4) est donc démontré.

□

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

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

Preuve. Comme tous les termes des matrices A_t sont positifs, il est clair que γ est supérieur au coefficient de Lyapounov de la suite obtenue en remplaçant les coefficients des q premieres lignes et des q premieres colonnes par 0 dans les matrices A_t. En utilisant la Remarque 2 du Théorème (1.2.3) 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 $\lambda \neq 0$,

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

où $\mathcal{B}(z) = 1 - \beta_1 z - \cdots - \beta_p z^p$. On en déduit que si $\gamma < 0$ alors $\mathcal{B}(z) = 0$ a toutes ses racines en dehors du cercle unité, d'où l'équivalence entre (ii) et (iii).
A présent, montrons que $(i) \Leftrightarrow (ii)$. On a $\mathcal{B}(0) = 1$ et $\mathcal{B}(1) = 1 - \sum_{j=1}^{p} \beta_j$, donc si $\sum_{j=1}^{p} \beta_j \geq 1$, alors $\mathcal{B}(1) \leq 0$ et, par continuité, il existe une racine dans]0,1]. Ainsi $(ii) \Rightarrow (i)$
Inversement si $\sum_{j=1}^{p} \beta_j < 1$ et si $\mathcal{B}(z_0) = 0$ pour un z_0 de module inférieur ou égal à 1 alors $1 = \sum_{j=1}^{p} \beta_j z_0^j = \left|\sum_{j=1}^{p} \beta_j z_0^j\right| \leq \sum_{j=1}^{p} \beta_j |z_0|^j \leq \sum_{j=1}^{p} \beta_j$, ce qui est impossible, par suite $(i) \Rightarrow (ii)$.

$\square$

1.3 Existence des moments d'ordre 2s

Nous concluons cette partie avec un résultat établissant que la condition de stationnarité stricte implique également l'existence de certains moments. Nous donnons au préalable 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 $g(v) = \dfrac{e^v - 1}{v}$ 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^-} integrable 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 $M(s) = \mathbb{E}(X^s) < 1$.

$\square$

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

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

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

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

en utilisant la norme multiplicative $\|A\| = \sum_{i,j} |A(i,j)|$, la positivité des éléments des A_i, l'indépendance et l'équidistribution des A_i. Le lemme (1.3.1) 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

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

$\square$

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

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

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 utilisant le fait que la norme est multiplicative. La solution stationnaire est définie par (1.16) et 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 $\varepsilon_t^{2s} \leq \|Z_t\|^s$.

$\square$

1.4 Extension des Modèles ARCH et GARCH

Les processus ARCH ont donné lieu à de nombreuses extensions dans la littérature statistique. Dans cette section, nous présentons quelques processus de type GARCH développés qui sont très utilisés dans le domaine de la finance.

1.4.1 Processus ARCH-M

Le processus ARCH-M (ARCH in Mean), proposé par Engle, Lilien et Robins (1987) 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 :

$$\begin{cases} Y_t = X_t\beta + \delta g\left(\sigma_t^2\right) + \varepsilon_t \\ \sigma_t^2 = \omega + \alpha\varepsilon_{t-1}^2 \end{cases} \tag{1.29}$$

où $(X_t)_t$ 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 $Y_t \mid \mathcal{F}_{t-1} \longrightarrow \mathcal{N}\left(\mu_t, \sigma_t^2\right)$, 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.

1.4.2 Modèles ARCH / GARCH asymétriques

Cette partie couvre les modèles ARCH non linéaires et plus particulièrement la prise en compte des phénomènes asymétries. Deux grandes classes de modèles ont été proposés :

Les modèles EGARCH

Proposé par Nelson (1991), le processus Exponential GARCH ou EGARCH(p,q) donne à la variance conditionnelle la définition suivante :

Définition 1.4.1. Un processus $(\varepsilon_t)_{t\in\mathbb{Z}}$ satisfait une representation EGARCH(p,q) si et seulement si :

$$\begin{cases} \varepsilon_t = \sigma_t\eta_t \\ \log\left(\sigma_t^2\right) = \omega + \sum_{i=1}^{q} \alpha_i g\left(\eta_{t-i}\right) + \sum_{j=1}^{p} \beta_j \log\left(\sigma_{t-j}^2\right) \end{cases} \tag{1.30}$$

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

$$g\left(\eta_{t-i}\right) = \theta\eta_{t-i} + \gamma\left(\left|\eta_{t-i}\right| - \mathbb{E}\left(\left|\eta_{t-i}\right|\right)\right), \tag{1.31}$$

où θ, γ, α_i et β_j sont des réels.

Si on remplace la fonction g dans l'equation (1.30), si on pose $a_i = \alpha_i\theta$ et $b_i = \alpha_i\gamma$, la variance conditionnelle de ε_t peut se réécrire sous la forme :

$$\log\left(\sigma_t^2\right) = \omega + \sum_{i=1}^{q} a_i\eta_{t-i} + \sum_{i=1}^{q} b_i\left(\left|\eta_{t-i}\right| - \mathbb{E}\left(\left|\eta_{t-i}\right|\right)\right) + \sum_{j=1}^{p} \beta_j \log\left(\sigma_{t-j}^2\right). \tag{1.32}$$

1.4.3 Modèles GARCH à seuil (TGARCH)

Les modèles TARCH sont définis par Zakoïan (1991) et TGARCH en (1994).Une façon naturelle d'introduire l'asymétrie est de spécifier la variance conditionnelle en fonction des composantes positive et négative des innovations passées.
Notons $\varepsilon_t^+ = \max(\varepsilon_t, 0)$, $\varepsilon_t^- = \min(\varepsilon_t, 0)$, en remarquant que $\varepsilon_t = \varepsilon_t^+ + \varepsilon_t^-$.

Définition 1.4.2. Un processus $(\varepsilon_t)_{t\in\mathbb{Z}}$ satisfait une representation GARCH(p,q) à seuil (Threshold GARCH(p,q)) si et seulement si :

$$\begin{cases} \varepsilon_t = \sigma_t \eta_t \\ \sigma_t = \omega + \sum_{i=1}^{q} \alpha_{i,+}\mathbb{I}_{\varepsilon_{t-i}\geq 0}\varepsilon_{t-i} + \alpha_{i,-}\mathbb{I}_{\varepsilon_{t-i}<0}\varepsilon_{t-i} + \sum_{j=1}^{p} \beta_j \sigma_{t-j} \end{cases} \tag{1.33}$$

où ω, $\alpha_{i,+}$, $\alpha_{i,-}$ et β_j sont des réels, et (η_t) une suite de variables *iid* telles que $\mathbb{E}(\eta_t) = 0$ et $\mathbf{V}(\eta_t) = 1$.
$\mathbb{I}_{\varepsilon_{t-i}<0}$ 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 ε_t.

Remarque 1.4.2. Il existe de nombreuses autres extensions des processus ARCH. On citera, entre autres, le processus IGARCH(p,q) (GARCH intégré), introduit et étudié par Engel 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 citera encore les processus GJR-ARCH et GJR-GARCH (Glosten, Jagannathan et Runkle, 1993), et Q-GARCH (Q pour Quadratic) 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 (ε_t) est un ARCH(∞) s'il existe une suite des variables (η_t) *iid* tels que $\mathbb{E}(\eta_t) = 0$ et $\mathbb{E}(\eta_t^2) = 1$ et une suite de constante $\phi_i \geq 0$ $i = 1, \cdots$, 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

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*

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

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 :*

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

Le processus (ε_t) défini par (1.35) est l'unique solution strictement stationnaire et non anticipative de modèle (1.34) tels que $\mathbb{E}\,|\varepsilon_t|^{2s} < \infty$.

Preuve. Soit S_t la variable aléatoire définie par :

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

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

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

L'indépendances de η_t nous assure

$$\begin{aligned} \mathbb{E}\left(S_t^s\right) &\leq \phi_0^s + \phi_0^s \sum_{k=1}^{\infty} \sum_{i_1, i_2, \cdots, i_k \geq 1} \phi_{i_1}^s \cdots \phi_{i_k}^s \mathbb{E}\left(\eta_{t-i_1}^{2s} \cdots \eta_{t-i_1-\cdots-i_k}^{2s}\right) \\ &\leq \phi_0^s \left[1 + \sum_{k=1}^{\infty} (A_s \mu_{2s})^k\right] = \frac{\phi_0^s}{1 - A_s \mu_{2s}}, \end{aligned}$$

ce qui montre que $S_t < \infty$ p.s.

Tous les termes de la somme étant positifs, on a :

$$\begin{aligned}
\sum_{i=1}^{\infty}\phi_i S_{t-i}\eta_{t-i}^2 &= \sum_{i_0=1}^{\infty}\phi_{i_0}\eta_{t-i_0}^2\left(\phi_0+\phi_0\sum_{k=1}^{\infty}\sum_{i_1,i_2,\cdots,i_k\geq 1}\phi_{i_1}\cdots\phi_{i_k}\eta_{t-i_0-i_1}^2\cdots\eta_{t-i_0-i_1-\cdots-i_k}^2\right)\\
&= \phi_0\sum_{i_0=1}^{\infty}\phi_{i_0}\eta_{t-i_0}^2+\phi_0\sum_{i_0=1}^{\infty}\phi_{i_0}\eta_{t-i_0}^2\sum_{k=1}^{\infty}\sum_{i_1,i_2,\cdots,i_k\geq 1}\phi_{i_1}\cdots\phi_{i_k}\eta_{t-i_0-i_1}^2\cdots\eta_{t-i_0-i_1-\cdots-i_k}^2\\
&= \phi_0\sum_{k=0}^{\infty}\sum_{i_0,i_1,\cdots,i_k\geq 1}\phi_{i_0}\cdots\phi_{i_k}\eta_{t-i_0}^2\eta_{t-i_0-i_1}^2\cdots\eta_{t-i_0-i_1-\cdots-i_k}^2
\end{aligned}$$

Par conséquent, $S_t = \phi_0 + \sum_{i=1}^{\infty} S_{t-i}\eta_{t-i}^2$.

La solution strictement stationnaire de modèle (1.34) est obtenue par $\varepsilon_t = (S_t)^{1/2}\eta_t$, de plus $\mathbb{E}|\varepsilon_t|^{2s} \leq \dfrac{\mu_{2s}\phi_0^s}{1-A_s\mu_{2s}}$.

Pour la deuxième partie du théorème, on a :

$$\begin{aligned}
\mathbb{E}|\varepsilon_t|^{2s} &= \mathbb{E}|(S_t)^{1/2}\eta_t|^{2s} \leq \mathbb{E}|S_t^s|\mathbb{E}|\eta_t^{2s}|\\
&\leq \mu_{2s}\mathbb{E}|S_t^s| \leq \frac{\mu_{2s}\phi_0^s}{1-A_s\mu_{2s}}.
\end{aligned}$$

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 :

$$
\begin{aligned}
\sigma_t^2 &= \phi_0 + \sum_{i=1}^{q} \phi_i \varepsilon_{t-i}^2 \\
&= \phi_0 + \sum_{i=1}^{q} \phi_i \eta_{t-i}^2 \sigma_{t-i}^2 \\
&= \phi_0 + \sum_{i_1=1}^{q} \phi_{i_1} \eta_{t-i_1}^2 \left(\phi_0 + \sum_{i_2=1}^{q} \phi_{i_2} \eta_{t-i_1-i_2}^2 \sigma_{t-i_1-i_2}^2 \right) \\
&= \phi_0 + \phi_0 \sum_{i_1=1}^{q} \phi_{i_1} \eta_{t-i_1}^2 + \sum_{i_1=1}^{q} \phi_{i_1} \eta_{t-i_1}^2 \sum_{i_2=1}^{q} \phi_{i_2} \eta_{t-i_1-i_2}^2 \sigma_{t-i_1-i_2}^2 \\
&= \phi_0 + \phi_0 \sum_{i_1=1}^{q} \phi_{i_1} \eta_{t-i_1}^2 + \sum_{i_1=1}^{q} \phi_{i_1} \eta_{t-i_1}^2 \sum_{i_2=1}^{q} \phi_{i_2} \eta_{t-i_1-i_2}^2 \left(\phi_0 + \sum_{i_3=1}^{q} \phi_{i_3} \eta_{t-i_1-i_2-i_3}^2 \sigma_{t-i_1-i_2-i_3}^2 \right) \\
&= \phi_0 + \phi_0 \sum_{i_1=1}^{q} \phi_{i_1} \eta_{t-i_1}^2 + \phi_0 \sum_{i_1=1}^{q} \phi_{i_1} \eta_{t-i_1}^2 \sum_{i_2=1}^{q} \phi_{i_2} \eta_{t-i_1-i_2}^2 \\
&+ \sum_{i_1=1}^{q} \phi_{i_1} \eta_{t-i_1}^2 \sum_{i_2=1}^{q} \phi_{i_2} \eta_{t-i_1-i_2}^2 \sum_{i_3=1}^{q} \phi_{i_3} \eta_{t-i_1-i_2-i_3}^2 \sigma_{t-i_1-i_2-i_3}^2 \\
&\vdots \\
&= \phi_0 + \phi_0 \sum_{k=1}^{q} \sum_{i_1, i_2, \cdots, i_k \geq 1} \phi_{i_1} \phi_{i_2} \cdots \phi_{i_k} \eta_{t-i_1} \cdots \eta_{t-i_1-i_2-\cdots-i_k} \\
&+ \sum_{i_1, i_2, \cdots, i_{q+1} \geq 1} \phi_{i_1} \phi_{i_2} \cdots \phi_{i_{q+1}} \eta_{t-i_1} \cdots \eta_{t-i_1-i_2-\cdots-i_q} \varepsilon_{t-i_1-i_2-\cdots-i_{q+1}} \\
&= S_{t,q} + R_{t,q}.
\end{aligned}
$$

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

$$
\begin{aligned}
\mathbb{E}\left(R_{t,q}^s\right) &\leq \sum_{i_1, i_2, \cdots, i_{q+1} \geq 1} \phi_{i_1}^s \phi_{i_2}^s \cdots \phi_{i_{q+1}}^s \mathbb{E}\left(|\eta_{t-i_1}|^{2s} \cdots |\eta_{t-i_1-i_2-\cdots-i_q}|^{2s} |\varepsilon_{t-i_1-i_2-\cdots-i_{q+1}}|^{2s} \right) \\
&\leq \left(A_s \mu_{2s}\right)^q A_s \mathbb{E}|\varepsilon_t|^{2s} < \infty.
\end{aligned}
$$

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

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 fonction affine de son passé, sous les condition $\omega > 0$, α_i, $\beta_j \geq 0$, $i = \overline{1,q}$, $j = \overline{1,p}$.
Représentation ARCH(∞) d'un GARCH(1,1).

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

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

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 de la stationnarité stricte peut se réécrire sous la forme :

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

Preuve.

1. $A_s \mu_{2s} = \dfrac{\alpha^s \mu_{2s}}{1-\beta^s} < 1 \Leftrightarrow \alpha^s \mu_{2s} < 1 - \beta^s \Leftrightarrow \alpha^s \mu_{2s} + \beta^s < 1.$

2. En itérant dans l'équation caractérisait le GARCH (1,1). Nécessairement si (ε_t) représente la solution stationnaire et non anticipative du modèle GARCH(1,1), alors pour $q \geq 1$

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

D'après le corollaire (1.3.1), il existe $s \in]0,1[$ tel que $\mathbb{E}\left(\sigma_t^{2s}\right) = c < \infty$, alors
$\sum_{q \geq 1} \mathbb{E} \left(\beta^q \sigma_{t-q}^2 \right)^s = \sum_{q \geq 1} \beta^{qs} \mathbb{E} \left(\sigma_{t-q}^{2s} \right) = \dfrac{c\beta^s}{1-\beta^s} < \infty.$

$\beta^q \sigma_{t-q}^2 \longrightarrow 0$ p.s. quand $q \longrightarrow \infty$.

D'où on obtient
$\sigma_t^2 = \dfrac{\omega}{1-\beta} + \alpha \sum_{i=1}^{\infty} \beta^{i-1}\varepsilon_{t-i}^2.$

□

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

Théorème 1.5.3. [Représentation ARCH(∞) d'un GARCH(p,q)] *Si* (ε_t) *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 $\underline{\sigma}_t^2 = (\sigma_t^2, \cdots, \sigma_{t-p+1}^2)'$, $\underline{c}_t = (\omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2, 0, \cdots, 0)'$ et B une matrice définie par

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

Le corollaire (1.2.1), 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 (1.3.1).En itérant (1.39) on obtient :

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

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

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

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

$\square$

Chapitre 2

Estimation des modèles ARCH

Il existe plusieurs méthodes pour l'estimation des modèles ARCH. Dans ce chapitre nous présentons les trois suivantes : la méthode de quasi-maximum de vraisemblance, les moindres carrés, et la méthode des deux phases. Ainsi, nous établissons les propriétés asymptotiques (convergence presque sûr et normalité asymptotique) de ces estimateurs en imposant des conditions sur les moments.

2.1 Méthode des Moindres Carrés Ordinaires (M.C.O)

Nous considérons l'estimation par les moindres carrés ordinaires (M.C.O) du modèle ARCH(q) par :

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

avec $\omega_0 > 0$; $\alpha_{0i} \geq 0$; $1 \leq i \leq q$,

et où (η_t) est une suite de variables *iid*, $E(\eta_t) = 0$, $V(\eta_t) = 1$.

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

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

$$\varepsilon_t^2 = \omega_0 + \sum_{i=1}^{q} \alpha_{0i} \varepsilon_{t-i}^2 + \mathrm{u}_t, \tag{2.2}$$

où $\mathrm{u}_t = \varepsilon_t^2 - \sigma_t^2 = (\eta_t^2 - 1)\sigma_t^2$.

La suite $(\mathrm{u}_t, \mathcal{F}_{t-1})_t$ constitue donc une différence de martingale.

On suppose que l'on dispose d'observations $\varepsilon_1, \cdots, \varepsilon_n$, réalisations partielles du processus

(ε_t), et de valeurs initiales $\varepsilon_0, \cdots, \varepsilon_{1-q}$. Par exemple ces valeurs initiales peuvent être choisies nulles. Soit le vecteur

$$\mathrm{Z}'_{t-1} = \left(1, \varepsilon^2_{t-1}, \cdots, \varepsilon^2_{t-q}\right),$$

on déduit de (2.2) le système

$$\varepsilon^2_t = \mathrm{Z}'_{t-1}\theta_0 + \mathrm{u_t}, \quad t = 1, \cdots, n, \tag{2.3}$$

soit

$$Y = X\theta_0 + \mathrm{U},$$

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

$$X = \begin{pmatrix} \mathrm{Z}'_{n-1} \\ \vdots \\ \mathrm{Z}'_0 \end{pmatrix}, Y = \begin{pmatrix} \varepsilon^2_n \\ \vdots \\ \varepsilon^2_1 \end{pmatrix}, \mathrm{U} = \begin{pmatrix} u_n \\ \vdots \\ u_1 \end{pmatrix}.$$

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 :

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

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 : $\mathbb{E}_{\theta_0}(\varepsilon^4_t) < \infty$.
H3 : $\mathrm{P}[\eta^2_t = 1] \neq 1$.

Théorème 2.1.1. [Convergence des estimateurs M.C.O pour un ARCH] *Soit* $(\hat{\theta}_n)$ *une suite d'estimateurs satisfaisant* (2.4). *Sous les hypothèses H1-H3,*

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

Preuve. La preuve comporte plusieurs étapes.
i) Nous avons vu (Théorème de la stationnarité stricte du modèle GARCH(p,q)) que l'unique solution stationnaire non anticipative (ε_t) est ergodique. Le processus (Z_t) est

également ergodique car (Z_t) s'écrit comme fonction mesurable des ε_{t-i}. Le théorème ergodique appliqué au processus strictement stationnaire (Z_t) entraîne

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

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

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

ii) Montrons par l'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 non 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 variables $\varepsilon_t^2, \cdots, \varepsilon_{t-q+1}^2$.
On peut supposer sans perte de généralité que, dans cette combinaison, le coefficient de $\varepsilon_t^2 = \eta_t^2\sigma_t^2$ est 1. Donc η_t s'exprime $p.s$ comme fonction mesurable des variables $\varepsilon_{t-1}, \cdots, \varepsilon_{t-q}$.
Or, d'après le caractère non anticipatif de la solution, η_t^2 est indépendante de ces variables. 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.
iii) Il découle de ce qui précède que $\frac{1}{n}X'X$ est $p.s$ inversible, pour n assez grand et que $p.s.$ quand $n \to \infty$,

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

iv) Rappelons que le processus $\mathrm{u_t}$ est l'innovation forte de (ε_t^2). On a donc, en particulier, les relations d'orthogonalité

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

c'est-à-dire

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

d'où l'on déduit, en utilisant (2.3),

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

D'après *ii*) et *iii*), $\hat{\theta}_n$ 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 $\mathbf{A}$ a été établie dans la preuve du Théorème (2.1.1), celle de $\mathbf{I}$ 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* $\mathbf{H1} - \mathbf{H4}$*,*

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

Preuve. On a, d'après (2.3)

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

d'où

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

Soit $\lambda \in \mathbb{R}^{q+1}$, $\lambda \neq 0$. La suite $(\lambda' Z_{t-1}\mathrm{u_t}, \mathcal{F}_t)$ est une différence de martingale stationnaire, ergodique et de carré intégrable ($\mathbb{E}_{\theta_0}\left(\lambda' Z_{t-1}\mathrm{u_t} \mid \mathcal{F}_{t-1}\right) = 0$), de variance
$\mathbf{V}_{\theta_0}\left(\lambda' Z_{t-1}\mathrm{u_t}\right) = \lambda'\mathbb{E}_{\theta_0}\left\{Z_{t-1}Z'_{t-1}\mathrm{u_t}^2\right\}\lambda = \lambda'\mathbb{E}_{\theta_0}\left\{Z_{t-1}Z'_{t-1}(\eta_t^2-1)^2\sigma_t^4\right\}\lambda = (\mu_4-1)\lambda'\mathbf{I}\lambda$.
Par application d'un T.C.L pour différence de martingale stationnaire, on en déduit que, pour tout $\lambda \neq 0$

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

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

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

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 un raisonnement classique, que $\sqrt{n}(\hat{\theta}_n - \theta_0)$ est asymptotiquement normal, de moyenne le vecteur nul, et de variance la matrice du théorème.

□

2.2 Méthode du maximum de vraisemblance

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

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

où $\omega > 0$, $\alpha_i \geq 0$, $i = 1, \cdots, q$, $\beta_j \geq 0$, $j = 1, \cdots, p$, (η_t) est une suite de variables *iid*.
Les ordres p et q sont supposés connus. Afin de définir l'estimateur du Q.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 quasi-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 $\varepsilon_0, \cdots, \varepsilon_{1-q}, \tilde{\sigma}_0^2, \cdots, \tilde{\sigma}_{1-p}^2$ que nous allons préciser, la vraisemblance conditionnelle gaussienne $L_n(\theta)$ s'écrit :

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

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

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

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

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

On peut prendre comme valeurs initiales

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

ou encore

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

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

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

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

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

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

$$\hat{\theta}_n = \arg\min_{\theta \in \Theta} \tilde{\mathbf{I}}_n(\theta). \tag{2.16}$$

Équations de vraisemblance

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

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

avec

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

Ces équations s'interprètent, pour n grand, comme des relations d'orthogonalité.
En effet, comme nous le verrons plus précisément dans la partie suivante, le terme de gauche de l'égalité précédente se comporte asymptotiquement comme

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

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'espérance 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 $\dfrac{1}{\sigma_t^4(\theta_0)}\dfrac{\partial \sigma_t^2(\theta_0)}{\partial \theta}$ 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 ergodique.

2.2.1 Propriétés asymptotiques de l'estimateur du Q.M.V

La convergence forte

On a déjà vue dans le premier chapitre que le modèle (2.8) possède une solution stationnaire au sens stricte si et seulement si la suite de matrices $\mathbf{A}_0 = (A_{0t})$ admet un coefficient de Lyapounov strictement négatif $\gamma(\mathbf{A}_0) < 0$, 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

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

Notons

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

a toutes ses racines à l'extérieur du cercle unité, et $\alpha_i \geq 0$, $1 \leq i \leq q$.

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

H1 : $\theta_0 \in \Theta$, et Θ est compact.

H2 : $\gamma(A_0) < 0$ et $\forall\theta \in \Theta$, $\sum_{j=1}^{p} \beta_j < 1$.

H3 : η_t^2 a une loi non dégénérée et $\mathbb{E}(\eta_t^2) = 1$.

H4 : si $p > 0$, $\mathcal{A}_\theta(z)$ et $\mathcal{B}_\theta(z)$ n'ont pas de racine commune, $\mathcal{A}_\theta(1) \neq 0$, et $\alpha_{0q} + \beta_{0p} \neq 0$.

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

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

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 $0 < \rho_2 < 1$, $i_1 \geq 0$, $i_2 \geq 0$,

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

Schema de la preuve.

La demonstration repose sur une representation vectorielle autorégressive d'ordre un 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 l'espérance de $\ell_t(\theta_0)$ est bien définie dans $\mathbb{R}$ 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 fait que $\mathbb{E}_{\theta_0}(\ell_t^+(\theta))$ peut être égal à $+\infty$. Les hypothèses **H3** et **H4** sont cruciales pour établir l'identifiabilité : la première exclut l'existence d'une combinaison linéaire constante entre les ε_{t-j}^2, $j \geq 0$. On utilise également l'hypothèse d'absence de racines communes. L'ergodicité de $\ell_t(\theta)$ et un argument de compacité permettent de conclure.
Il sera commode de réécrire l'équation

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_{0j} \sigma_{t-j}^2, \quad \forall t, \tag{2.21}$$

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 :
a) $\lim_{n\to\infty} \sup_{\theta\in\Theta} |\mathbf{I}_n(\theta) - \tilde{\mathbf{I}}_n(\theta)| = 0, \quad p.s.$
b) $(\exists\ t \in \mathbb{Z}$ tel que $\sigma_t^2(\theta) = \sigma_t^2(\theta_0)\ \mathbf{P}_{\theta_0}\ p.s.\ ) \Longrightarrow \theta = \theta_0$.
c) $\mathbb{E}_{\theta_0}|\ell_t(\theta_0)| < \infty$ si $\theta = \theta_0$, et si $\theta \neq \theta_0$, $\mathbb{E}_{\theta_0}(\ell_t(\theta)) > \mathbb{E}_{\theta_0}(\ell_t(\theta_0))$.
d) Pour tout $\theta \neq \theta_0$ il existe un voisinage $\mathbf{V}(\theta)$ tel que $\liminf_{n\to\infty} \inf_{\theta^*\in\mathbf{V}(\theta)} \tilde{\mathbf{I}}_n(\theta^*) > \mathbb{E}_{\theta_0}(\ell_1(\theta_0))$, $p.s.$

a) Oubli asymptotique des valeurs initiales.

D'après le corollaire (1.2.1), 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 :

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

Soit $\underline{\tilde{\sigma}}_t^2$ le vecteur obtenu en remplaçant σ_{t-i}^2 par $\tilde{\sigma}_{t-i}^2$ dans $\underline{\sigma}_t^2$, et soit $\underline{\tilde{c}}_t$ le vecteur obtenu en remplaçant $\varepsilon_0^2, \cdots, \varepsilon_{1-q}^2$ par les valeurs initiales (2.13) ou (2.14). Nous avons

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

On déduit de (2.24) que presque sûrement

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

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),

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

L'existence d'un moment d'ordre $s > 0$ pour ε_t^2 , donné par le corollaire (1.3.1), et en utilisant le lemme de Borel–Cantelli, et l'inégalité de markov. Soit $\delta > 0$,

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

alors la série de terme générale $\mathbb{P}(\rho^t\varepsilon_t^2 > \delta)$ est convergente. Ce qui permet d'affirmer que $\rho^t\varepsilon_t^2 \to 0$ $p.s$. Par suite, en utilisant le lemme de Cesàro, on en déduit a).

b) Identifiabilité du paramètre.

Supposons que $\sigma_t^2(\theta) = \sigma_t^2(\theta_0)$ $\mathbf{P}_{\theta_0}$ *p.s.* Le polynôme $\mathcal{B}_\theta(B)$ est inversible sous l'hypothèse **H2**, par le corollaire (1.2.1). En utilisant (2.21), on obtient :

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

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 **H3**,
$\varepsilon_t^2 - \mathbb{E}_{\theta_0}(\varepsilon_t^2 \mid \varepsilon_{t-1}^2, \cdots) = \sigma_t^2(\theta_0)(\eta_t^2 - 1) \neq 0$, 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 d'absence de racines communes **H4**, ceci entraîne $\mathcal{A}_\theta(z) = \mathcal{A}_{\theta_0}(z)$, $\mathcal{B}_\theta(z) = \mathcal{B}_{\theta_0}(z)$ et $\omega = \omega_0$. On a donc montré b).

c) 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 $\mathbb{E}_{\theta_0}(\mathbf{I}_n(\theta)) = \mathbb{E}_{\theta_0}(\ell_t(\theta))$ 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 $\mathbb{E}_{\theta_0}(\ell_t^+(\theta_0)) < \infty$. Utilisant l'inégalité de Jensen et, à nouveau, 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

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

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 :

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

avec égalité si et seulement si $\sigma_t^2(\theta_0)/\sigma_t^2(\theta) = 1$ $\mathbf{P}_{\theta_0}$- *p.s.*, c'est-à-dire, étant donné b), si et seulement si $\theta = \theta_0$.

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

Pour tout $\theta \in \Theta$ et tout entier positif k, soit $\mathbf{V}_k(\theta)$ la boule ouverte de centre θ et de rayon $1/k$. En raison de a),

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

Pour obtenir la convergence de cette moyenne empirique on ne peut utiliser le théorème ergodique standard car nous avons vu que $\ell_t(\theta^*)$ 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 $\left\{\inf_{\theta^*\in\mathbf{V}_k(\theta)\cap\Theta} \ell_t(\theta^*)\right\}$, permet d'affirmer que

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

D'après le théorème de Beppo-Levi, $\mathbb{E}_{\theta_0}\left(\inf_{\theta^*\in\mathbf{V}_k(\theta)\cap\Theta} \ell_1(\theta^*)\right)$ tend en croissant vers $\mathbb{E}_{\theta_0}(\ell_1(\theta))$ quand $k \to \infty$. Étant donné (2.31), nous avons montré d).

La fin de la preuve du théorème utilise un argument de compacité. Remarquons d'abord que pour tout voisinage $\mathbf{V}(\theta_0)$ de θ_0

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

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

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

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

□

Normalité asymptotique

On considère les hypothèses supplémentaires suivantes :
H5 : $\theta_0 \in \mathring{\Theta}$, où $\mathring{\Theta}$ est l'intérieur de Θ.
H6 : $\kappa_\eta = \mathbb{E}(\eta_t^4) < \infty$.
La loi limite de $\hat{\theta}_n$ est donnée par le résultat suivant :
Théorème 2.2.2. [Normalité asymptotique des estimateurs du Q.M.V] *Sous les hypothèses* **H1 − H6**,

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

où

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

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 :

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

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

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

et que

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

La preuve du théorème en découlera immédiatement. Nous allons a nouveau décomposer la démonstration en plusieurs points.

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

b) J est inversible et $\mathbf{V}_{\theta_0}\left\{\frac{\partial \ell_t(\theta_0)}{\partial\theta}\right\}=(\kappa_\eta-1)J.$

c) Il existe un voisinage $\mathbf{V}(\theta_0)$ de θ_0 tel que, pour tous i, j, $k\in\{1,\cdots,p+q+1\}$,

$$\mathbb{E}_{\theta_0}\sup_{\theta\in\mathbf{V}(\theta_0)}\left|\frac{\partial^3\ell_t(\theta)}{\partial\theta_i\partial\theta_j\partial\theta_k}\right|<\infty.$$

d) $\left\|n^{-1/2}\sum_{t=1}^{n}\left\{\frac{\partial \ell_t(\theta_0)}{\partial\theta}-\frac{\partial \tilde{\ell}_t(\theta_0)}{\partial\theta}\right\}\right\|$ et $\sup_{\theta\in\mathbf{V}(\theta_0)}\left\|n^{-1}\sum_{t=1}^{n}\left\{\frac{\partial^2 \ell_t(\theta_0)}{\partial\theta\partial\theta'}-\frac{\partial^2 \tilde{\ell}_t(\theta_0)}{\partial\theta\partial\theta'}\right\}\right\|,$

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

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

f) $n^{-1}\sum_{t=1}^{n}\frac{\partial^2}{\partial\theta_i\partial\theta_j}\ell_t(\theta^*_{ij})\longrightarrow J(i,j)p.s.$

a) Intégrabilité des dérivées du critère en θ_0.

Puisque $\ell_t(\theta)=\varepsilon_t^2/\sigma_t^2+\log\sigma_t^2$, 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 montrer a) il suffira donc de montrer

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

D'après (2.25) nous avons

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

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

où $\underline{1} = (1,0,\cdots,0)'$, $\underline{\varepsilon}_t^2 = (\varepsilon_t^2,0,\cdots,0)'$, $B^{(j)}$ est une matrice $p \times p$ qui possède un 1 en position $(1,j)$, et des 0 partout ailleurs. Remarquons que, d'après la positivité des coefficients et (2.40)-(2.41), les dérivées de σ_t^2 sont positives ou nulles. D'après (2.40), il est clair que $\partial\sigma_t^2/\partial\omega$ est bornée. Puisque $\sigma_t^2 \geq \omega > 0$, il en est de même pour $\{\partial\sigma_t^2/\partial\omega\}/\sigma_t^2$. Cette variable possède donc des moments de tous ordres. D'après la seconde égalité de (2.40) et la positivité de tous les termes considérés, nous avons

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

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

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

En utilisant (2.24), nous avons $\|B^K\| \leq K\rho^k$ pour tout k. De plus ε_t^2 possédant 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 :

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

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 :

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

On a donc

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

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 bien sûr 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

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

Les arguments utilisés pour montrer (2.43) donnent alors

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

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

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

car $\beta_j B^{(j)} \leq B$ Comme pour (2.43), on en déduit

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

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 puisque par (2.42), les variables $\{\partial^2\sigma_t^2/\partial\alpha_i\}/\sigma_t^2$ sont bornées en θ_0, il est clair 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 élémentaire $x/(1+x) \leq x^{s/2}$ 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-Schwarz permet de conclure que le troisième espérance de (2.39) existe.

b) Inversibilité de J et lien avec la variance de la dérivée du critère.

En utilisant a) et à nouveau l'indépendance entre $\eta_t^2 = \varepsilon_t^2/\sigma_t^2(\theta_0)$ et σ_t^2 ainsi que ses dérivées, nous avons d'après (2.36)

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

De plus, étant donné (2.39), J existe et vérifie bien (2.33). Nous avons également

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

Supposons maintenant que J soit non inversible. Alors il existe un vecteur non nul λ de $\mathbb{R}^{p+q+1}$ tel que $\lambda'\{\partial\sigma_t^2(\theta_0)/\partial\theta\}$ *p.s.* D'après (2.21) et la stationnarité de $\{\partial\sigma_t^2(\theta_0)/\partial\theta\}_t$, 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 $\lambda = (\lambda_0, \lambda_1, \cdots, \lambda_{q+p})'$. 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 $\lambda = 0$, ce qui termine la preuve de b).

c) Intégrabilité uniforme des dérivées d'ordre 3 du critère.

En dérivant (2.37), nous obtenons

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

Commençons par étudier l'intégrabilité de $\{1-\varepsilon_t^2/\sigma_t^2\}$. C'est le terme le plus délicat à traiter. En effet nous n'avons pas intégrabilité de $\varepsilon_t^2/\sigma_t^2$ uniformément sur Θ: en $\theta = (\omega, 0')$, le rapport $\varepsilon_t^2/\sigma_t^2$ n'est intégrable que si $\mathbb{E}(\varepsilon_t^2)$ existe. Nous allons cependant montrer l'intégrabilité de $\{1-\varepsilon_t^2/\sigma_t^2\}$ uniformément en θ au voisinage de θ_0. Soit Θ^* un compact contenant θ_0 et contenu dans l'intérieur de Θ ($\forall\ \theta\in\Theta^*$, on a $\theta\geq\theta_*>0$ composante par composante). Notons B_0 la matrice B (définie en (2.23)) évaluée au point $\theta=\theta_0$. Pour tout $\delta>0$, il existe un voisinage $\mathbf{V}(\theta_0)$ de θ_0, entièrement contenu dans Θ^*, tel que pour tout $\theta\in\mathbf{V}(\theta_0)$, $B_0\leq(1+\delta)B$ (i.e. $B_0(i,j)\leq(1+\delta)B(i,j)$ pour tout i et tout j).
Notons que, puisque $\mathbf{V}(\theta_0)\subset\Theta^*$, on a $\sup\limits_{\theta\in\mathbf{V}(\theta_0)} 1/\alpha_i<\infty$. De (2.25), on tire

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

et, en utilisant encore $x/(1+x)\leq x^s$ pour tout $x\geq 0$ et tout $s\in]0,1[$,

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

Si on choisit s tel que $\mathbb{E}(\varepsilon_t^{2s})$ et, par exemple, $\delta=(1-\rho^s)/(2\rho^s)$ alors l'espérance de la série précédente est finie. On en déduit qu'il existe un voisinage $\mathbf{V}(\theta_0)$ de θ_0 tel que

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

En utilisant (2.49), en conservant le même choix de δ mais en choisissant s tel que $\mathbb{E}(\varepsilon_t^{4s})$, l'inégalité triangulaire donne

$$\left\| \sup_{\theta \in \mathbf{V}(\theta_0)} \frac{\varepsilon_t^2}{\sigma_t^2} \right\|_2 = \kappa_n^{1/2} \left\| \sup_{\theta \in \mathbf{V}(\theta_0)} \frac{\sigma_t^2(\theta_0)}{\sigma_t^2} \right\|_2$$
$$\leq \kappa_n^{1/2} K + \kappa_n^{1/2} K q \sum_{k=0}^{\infty} (1+\delta)^k \rho^{ks} \|\varepsilon_t^{2s}\|_{2s} < \infty. \tag{2.51}$$

Étudions maintenant le deuxième terme entre accolades dans (2.49). En dérivant (2.44), (2.45) et (2.46), à l'aide des arguments utilisés pour montrer (2.41), on obtient :

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

quand les indices i_1, i_2 et i_3 ne sont pas tous dans $\{q+1, q+2, \cdots, q+1+p\}$ (i.e. quand on dérive par au moins un paramètre autre qu'un des β_j). En reprenant les arguments utilisés pour montrer (2.42) et (2.46), puis (2.43), 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 $s \in]0,1[$. Comme $\mathbb{E}_{\theta_0} \{ \sup_{\theta \in \Theta^*} \underline{c}_{t-k}(1) \}^{2s} < \infty$ pour un $s > 0$, on en déduit

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

Remarquons que la relation (2.52) la puissance 2 peut être remplacée par une puissance d arbitrairement grande :

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

En utilisant l'inégalité de Cauchy-Schwarz, (2.50) et (2.52), on obtient :

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

Les autres termes entre parenthèses dans (2.48) se traitent de la même manière. On montre en particulier que

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

pour tout entier d. Ceci permet par exemple d'établir à l'aide de Hölder que

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

Les autres termes de la somme dans (2.48) se traitent de la même manière. Ainsi on obtient c).

d) Oubli asymptotique des valeurs initiales.

En utilisant (2.26), on obtient les équations analogues à (2.40)-(2.41) pour les dérivées de $\underline{\tilde{\sigma}}_t^2$:

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

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

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

où $\partial \underline{\tilde{\sigma}}_0^2 / \partial \omega$ vaut $(0, \cdots, 0)'$ quand les conditions initiales sont données par (2.14) et vaut $(1, \cdots, 1)'$ quand les conditions initiales sont données 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

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

En utilisant (2.27) on obtient :

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

Puisque

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

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

$$\begin{aligned}
\left|\frac{\partial \ell_t(\theta_0)}{\partial \theta_i} - \frac{\partial \tilde{\ell}_t(\theta_0)}{\partial \theta_i}\right| &\leq \left|\left\{\frac{\varepsilon_t^2}{\tilde{\sigma}_t^2} - \frac{\varepsilon_t^2}{\sigma_t^2}\right\}\left\{\frac{1}{\sigma_t^2}\frac{\partial \sigma_t^2}{\partial \theta_i}\right\} + \left\{1 - \frac{\varepsilon_t^2}{\tilde{\sigma}_t^2}\right\}\left\{\frac{1}{\sigma_t^2} - \frac{1}{\tilde{\sigma}_t^2}\right\}\left\{\frac{\partial \sigma_t^2}{\partial \theta_i}\right\}\right|(\theta_0) \\
&+ \left|\left\{1 - \frac{\varepsilon_t^2}{\tilde{\sigma}_t^2}\right\}\left\{\frac{1}{\tilde{\sigma}_t^2}\right\}\left\{\frac{\partial \sigma_t^2}{\partial \theta_i} - \frac{\partial \tilde{\sigma}_t^2}{\partial \theta_i}\right\}\right|(\theta_0) \\
&\leq K\rho^t(1+\eta_t^2)\left|1 + \left\{\frac{1}{\sigma_t^2(\theta_0)}\frac{\partial \sigma_t^2(\theta_0)}{\partial \theta_i}\right\}\right|.
\end{aligned}$$

Par suite

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

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

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

ce qui, par (2.60), montre la première partie de d).

Regardons maintenant l'influence asymptotique des valeurs initiales sur les dérivées secondes du critère en un voisinage de θ_0. D'après (2.37) et les majorations précédentes, nous avons

$$\begin{aligned}
&\sup_{\theta \in \mathbf{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 \tilde{\ell}_t(\theta_0)}{\partial \theta_i \partial \theta_j}\right\}\right| \\
&\leq n^{-1}\sum_{t=1}^{n}\sup_{\theta \in \mathbf{V}(\theta_0)}\left\{\frac{\varepsilon_t^2}{\tilde{\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}{\tilde{\sigma}_t^2}\right\}\left\{\left(\frac{1}{\sigma_t^2} - \frac{1}{\tilde{\sigma}_t^2}\right)\frac{\partial^2 \sigma_t^2}{\partial \theta_i \partial \theta_j} + \frac{1}{\tilde{\sigma}_t^2}\left(\frac{\partial^2 \sigma_t^2}{\partial \theta_i \partial \theta_j} - \frac{\partial^2 \tilde{\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}{\tilde{\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}{\tilde{\sigma}_t^2} - 1\right\}\left\{\left(\frac{1}{\sigma_t^2} - \frac{1}{\tilde{\sigma}_t^2}\right)\frac{\partial \sigma_t^2}{\partial \theta_i} + \frac{1}{\tilde{\sigma}_t^2}\left(\frac{\partial \sigma_t^2}{\partial \theta_i} - \frac{\partial \tilde{\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}{\tilde{\sigma}_t^2} - 1\right\}\left\{\frac{1}{\tilde{\sigma}_t^2}\frac{\partial \tilde{\sigma}_t^2}{\partial \theta_i}\right\}\left\{\left(\frac{1}{\sigma_t^2} - \frac{1}{\tilde{\sigma}_t^2}\right)\frac{\partial \sigma_t^2}{\partial \theta_j} + \frac{1}{\tilde{\sigma}_t^2}\left(\frac{\partial \sigma_t^2}{\partial \theta_j} - \frac{\partial \tilde{\sigma}_t^2}{\partial \theta_j}\right)\right\} \\
&\leq K n^{-1}\sum_{t=1}^{n}\rho^t \Upsilon_t,
\end{aligned}$$

où
$$\Upsilon_t = \sup_{\theta \in \mathbf{V}(\theta_0)} \left\{ 1 + \frac{\varepsilon_t^2}{\sigma_t^2} \right\} \left\{ 1 + \frac{1}{\sigma_t^2} \frac{\partial^2 \sigma_t^2}{\partial \theta_i \partial \theta_j} + \frac{1}{\sigma_t^2} \frac{\partial \sigma_t^2}{\partial \theta_i} \frac{1}{\sigma_t^2} \frac{\partial \sigma_t^2}{\partial \theta_j} \right\}.$$
D'après (7.58), (7.61) et l'inégalité de Hölder, on voit que, pour un certain voisinage $\mathbf{V}(\theta_0)$, l'espérance de Υ_t est une constante finie. En utilisant a nouveau l'inégalité de Markov on montre alors la seconde convergence de d).

e) 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 retrouver directement à partir de (2.36), en utilisant le fait que $\sigma_t^2(\theta_0) \in \varepsilon_u, u < t$ et $\mathbb{E}_{\theta_0}(\varepsilon_t^2 \mid \varepsilon_u, u < t) = \sigma_t^2(\theta_0)$:
$$\mathbb{E}_{\theta_0}\left(\frac{\partial}{\partial \theta} \ell_t(\theta_0) \mid \varepsilon_u, u < t \right) = \frac{1}{\sigma_t^4(\theta_0)} \left(\frac{\partial}{\partial \theta} \sigma_t^2(\theta_0) \right) \mathbb{E}_{\theta_0}(\sigma_t^2(\theta_0) - \varepsilon_t^2 \mid \varepsilon_u, u < t) = 0.$$
Notons également que, d'après (2.48), $\mathbf{V}_{\theta_0}(\partial \ell_t(\theta_0)/\partial \theta) = (\kappa_\eta - 1)J$ est finie. D'après l'inversibilité de J et les hypothèses sur la loi de η_t (qui entraînent $0 < \kappa_\eta - 1 < \infty$), cette matrice de variance est non dégénérée. Nous en 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 :
$$\begin{aligned}
\mathbf{V}_{\theta_0}(\lambda' \frac{\partial \ell_t(\theta_0)}{\partial \theta}) &= \lambda' \mathbb{E}_{\theta_0}(\frac{\partial \ell_t(\theta_0)}{\partial \theta}) \lambda \\
&= \lambda' \mathbb{E}_{\theta_0} \left\{ (1 - \eta_t^2)^2 \frac{1}{\sigma_t^4(\theta_0)} \frac{\partial \sigma_t^2(\theta_0)}{\partial \theta} \frac{\partial \sigma_t^2(\theta_0)}{\partial \theta'} \right\} \lambda \\
&= (\kappa_\eta - 1) \lambda' J \lambda.
\end{aligned}$$
En appliquant le (T.C.L) pour les différence de martingale, on aura
$$\frac{1}{\sqrt{n}} \sum_{t=1}^{n} \left\{ \lambda' \frac{\partial \ell_t(\theta_0)}{\partial \theta} \right\} \xrightarrow{\mathcal{L}} \mathcal{N}(0,\, (\kappa_\eta - 1) \lambda' J \lambda),$$
par suite, en appliquant le théorème de Wold-Cramer
$$\frac{1}{\sqrt{n}} \sum_{t=1}^{n} \{ \frac{\partial \ell_t(\theta_0)}{\partial \theta})\} \xrightarrow{\mathcal{L}} \mathcal{N}(0,\, (\kappa_\eta - 1) J).$$

f) 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 tous i et j
$$n^{-1} \sum_{t=1}^{n} \frac{\partial^2}{\partial \theta_i \partial \theta_j} \ell_t(\theta_{ij}^*) = n^{-1} \sum_{t=1}^{n} \frac{\partial^2}{\partial \theta_i \partial \theta_j} \ell_t(\theta_0) + n^{-1} \sum_{t=1}^{n} \frac{\partial}{\partial \theta'} \left\{ \frac{\partial^2}{\partial \theta_i \partial \theta_j} \ell_t(\tilde{\theta}_{ij}) \right\} (\theta_{ij}^* - \theta_0), \tag{2.61}$$

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

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

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

□

2.3 Méthode des deux phases

L'objectif de cette section est, d'une part l'élaboration d'une méthode en deux étapes, pour l'estimation des modèles ARCH, et d'autre 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 :

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

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

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

et $u_t=\eta_t^2-1$; $1\leq t\leq n$. Ainsi

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

où $\beta=[\beta_0,\,\beta_1,\,\cdots,\,\beta_p]$.

Étape 1: Construction de l'estimateur préliminaire

En remplaçant $\sigma_{t-1}^2(\beta)$ dans (2.62), on obtient :

$$Y_t = \mathbf{Z}_{t-1}'\beta + \sigma_{t-1}^2(\beta)\, u_t; \quad 1 \le t \le n, \tag{2.64}$$

où $\mathbb{E}\{\sigma_{t-1}^2(\beta)\, u_t\} = \mathbb{E}\left\{\sigma_{t-1}^2(\beta)\right\}\mathbb{E}\{u_t\} = 0\,;\, 1 \le t \le n.$

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 :

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

où $\mathbf{Z}$ est une matrice d'ordre $n\times$(1+p) dont la t-ème ligne est $\mathbf{Z}_{t-1}'$ et $\mathbf{Y}$ est le vecteur composé de Y_t; $1 \le t \le n$.

La démonstration de la normalité asymptotique de cet estimateur comme, celle des autres théorèmes de ce chapitre, repose sur l'application du théorème central limite approprié.

Étape 2: Construction de l'estimateur d'intérêt

A présent, nous allons utiliser $\hat{\beta}_{pr}$ pour construire un estimateur d'intérêt $\hat{\beta}$ de β comme suit :

En divisant (2.64) par $\sigma_{t-1}^2(\beta)$, on obtient :

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

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

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

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

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

Avant d'étudier les propriétés de l'estimateur d'intérêt, nous énoncerons d'abord un lemme qui traite l'estimateur préliminaire $\hat{\beta}_{pr}$. 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 et suffisantes d'existences des moments d'ordre r (r élevé) de Y sont donnés en fonction du paramètre β (voir Engle 1982; Thms 1 et 2).

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

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

La normalité asymptotique de $\hat{\beta}_{pr}$ 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 $\hat{\beta}$, on suppose que :

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

et

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

Notons que 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),*

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

Preuve. De (2.67),

$$\hat{\beta} = \left[\frac{1}{n}\sum_{t=1}^{n}\left\{\frac{\mathbf{Z}_{t-1}\mathbf{Z}_{t-1}^{'}}{\sigma_{t-1}^{4}(\hat{\beta}_{pr})}\right\}\right]^{-1}\left[\frac{1}{n}\sum_{t=1}^{n}\left\{\frac{\mathbf{Z}_{t-1}\mathbf{Y}_{t}}{\sigma_{t-1}^{4}(\hat{\beta}_{pr})}\right\}\right]$$

$$= \left[\frac{1}{n}\sum_{t=1}^{n}\left\{\frac{\mathbf{Z}_{t-1}\mathbf{Z}_{t-1}^{'}}{\sigma_{t-1}^{4}(\hat{\beta}_{pr})}\right\}\right]^{-1}\left[\frac{1}{n}\sum_{t=1}^{n}\left\{\frac{\mathbf{Z}_{t-1}(\mathbf{Z}_{t-1}^{'}\beta+\sigma_{t-1}^{2}(\beta)u_{t})}{\sigma_{t-1}^{4}(\hat{\beta}_{pr})}\right\}\right]$$

$$= \beta + \left[\frac{1}{n}\sum_{t=1}^{n}\left\{\frac{\mathbf{Z}_{t-1}\mathbf{Z}_{t-1}^{'}}{\sigma_{t-1}^{4}(\hat{\beta}_{pr})}\right\}\right]^{-1}\left[\frac{1}{n}\sum_{t=1}^{n}\left\{\frac{\mathbf{Z}_{t-1}\sigma_{t-1}^{2}(\beta)u_{t}}{\sigma_{t-1}^{4}(\hat{\beta}_{pr})}\right\}\right]$$

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

Montrons que si

$$\frac{1}{n}\sum_{t=1}^{n}\left\{\frac{1}{\sigma_{t-1}^{4}(\hat{\beta}_{pr})}-\frac{1}{\sigma_{t-1}^{4}(\hat{\beta})}\right\}\mathbf{Z}_{t-1}\mathbf{Z}_{t-1}^{'} = o_{p}(1), \tag{2.73}$$

et

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\sigma_{t-1}^{2}(\beta)\left\{\frac{1}{\sigma_{t-1}^{4}(\hat{\beta}_{pr})}-\frac{1}{\sigma_{t-1}^{4}(\beta)}\right\}\mathbf{Z}_{t-1}u_{t} = o_{p}(1), \tag{2.74}$$

on aura nécessairement

$$\sqrt{n}(\hat{\beta}-\beta) = \left[\frac{1}{n}\sum_{t=1}^{n}\left\{\frac{\mathbf{Z}_{t-1}\mathbf{Z}_{t-1}^{'}}{\sigma_{t-1}^{4}(\hat{\beta})}\right\}\right]^{-1}\left[\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\frac{\mathbf{Z}_{t-1}\sigma_{t-1}^{2}(\beta)u_{t}}{\sigma_{t-1}^{4}(\hat{\beta})}\right] + o_{p}(1).$$

Pour démontrer (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 $\mathbf{U} = [u_1, \cdots, u_k]'$, $\mathbf{V} = [v_1, \cdots, v_k]'$, et $\mathbf{W}$ un vecteur dont toutes les composantes sont positives alors

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

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

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

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

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

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

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

De (2.79)

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

Comme l'unique solution stationnaire non anticipative de (ε_t) est ergodique, et $\forall t$, $\mathbf{Z}_{t-1}$ s'écrit 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 $\mathbf{T}_2$,

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

Puisque

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

on obtient $\mathbf{T}_2=o_p(1)$.

Pour démontrer (2.74), on utilise l'équation (2.77) et en remplaçant $\sigma_{t-1}^2(\beta)$ par (2.63), on aura :

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

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

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

et

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

De (2.71), on conclut

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

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

□

2.4 Conclusion

Dans ce chapitre, nous avons présenté en trois sections les diverses méthodes d'estimation des modèles ARCH en mettant l'accent sur les propriétés asymptotiques de chaque estimateur.

Pour la méthode des moindres carrés et celle de Bose et 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 simples à obtenir.

Les estimateurs de maximum de vraisemblance des paramètres du modèle ARCH sont convergents et asymptotiquement normaux. La précision de ces estimateurs s'exprime en fonction de 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 Q.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 moments, cela explique pourquoi la méthode de Q.M.V est préférée.

Chapitre 3

Estimateur Q.M.V de paramètre α_0 d'un ARCH(1) non stationnaire

Dans ce chapitre, nous nous intéressons à l'étude de la consistence forte et de la normalité asymptotique de l'estimateur de quasi-maximum de vraisemblance du modèle ARCH(1) dans le cas non stationnaire, i.e, les paramètres du modèle se trouvent en dehors de la région de stationnarité. Lorsque la condition de stationnarité stricte n'est pas vérifiée, c'est à dire $\gamma \geq 0$, ε_t^2 converge presque sûre vers l'infini, ce qui va nous permettre de montrer la consistence forte de l'estimateur Q.M.V. Dans la première section du présent chapitre, le paramètre ω_0 est supposé fixé, α_0 est inconnu. Dans la seconde, ω_0 est considéré libre.
Soit le modèle ARCH(1) suivant :

$$\begin{cases} \varepsilon_t = \sqrt{h_t}\eta_t \\ h_t = \omega_0 + \alpha_0 \varepsilon_{t-1}^2 \end{cases} \tag{3.1}$$

où $\omega_0 > 0$, $\alpha_0 > 0$, et $(\eta_t)_t$ est une suite *iid* tel que $\mathbf{V}(\eta_t^2) = \mathbb{E}(\eta_t^4 - 1) = \xi < \infty$. Lorsque la contrainte de stationnarité stricte n'est pas vérifiée, on a nécessairement

$$\alpha_0 \geq \exp\{-\mathbb{E}\log \eta_t^2\}. \tag{3.2}$$

Notons que σ_t^2 converge vers l'infini, presque sûrement lorsque

$$\alpha_0 > \exp\{-\mathbb{E}\log \eta_t^2\}. \tag{3.3}$$

et seulement en probabilité lorsque l'inégalité (3.2) est vérifiée.
La section qui suit traite les propriétés asymptotiques de l'estimateur $\hat{\alpha}_n$ dans le cas où le paramètre ω_0 d'un ARCH(1) est connu. Sans perte de généralité, supposons $\omega_0 = 1$.

3.1 Cas où le paramètre ω est fixé

En fixant une valeur initiale pour ε_0^2, la vraisemblance conditionnelle gaussienne $L_n(\alpha)$ s'écrit:

$$L_n(\alpha) = \prod_{t=1}^{n} \frac{1}{\sqrt{2\Pi\sigma_t^2}} \exp\left(-\frac{\varepsilon_t^2}{2\sigma_t^2}\right),$$

où $\sigma_t^2(\alpha) = \omega_0 + \alpha\varepsilon_{t-1}^2$ pour tout $t = 1, \cdots, n$.
En prenant le logarithme de $L_n(\alpha)$, on aura :

$$\log L_n(\alpha) = -\frac{1}{2}\sum_{t=1}^{n}\left(\log\sigma_t^2 + \frac{\varepsilon_t^2}{\sigma_t^2}\right).$$

L'estimateur Q.M.V vérifie

$$\hat{\alpha}_n = \arg\min_{\alpha\in[0,+\infty[} \ell_n(\alpha), \quad \ell_n(\alpha) = -\frac{1}{2}\sum_{t=1}^{n}\left(\log\sigma_t^2 + \frac{\varepsilon_t^2}{\sigma_t^2}\right). \tag{3.4}$$

En dérivant l'équation (3.4), par rapport à α, on obtient :

$$\frac{\partial}{\partial\alpha}\ell_n(\alpha_0) = -\frac{1}{2}\sum_{t=1}^{n}\left(1 - \frac{\varepsilon_t^2}{\sigma_t^2}\right)\frac{\varepsilon_{t-1}^2}{\sigma_t^2}, \tag{3.5}$$

$$\frac{\partial^2}{\partial\alpha^2}\ell_n(\alpha_0) = \frac{1}{2}\sum_{t=1}^{n}\left(1 - 2\frac{\varepsilon_t^2}{\sigma_t^2}\right)\frac{\varepsilon_{t-1}^4}{\sigma_t^4}, \tag{3.6}$$

$$\frac{\partial^3}{\partial\alpha^3}\ell_n(\alpha_0) = -\sum_{t=1}^{n}\left(1 - 3\frac{\varepsilon_t^2}{\sigma_t^2}\right)\frac{\varepsilon_{t-1}^6}{\sigma_t^6}. \tag{3.7}$$

Le théorème qui suit établit la consistence forte et la normalité asymptotique de l'estimateur Q.M.V de α_0.

Théorème 3.1. *Soit le modèle ARCH(1) défini par* (3.37). *On suppose la condition de non stationnarité* (3.3) *vérifiée, et que* $\mathbf{V}(\eta_t^2) < \infty$. *Alors quand* $n \to \infty$,

$$\hat{\alpha}_n \xrightarrow{p.s} \alpha_0. \tag{3.8}$$

De plus

$$\sqrt{n}(\hat{\alpha}_n - \alpha_0) \xrightarrow{\mathcal{L}} \mathcal{N}(0,\xi\alpha_0^2). \tag{3.9}$$

La preuve de ce théorème repose sur les résultats des lemmes (3.1)-(3.5).

Lemme 3.1. *Supposons que* (3.3) *est vérifiée, alors*

$$\varepsilon_t^2 \xrightarrow{p.s} \infty \tag{3.10}$$

quand $t \to \infty$.

Lemme 3.2. *Si* (3.3) *est vérifiée, alors pour des nombres positifs $m \leq k$, quand $t \to \infty$ on a*

$$\frac{\varepsilon_{t-1}^{2m}}{(1+\alpha_0\varepsilon_{t-1}^2)^k} \xrightarrow{p.s} \begin{cases} \dfrac{1}{\alpha_0^m} & si \quad m = k, \\ \\ 0 \quad si & m < k, \end{cases} \tag{3.11}$$

de même, quand $n \to \infty$,

$$\frac{1}{n}\sum_{t=1}^{n} \frac{\varepsilon_{t-1}^{2m}}{(1+\alpha_0\varepsilon_{t-1}^2)^k} \xrightarrow{p.s} \begin{cases} \dfrac{1}{\alpha_0^m} & si \quad m = k, \\ \\ 0 \quad si & m < k. \end{cases} \tag{3.12}$$

Preuve du théorème (3.1).

a- Consistence forte

On considère la constante $\underline{\alpha} \in]0, \alpha_0[$. Montrons que $\hat{\alpha}_n > \underline{\alpha}$ pour n assez grand.
On note

$$\hat{\alpha}_n = \arg\min_{\alpha\in[0,\infty[} Q_n(\alpha), \quad Q_n(\alpha) = \frac{1}{n}\sum_{t=1}^{n}\{\ell_t(\alpha) - \ell_t(\alpha_0)\}.$$

On a

$$\begin{aligned} Q_n(\alpha) &= \frac{1}{n}\sum_{t=1}^{n} \eta_t^2 \left\{\frac{\sigma_t^2(\alpha_0)}{\sigma_t^2(\alpha)} - 1\right\} + \log\frac{\sigma_t^2(\alpha)}{\sigma_t^2(\alpha_0)} \\ &= \frac{1}{n}\sum_{t=1}^{n} \eta_t^2 \frac{(\alpha_0-\alpha)\varepsilon_{t-1}^2}{(1+\alpha\varepsilon_{t-1}^2)} + \log\frac{1+\alpha\varepsilon_{t-1}^2}{1+\alpha_0\varepsilon_{t-1}^2}. \end{aligned}$$

Puisque pour tout $x > 0$, $x \geq 1 + \log x$, on a :

$$\inf_{\alpha<\underline{\alpha}} Q_n(\alpha) \geq \log\frac{1}{n}\sum_{t=1}^{n}\eta_t^2 + \log\frac{(\alpha_0-\underline{\alpha})\varepsilon_{t-1}^2}{1+\alpha_0\varepsilon_{t-1}^2} + 1.$$

Pour tout $M > 0$, il exist un entier t_M tels que $\varepsilon_t^2 > M$, pour tout $t > t_M$, ce qui implique

$$\liminf_{n\to\infty} \inf_{\alpha<\underline{\alpha}} Q_n(\alpha) \geq \log\frac{(\alpha_0-\underline{\alpha})M}{1+\alpha_0 M} + 1.$$

Comme M est arbitrairement grand,

$$\liminf_{n\to\infty} \inf_{\alpha<\underline{\alpha}} Q_n(\alpha) \geq \log\frac{\alpha_0-\underline{\alpha}}{\alpha_0} + 1 > 0, \tag{3.13}$$

sous réserve que $\underline{\alpha} < (1-e^{-1})\alpha_0$. Si $\underline{\alpha}$ est choisi de sorte que la contrainte soit vérifiée, alors

$$\limsup_{n\to\infty} Q_n(\hat{\alpha}_n) \leq \limsup_{n\to\infty} Q_n(\alpha_0) = 0, \tag{3.14}$$

de (3.13) et (3.14)

$$\lim_{n\to\infty} \hat{\alpha}_n \geq \underline{\alpha} \quad p.s. \tag{3.15}$$

Nous définissons un critère O_n équivalent au critère Q_n. Comme $\varepsilon_{t-1}^2 \to \infty$ p.s. quand $t \to \infty$, pour $\alpha \neq 0$,

$$\lim_{n\to\infty} Q_n(\alpha) = \lim_{n\to\infty} O_n(\alpha),$$

où

$$O_n(\alpha) = \frac{1}{n}\sum_{t=1}^{n} \eta_t^2 \frac{(\alpha_0 - \alpha)}{\alpha} + \log \frac{\alpha}{\alpha_0}.$$

D'autre part, on a :

$$\lim_{n\to\infty} O_n(\alpha) = \frac{\alpha_0}{\alpha} - 1 + \log \frac{\alpha}{\alpha_0} > 0,$$

où $\alpha_0/\alpha \neq 1$.

A présent, montrons la convergence uniforme de $Q_n(\alpha) - O_n(\alpha)$ vers zero sur $\alpha \in]\underline{\alpha}, \infty[$. On a :

$$Q_n(\alpha) - O_n(\alpha) = \frac{1}{n}\sum_{t=1}^{n} \eta_t^2 \frac{\alpha - \alpha_0}{\alpha(1+\alpha\varepsilon_{t-1}^2)} + \log \frac{(1+\alpha\varepsilon_{t-1}^2)\alpha_0}{(1+\alpha_0\varepsilon_{t-1}^2)\alpha}.$$

Ainsi pour tous $M > 0$ et tout $\epsilon > 0$, presque sûrment

$$|Q_n(\alpha) - O_n(\alpha)| \leq (1+\epsilon)\frac{|\alpha - \alpha_0|}{\alpha^2 M} + \frac{|\alpha - \alpha_0|}{\alpha\alpha_0 M},$$

si n est assez grand. En plus les contraintes précédentes, supposent que $\underline{\alpha} < 1$.

On obtient $|\alpha - \alpha_0|/\alpha^2 M \leq \alpha_0/\underline{\alpha}^2 M$ pour $\underline{\alpha} < \alpha \leq \alpha_0$, et

$$\frac{|\alpha - \alpha_0|}{\alpha^2 M} \leq \frac{\alpha}{\alpha^2 M} \leq \frac{1}{\underline{\alpha} M} \leq \frac{1}{\underline{\alpha}^2 M},$$

pour tout $\alpha \geq \alpha_0$. On a :

$$\lim_{n\to\infty} \sup_{\alpha > \underline{\alpha}} |Q_n(\alpha) - O_n(\alpha)| \leq (1+\epsilon)\frac{1+\alpha_0}{\underline{\alpha}^2 M} + \frac{1+\alpha_0}{\underline{\alpha}\alpha_0 M}.$$

Comme M peut être choisi arbitrairement grand et ϵ arbitrairement petit, on a presque sûrement

$$\lim_{n\to\infty} \sup_{\alpha > \underline{\alpha}} |Q_n(\alpha) - O_n(\alpha)| = 0. \tag{3.16}$$

Pour la dernière étape de la preuve, soit deux constantes α_0^- et α_0^+ telles que $\underline{\alpha} < \alpha_0^- < \alpha_0 < \alpha_0^+$. En notant $\hat{\sigma}_\eta^2 = n^{-1}\sum_{t=1}^{n} \eta_t^2$, la solution de

$$\alpha_n^* = \arg\min_{\alpha} O_n(\alpha),$$

est $\alpha_n^* = \alpha_0 \hat{\sigma}_\eta^2$. Cette solution appartient à l'intervalle $]\alpha_0^-, \alpha_0^+[$ quand n est assez grand. Dans ce cas

$$\alpha_n^{**} = \arg \min_{\alpha \notin]\alpha_0^-, \alpha_0^+[} O_n(\alpha) \in \{\alpha_0^-, \alpha_0^+\},$$

est l'une des deux extrémités de l'intervalle $]\alpha_0^-, \alpha_0^+[$ et

$$\lim_{n\to\infty} O_n(\alpha_n^{**}) = \min\{\lim_{n\to\infty} O_n(\alpha_0^-), \lim_{n\to\infty} O_n(\alpha_0^+)\} > 0. \tag{3.17}$$

En utilisant (3.15), (3.16), (3.17), et le fait que $\min_\alpha Q_n(\alpha) \leq Q_n(\alpha_0) = 0$, on déduit

$$\lim_{n\to\infty} \arg\min_{\alpha\geq 0} Q_n(\alpha) \in]\alpha_0^-, \alpha_0^+[.$$

Puisque $]\alpha_0^-, \alpha_0^+[$ est un intervalle contenant α_0 qui peut être arbitrairement petit, et $\hat{\alpha}_n = \arg\min_{\alpha\geq 0} Q_n(\alpha)$, on obtient la convergence (3.8).

$\square$

b- La normalité asymptotique

Pour démontrer la normalité asymptotique de l'estimateur Q.M.V, les lemmes (3.3), (3.4) et (3.5) sont nécessaires.

Lemme 3.3. *Sous les hypothèses du théorème* (3.1)*, on a :*

$$(1/\sqrt{n})\frac{\partial}{\partial\alpha}\ell_n(\alpha_0) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \frac{\xi}{4\alpha_0^2}\right) \tag{3.18}$$

quand $n \to \infty$.

Preuve. Par définition, $(1/\sqrt{n})\frac{\partial}{\partial\alpha}\ell_n(\alpha_0) = (1/\sqrt{n})\sum_{t=1}^n s_t$, où

$$s_t = -\frac{1}{2}\left(1 - \frac{\varepsilon_t^2}{\sigma_t^2}\right)\frac{\varepsilon_{t-1}^2}{\sigma_t^2}.$$

La suite $(s_t, \mathcal{F}_t)$ est une différence de martingale, avec $\mathcal{F}_t = \sigma(\varepsilon_t, \varepsilon_{t-1}, \cdots, \varepsilon_0)$, alors

$$\mathbb{E}(s_t \mid \mathcal{F}_{t-1}) = -\frac{1}{2}\mathbb{E}(1-\eta_t^2)\frac{\varepsilon_{t-1}^2}{\sigma_t^2} = 0.$$

En utilisant (3.12), on obtient :

$$\frac{1}{n}\sum_{t=1}^n \mathbb{E}(s_t^2 \mid \mathcal{F}_{t-1}) = \frac{1}{n}\sum_{t=1}^n \left(\frac{\varepsilon_{t-1}^2}{1+\alpha_0\varepsilon_{t-1}^2}\right)^2 \frac{\xi}{4} \xrightarrow{p.s} \frac{\xi}{4\alpha_0^2} > 0,$$

où $\xi = \mathbb{E}(1-\eta_t^2)^2 < \infty$. De plus, s_t^2 est bornée par $\mu_t^2 = (1-\eta_t^2)^2/4\alpha_0^2$, et d'après le théorème central limite de Lindeberg pour les différences de martingales, $\forall\ \delta > 0$

$$\frac{1}{n}\sum_{t=1}^{n}\mathbb{E}(s_t^2\mathbb{I}\{|s_t| > \sqrt{n}\delta\}) \leq \mathbb{E}(\mu_t^2\mathbb{I}\{|\mu_t| > \sqrt{n}\delta\}) \stackrel{n\to\infty}{\longrightarrow} 0,$$

D'où

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{n} s_t \xrightarrow{\mathcal{L}} \mathcal{N}\left(0,\ \frac{\xi}{4\alpha_0^2}\right).$$

□

Lemme 3.4. *Sous les hypothèses du théorème* (3.1)*, on a :*

$$\frac{1}{n}\left(-\frac{\partial^2}{\partial\alpha^2}\ell_n(\alpha_0)\right) \xrightarrow{p.s} \frac{1}{2\alpha_0^2} > 0, \tag{3.19}$$

quand $n \to \infty$.

Preuve. Soit

$$\frac{1}{n}\left(-\frac{\partial^2}{\partial\alpha^2}\ell_n(\alpha_0)\right) = \frac{1}{2n}\sum_{t=1}^{n}\kappa_t\gamma_t,$$

avec $\kappa_t = 2\eta_t^2 - 1$ et $\gamma_t = \dfrac{\varepsilon_{t-1}^4}{\sigma_t^4} = \dfrac{\varepsilon_{t-1}^4}{(1+\alpha_0\varepsilon_{t-1}^2)^2}$.

En appliquant la loi forte de grands nombres, on obtient :

$$\frac{1}{n}\sum_{t=1}^{n}\kappa_t \xrightarrow{p.s} 1,$$

de (3.11), on trouve $\gamma_t \xrightarrow{p.s} \dfrac{1}{\alpha_0^2}$.

Finalement, on déduit le résultat de lemme (3.4).

□

Le lemme suivant nous donne la bornitude de la troisième dérivée de la fonction de vraisemblance, la preuve de ce lemme ne nécessite pas la non stationnarité de processus (ε_t).

Lemme 3.5. *On note par* $I(\alpha_0,\delta)$ *l'intervalle* $[\alpha_0-\delta, \alpha_0+\delta]$, $0 < \delta < \alpha_0$. *Alors par* $\frac{\partial^3}{\partial\alpha^3}\ell_n(\alpha_0)$ *donnée par* (3.7)

$$\sup_{\tilde{\alpha}\in I(\alpha_0,\delta)}\left|\frac{1}{n}\frac{\partial^3}{\partial\alpha^3}\ell_n(\tilde{\alpha})\right| \leq g(\alpha_0,\delta,n) \xrightarrow{p.s} \beta < \infty, \tag{3.20}$$

quand $n \to \infty$.

Preuve. Pour $\alpha_l = \alpha_0 - \delta$,

$$\begin{aligned}
\left|\frac{1}{n}\frac{\partial^3}{\partial\alpha^3}\ell_n(\tilde{\alpha})\right| &= \left|\frac{1}{n}\sum_{t=1}^{n}\left(3\frac{\varepsilon_t^2}{\sigma_t^2}-1\right)\frac{\varepsilon_{t-1}^6}{\sigma_t^6}\right| \\
&\leq \left|\frac{1}{n}\sum_{t=1}^{n}\left(3\frac{\varepsilon_t^2}{\sigma_t^2}-1\right)\right|\frac{1}{\alpha_l^3} \\
&= \left|\frac{1}{n}\sum_{t=1}^{n}\left(3\frac{\{1+\alpha_0\varepsilon_{t-1}^2\}\eta_t^2}{1+\tilde{\alpha}\varepsilon_{t-1}^2}-1\right)\right|\frac{1}{\alpha_l^3} \\
&\leq \frac{1}{n}\sum_{t=1}^{n}\left(3\left\{1+\frac{\alpha_0}{\alpha_l}\right\}\eta_t^2+1\right)\frac{1}{\alpha_l^3} := g(\alpha,\delta,n).
\end{aligned}$$

Par l'application de la loi des grands nombres, on obtient :

$$\left|\frac{1}{n}\frac{\partial^3}{\partial\alpha^3}\ell_n(\tilde{\alpha})\right| \xrightarrow{p.s} \beta,$$

où $\beta = \dfrac{4\alpha_l + 3\alpha_0}{\alpha_l^4} < \infty$.

□

La dérivée du critère s'annule en $\hat{\alpha}_n$. Un développement limité de cette dérivée autour de α_0 donne alors

$$\frac{1}{\sqrt{n}}\frac{\partial}{\partial\alpha}\ell_n(\alpha_0) + \frac{1}{n}\frac{\partial^2}{\partial\alpha^2}\ell_n(\alpha_0)\sqrt{n}(\hat{\alpha}_n-\alpha_0) + \frac{1}{n}\frac{\partial^3}{\partial\alpha^3}\ell_n(\alpha^*)\frac{\sqrt{n}}{2}(\hat{\alpha}_n-\alpha_0)^2 = 0,$$

où α^* est entre α et α_0. On déduit alors de (3.18), (3.19) et (3.20), et du théorème de Slutsky que

$$\sqrt{n}(\hat{\alpha}_n - \alpha) \xrightarrow{\mathcal{L}} \mathcal{N}(0,\, \xi\alpha^2).$$

□

On s'intéresse, dans la section qui suit, aux propriétés asymptotiques de l'estimateur $\hat{\alpha}_n$ dans le cas où le paramètre ω_0 du modèle ARCH(1) est considéré libre.

3.2 Cas où le paramètre ω est libre

En fixant une valeur initiale pour ε_0^2, la vraisemblance conditionnelle gaussienne $L_n(\theta)$ s'écrit :

$$L_n(\theta) = \prod_{t=1}^{n}\frac{1}{\sqrt{2\Pi\sigma_t^2}}\exp\left(-\frac{\varepsilon_t^2}{2\sigma_t^2}\right),$$

l'estimateur du quasi-maximum de vraisemblance du modèle ARCH(1) donné par (3.37), est défini comme une solution mesurable de l'équation

$$(\hat{\omega}_n,\hat{\alpha}_n) = \arg\min_{\theta\in\Theta} \frac{1}{n}\sum_{t=1}^{n} \ell_t(\theta), \quad \ell_t(\theta) = \frac{\varepsilon_t^2}{\sigma_t^2} + \log\sigma_t^2, \tag{3.21}$$

où $\theta = (\omega, \alpha)$, Θ est un compact de $]0, \infty[^2$, et $\sigma_t^2(\theta) = \omega + \alpha\varepsilon_{t-1}^2$ pour $t = 1, \cdots, n$. La convergence presque sûre de ε_t^2 vers l'infini va nous permettre de montrer la consistence forte de l'estimateur Q.M.V. Le lemme suivant précise la vitesse de convergence de ε_t^2 vers l'infini sous (3.3).

Lemme 3.6. *Soit le modèle ARCH(1) défini par* (3.37) *avec la condition initiale* $\varepsilon_0^2 \geq 0$ *quelconque. On suppose la condition de non stationnarité* (3.3) *vérifiée. Alors quand* $n \to \infty$ *presque sûrement,*

$$\frac{1}{h_n} = o(\rho^n) \quad et \quad \frac{1}{\varepsilon_n} = o(\rho^n),$$

pour toute constante ρ *telle que*

$$1 > \rho > \exp\{-\mathbb{E}\log\eta_t^2\}/\alpha_0. \tag{3.22}$$

Preuve. On a :

$$\rho^n h_n = \rho^n\omega_0\left\{1 + \sum_{t=1}^{n-1}\alpha_0^t\eta_{n-1}^2\cdots\eta_{n-t}^2\right\} + \rho^n\alpha_0^n\eta_{n-1}^2\cdots\eta_1^2\varepsilon_0^2 \geq \rho^n\omega_0\prod_{t=1}^{n-1}\alpha_0\eta_t^2. \tag{3.23}$$

D'où

$$\begin{aligned}\liminf_{n\to\infty}\frac{1}{n}\log\rho^n h_n &\geq \lim_{n\to\infty}\frac{1}{n}\left\{\log\rho\omega_0 + \sum_{t=1}^{n-1}\log\rho\alpha_0\eta_t^2\right\}\\ &= \mathbb{E}(\log\rho\alpha_0\eta_1^2),\end{aligned}$$

en utilisant (3.22) pour la dernière inégalité. Par suite $\log\rho^n h_n$, et donc $\rho^n h_n$, tend presque sûrement vers $+\infty$ quand $n \to \infty$. Pour toute fonction f à valeurs réelles, soit $f^+(x) = \max\{f(x),0\}$ et $f^-(x) = \max\{-f(x),0\}$, de sorte que $f(x) = f^+(x) - f^-(x)$. Puisque $\mathbb{E}(\log^+\eta_1^2) \leq \mathbb{E}(\eta_1^2) = 1$, on obtient $\mathbb{E}(|\log\eta_1^2|) = \infty$ si et seulement si $\mathbb{E}(\log\eta_1^2) = -\infty$. ainsi (3.3) implique $\mathbb{E}(|\log\eta_1^2|) < \infty$, ce qui nécessite que $\log\eta_n^2/n \to 0$ presque sûrement quand $n \to \infty$. En utilisant (3.23), $\liminf_{n\to\infty} n^{-1}\log\rho^n\eta_n^2 h_n \geq \mathbb{E}(\log\rho\alpha_0\eta_1^2)$, et $\rho^n\varepsilon_n^2 = \rho^n\eta_n^2 h_n \to \infty$ presque sûrement.

□

Ce résultat permet d'obtenir la convergence forte et la normalité asymptotique de l'estimateur Q.M.V de α_0.

Théorème 3.2.1. *Sous les hypothèses du lemme* (3.6)*, l'estimateur Q.M.V défini par* (3.21) *vérifie*

$$\hat{\alpha}_n \xrightarrow{p.s} \alpha_0, \tag{3.24}$$

et lorsque θ_0 appartient a l'interieur de Θ

$$\sqrt{n}(\hat{\alpha}_n - \alpha_0) \xrightarrow{\mathcal{L}} \mathcal{N}(0,\, \xi\alpha_0^2) \tag{3.25}$$

quand $n \to \infty$.

Dans la preuve de ce théorème on montre que le vecteur du score satisfait

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\frac{\partial}{\partial\theta}\ell_t(\theta_0) \xrightarrow{\mathcal{L}} \mathcal{N}\left\{0,\ J = \xi\begin{pmatrix} 0 & 0 \\ 0 & \alpha_0^{-2}\end{pmatrix}\right\}. \tag{3.26}$$

Preuve du théorème (3.2.1)**.**

La consistence forte

On note

$$(\hat{\omega}_n, \hat{\alpha}_n) = \arg\min_{\theta\in\Theta} Q_n(\theta),$$

où

$$Q_n(\theta) = \frac{1}{n}\sum_{t=1}^{n}\{\ell_t(\theta) - \ell_t(\theta_0)\}.$$

On a

$$\begin{aligned} Q_n(\theta) &= \frac{1}{n}\sum_{t=1}^{n}\eta_t^2\left\{\frac{\sigma_t^2(\theta_0)}{\sigma_t^2(\theta)} - 1\right\} + \log\frac{\sigma_t^2(\theta_0)}{\sigma_t^2(\theta)} \\ &= \frac{1}{n}\sum_{t=1}^{n}\eta_t^2\frac{(\omega_0-\omega)+(\alpha_0-\alpha)\varepsilon_{t-1}^2}{\omega+\alpha\varepsilon_{t-1}^2} + \log\frac{\omega+\alpha\varepsilon_{t-1}^2}{\omega_0+\alpha_0\varepsilon_{t-1}^2}. \end{aligned}$$

Pour tout $\theta \in \Theta$, nous avons $\alpha \neq 0$. En posant

$$O_n(\alpha) = \frac{1}{n}\sum_{t=1}^{n}\eta_t^2\frac{(\alpha_0-\alpha)}{\alpha} + \log\frac{\alpha}{\alpha_0},$$

et

$$d_t = \frac{\alpha(\omega_0+\omega) - \omega(\alpha_0-\alpha)}{\alpha(\omega+\alpha\varepsilon_{t-1}^2)},$$

nous avons

$$Q_n(\theta) - O_n(\alpha) = \frac{1}{n}\sum_{t=1}^{n}\eta_t^2 d_{t-1} + \frac{1}{n}\sum_{t=1}^{n}\log\frac{(\omega+\alpha\varepsilon_{t-1}^2)\alpha_0}{(\omega_0-\alpha_0\varepsilon_{t-1}^2)\alpha} \xrightarrow{p.s} 0,$$

puisque, d'après le lemme (3.6), $\varepsilon_t^2 \to \infty$ presque sûrement quand $t \to \infty$. Il est facile de voir que cette convergence est uniforme sur le compact Θ :

$$\lim_{n\to\infty}\sup_{\theta\in\Theta}|Q_n(\theta) - O_n(\alpha)| = 0 \quad p.s. \tag{3.27}$$

Soit deux constantes α_0^- et α_0^+ telles que $0 < \alpha_0^- < \alpha_0 < \alpha_0^+$. En notant $\hat{\sigma}_\eta^2 = n^{-1}\sum_{t=1}^{n}\eta_t^2$, la solution de

$$\alpha_n^* = \arg\min_{\alpha} O_n(\alpha),$$

est $\alpha_n^* = \alpha_0\hat{\sigma}_\eta^2$. Cette solution appartient à l'intervalle $]\alpha_0^-, \alpha_0^+[$ quand n est assez grand. Dans ce cas

$$\alpha_n^{**} = \arg\min_{\alpha\notin]\alpha_0^-,\alpha_0^+[} O_n(\alpha) \in \{\alpha_0^-, \alpha_0^+\}$$

et

$$\lim_{n\to\infty} O_n(\alpha_n^{**}) = \min\{\lim_{n\to\infty} O_n(\alpha_0^-), \lim_{n\to\infty} O_n(\alpha_0^+)\} > 0. \tag{3.28}$$

En raison de (3.27) et (3.28)

$$\begin{aligned}
\lim_{n\to\infty}\sup_{\theta\in\Theta}|Q_n(\theta) - O_n(\alpha)| &\geq \lim_{n\to\infty}\min_{\theta\in\Theta,\alpha\notin]\alpha_0^-,\alpha_0^+[} Q_n(\theta) - \lim_{n\to\infty}\min_{\theta\in\Theta,\alpha\notin]\alpha_0^-,\alpha_0^+[} O_n(\alpha^{**}) \\
\lim_{n\to\infty}\min_{\theta\in\Theta,\alpha\notin]\alpha_0^-,\alpha_0^+[} O_n(\alpha^{**}) &\geq \lim_{n\to\infty}\min_{\theta\in\Theta,\alpha\notin]\alpha_0^-,\alpha_0^+[} Q_n(\theta),
\end{aligned}$$

ce qui montre que presque sûrement

$$\lim_{n\to\infty}\min_{\theta\in\Theta,\alpha\notin]\alpha_0^-,\alpha_0^+[} Q_n(\theta) > 0.$$

Comme $\min_{\theta} Q_n(\theta) \leq Q_n(\theta_0) = 0$, on en déduit que, presque sûrement, $\hat{\theta}_n$ appartient à $V(\theta_0)$ pour n assez grand.

$$\lim_{n\to\infty}\arg\min_{\theta\in\Theta} Q_n(\theta) \in]0,\infty[\times]\alpha_0^-, \alpha_0^+[.$$

Puisque $]\alpha_0^-, \alpha_0^+[$ est un intervalle contenant α_0 qui peut être arbitrairement petit, et $\hat{\theta}_n = \arg\min_{\theta\in\Theta} Q_n(\theta)$, on obtient la convergence (3.24).

□

Pour démontrer la normalité asymptotique de l'estimateur Q.M.V, nous avons besoin du résultat intermédiaire suivant :

Lemme 3.7. *Sous les hypothèses du théorème* (3.2.1), *on a :*

$$\sum_{t=1}^{\infty} \sup_{\theta\in\Theta} \left| \frac{\partial}{\partial\omega}\ell_t(\theta) \right| < \infty \quad p.s., \tag{3.29}$$

$$\sum_{t=1}^{\infty} \sup_{\theta\in\Theta} \left\| \frac{\partial^2}{\partial\omega\partial\theta}\ell_t(\theta) \right\| < \infty \quad p.s., \tag{3.30}$$

$$\sup_{\theta\in\Theta} \left| \frac{1}{n}\sum_{t=1}^{n} \frac{\partial^2}{\partial\alpha^2}\ell_t(\omega,\alpha_0) - \frac{1}{\alpha_0^2} \right| = o(1) \quad p.s., \tag{3.31}$$

$$\frac{1}{n}\sum_{t=1}^{n} \sup_{\theta\in\Theta} \left| \frac{\partial^3}{\partial\alpha^3}\ell_t(\theta) \right| = O(1) \quad p.s. \tag{3.32}$$

Preuve. En utilisant le lemme (3.6), il existe une variable aléatoire réelle K et une constante $\rho \in]0,1[$ indépendantes de θ et de t telles que

$$\begin{aligned}
\left| \frac{\partial}{\partial\omega}\ell_t(\theta) \right| &= \left| \frac{1}{\sigma_t^2(\theta)} \frac{\partial\sigma_t^2(\theta)}{\partial\omega} \left(1 - \frac{\varepsilon_t^2}{\sigma_t^2(\theta)}\right) \right| \\
&= \left| \frac{-(\omega_0 + \alpha_0\varepsilon_{t-1}^2)\eta_t^2}{(\omega + \alpha\varepsilon_{t-1}^2)^2} + \frac{1}{\omega + \alpha\varepsilon_{t-1}^2} \right| \\
&\leq K\rho^t(\eta_t^2 + 1).
\end{aligned}$$

Comme la série $\sum_{t=1}^{\infty} K\rho^t(\eta_t^2+1)$ a une espérance finie, alors elle est presque sûrement finie, ce qui achève la démonstration de (3.29), (3.30) se démontre de la même manière. Pour

(3.31), on a :

$$\begin{aligned}
\frac{\partial^2 \ell_t(\omega,\alpha_0)}{\partial \alpha^2} - \frac{1}{\alpha_0^2} &= \left\{ 2\frac{(\omega_0+\alpha_0\varepsilon_{t-1}^2)\eta_t^2}{\omega+\alpha_0\varepsilon_{t-1}^2} - 1 \right\} \frac{\varepsilon_{t-1}^4}{(\omega+\alpha_0\varepsilon_{t-1}^2)^2} - \frac{1}{\alpha_0^2} \\
&= \left\{ \frac{2(\omega_0+\alpha_0\varepsilon_{t-1}^2)\eta_t^2 - (\omega+\alpha_0\varepsilon_{t-1}^2) + 2\omega\eta_t^2 - 2\omega\eta_t^2}{\omega+\alpha_0\varepsilon_{t-1}^2} \right\} \frac{\varepsilon_{t-1}^4}{(\omega+\alpha_0\varepsilon_{t-1}^2)^2} - \frac{1}{\alpha_0^2} \\
&= \left\{ \frac{2(\omega_0+\omega)\eta_t^2 + (2\eta_t^2-1)(\omega+\alpha_0\varepsilon_{t-1}^2)}{\omega+\alpha_0\varepsilon_{t-1}^2} \right\} \frac{\varepsilon_{t-1}^4}{(\omega+\alpha_0\varepsilon_{t-1}^2)^2} - \frac{1}{\alpha_0^2} \\
&= (2\eta_t^2-1)\frac{\varepsilon_{t-1}^4}{(\omega+\alpha_0\varepsilon_{t-1}^2)^2} - \frac{2\eta_t^2 - 2\eta_t^2 + 2 - 1}{\alpha_0^2} + r_{1,t} \\
&= (2\eta_t^2-1)\frac{\varepsilon_{t-1}^4}{(\omega+\alpha_0\varepsilon_{t-1}^2)^2} + \frac{2(\eta_t^2-1)}{\alpha_0^2} - \frac{(2\eta_t^2-1)}{\alpha_0^2} + r_{1,t} \\
&= 2(\eta_t^2-1)\frac{1}{\alpha_0^2} + r_{1,t} + r_{2,t},
\end{aligned}$$

où

$$\sup_{\theta\in\Theta} |r_{1,t}| = \sup_{\theta\in\Theta} \left| \frac{2(\omega_0-\omega)\eta_t^2}{(\omega+\alpha_0\varepsilon_{t-1}^2)} \frac{\varepsilon_{t-1}^4}{(\omega+\alpha_0\varepsilon_{t-1}^2)^2} \right| = o(1) \quad p.s.$$

et

$$\begin{aligned}
\sup_{\theta\in\Theta} |r_{2,t}| &= \sup_{\theta\in\Theta} \left| (2\eta_t^2-1) \left\{ \frac{\varepsilon_{t-1}^4}{(\omega+\alpha_0\varepsilon_{t-1}^2)^2} - \frac{1}{\alpha_0^2} \right\} \right| \\
&= \sup_{\theta\in\Theta} \left| (2\eta_t^2-1) \left\{ \frac{\omega^2 + 2\alpha_0\varepsilon_{t-1}^2}{\alpha_0^2(\omega+\alpha_0\varepsilon_{t-1}^2)^2} \right\} \right| \\
&= o(1) \quad p.s.
\end{aligned}$$

quand $t \to \infty$. On en déduit (3.30). Quant à (3.32), il suffit de remarquer que

$$\begin{aligned}
\left| \frac{\partial^3}{\partial \alpha^3} \ell_t(\theta) \right| &= \left| \left\{ 2 - 6\frac{(\omega_0+\alpha_0\varepsilon_{t-1}^2)\eta_t^2}{\omega+\alpha\varepsilon_{t-1}^2} \right\} \left(\frac{\varepsilon_{t-1}^2}{\omega+\alpha\varepsilon_{t-1}^2} \right)^3 \right| \\
&\leq \left\{ 2 + 6\left(\frac{\omega_0}{\omega} + \frac{\alpha_0}{\alpha} \right) \eta_t^2 \right\} \frac{1}{\alpha^3}.
\end{aligned}$$

□

Normalité asymptotique

Signalons que nous ne savons pas a priori si la dérivée du critère s'annule en $\hat{\theta}_n = (\hat{\omega}_n, \hat{\alpha}_n)$ car nous n'avons que la convergence presque sûr de $\hat{\alpha}_n$ vers α_0. Ainsi le minimum du critère pourrait parfois se trouver sur le bord de Θ, même asymptotiquement. Par contre la dérivée partielle par rapport à la seconde coordonnée doit asymptotiquement s'annuler en l'optimum puisque $\hat{\alpha}_n \to \alpha_0$ et $\theta_0 \in \mathring{\Theta}$. Un développement limité de la dérivée du critère donne

$$\begin{pmatrix} \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \frac{\partial}{\partial \omega} \ell_t(\hat{\theta}_n) \\ \\ 0 \end{pmatrix} = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \frac{\partial}{\partial \theta} \ell_t(\theta_0) + J_n \sqrt{n}(\hat{\theta}_n - \theta_0), \tag{3.33}$$

où J_n est une matrice 2×2 dont les éléments sont de la forme

$$J_n(i,j) = \frac{1}{n} \sum_{t=1}^{n} \frac{\partial^2}{\partial \theta_i \partial \theta_j} \ell_t(\theta^*_{i,j}),$$

avec $\theta^*_{i,j} = (\omega^*_{i,j}, \alpha^*_{i,j})$ entre $\hat{\theta}_n$ et θ_0. D'après le lemme (3.6) qui montre que $\varepsilon_t^2 \to \infty$ presque sûrement et d'après le théorème central limite de Lindeberg pour les différences de martingales

$$\begin{aligned} \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \frac{\partial}{\partial \alpha} \ell_t(\theta_0) &= \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \frac{\varepsilon_{t-1}^2}{\sigma_t^2(\theta_0)} - \frac{\varepsilon_t^2 \varepsilon_{t-1}^2}{\sigma_t^4(\theta_0)} \\ &= \frac{1}{\sqrt{n}} \sum_{t=1}^{n} (1 - \eta_t^2) \frac{\varepsilon_{t-1}^2}{\omega_0 + \alpha_0 \varepsilon_{t-1}^2} \\ &= \frac{1}{\sqrt{n}} \sum_{t=1}^{n} (1 - \eta_t^2) \frac{1}{\alpha_0} + o_p(1) \\ &\xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \frac{\xi}{\alpha_0^2}\right). \end{aligned} \tag{3.34}$$

La relation (3.30) du lemme (3.7) et la compacite de Θ nous assurent que

$$J_n(2,1)\sqrt{n}(\hat{\omega}_n - \omega_0) \leq \sum_{t=1}^{\infty} \sup_{\theta \in \Theta} \left\| \frac{\partial^2}{\partial \omega \partial \theta} \ell_t(\theta) \right\| \frac{1}{\sqrt{n}} (\hat{\omega}_n - \omega_0) \xrightarrow{p.s} 0. \tag{3.35}$$

En faisant un nouveau développement limité de la fonction

$$\alpha \longmapsto \frac{1}{n} \sum_{t=1}^{n} \frac{\partial^2}{\partial \alpha^2} \ell_t(\omega^*_{2,2}, \alpha),$$

nous obtenons

$$J_n(2,2) = \frac{1}{n}\sum_{t=1}^{n}\frac{\partial^2}{\partial\alpha^2}\ell_t(\omega^*_{2,2},\alpha_0) + \frac{1}{n}\sum_{t=1}^{n}\frac{\partial^3}{\partial\alpha^3}\ell_t(\omega^*_{2,2},\alpha^*)(\alpha^*_{2,2}-\alpha_0),$$

où α^* est entre $\alpha^*_{2,2}$ et α_0. En utilisant (3.31), (3.32) et (3.24) on obtient

$$J_n(2{,}2) \xrightarrow{p.s} \frac{1}{\alpha_0^2}. \tag{3.36}$$

On conclut avec la seconde ligne de (3.33) et avec (3.34), (3.35) et (3.36).

□

Les résultats obtenus dans ce chapitre sont très intéressants d'un point de vue théorique, car même si la condition de stationnarité stricte n'est pas vérifiée l'estimateur Q.M.V de paramètre α_0 est fortement consistent et asymptotiquement gaussien. En réalité, il n'est pas possible d'estimer simultanément ω_0 et α_0 sous l'hypothèse de non stationnarité (3.3), et on ne dispose d'aucun résultat asymptotique pour le paramètre ω_0, et seulement l'estimateur de α_0 qui confirme les propriétés asymptotiques.

3.3 Simulation

Afin d'illustrer les résultats théorique de ce chapitre, nous présentons une étude par simulation. Elle est à base du langage de programmation **R** téléchargeable sur le site **http://cran.rproject.org/**.
Soit le modèle ARCH(1) suivant :

$$\begin{cases} \varepsilon_t = \sqrt{h_t}\eta_t \\ h_t = \omega_0 + \alpha_0\varepsilon^2_{t-1} \end{cases} \tag{3.37}$$

où $\omega_0 > 0$, $\alpha_0 > 0$, et supposons que la loi des variables η_t soit une $\mathcal{N}(0{,}1)$. Le vecteur des paramètres est $\theta = (\omega,\alpha)'$. La contrainte de non stationnarité stricte, s'écrit d'après (3.3)

$$\alpha_0 > \exp\{-\mathbb{E}(\log\eta_t^2)\} \simeq 3.56.$$

Afin que la vraie valeur $\theta_0 = (\omega_0,\alpha_0)$ appartienne à Θ, on peut prendre un espace des paramètres de la forme $\Theta = [0, +\infty] \times [0.0001, +\infty]$.
Afin d'illustrer les performances de l'estimateur du maximum de vraisemblance, nous allons procéder à la simulation des erreurs d'estimation du maximum de vraisemblance des paramètres du modèle ARCH(1) en distinguant la forte et la faible stationnarité de la non

stationnarité.
A chaque itération de la simulation, on génère n réalisations d'un ARCH(1) pour des valeurs différentes des paramètres (ω, α) et on calcule la valeur de l'estimateur du maximum de vraisemblance. Pour chaque valeur du paramètre, on effectue 100 répliques. A l'aide de la boite à moustache, on déduit la qualité de l'estimateur $(\hat{\omega}, \hat{\alpha})$. Le programme est à base du langage de programmation **R**. Signalons au passage, le package `fGarch` n'inclut pas le cas non stationnaire ce qui exclut l'utilisation des commandes `garchSim` et `garchFit` qui permettent de simuler et d'estimer les paramètres du ARCH. Cet état de fait nous a conduit à créer notre propre générateurs (simulation) de données.

Interprétation des résultats

De la figure 3.1, on constate que les résultats obtenus pour $\theta_0 = (1, 0.95)$ et $\theta_0 = (1, 1.5)$ sont similaires. Précisons que le cas $\theta_0 = (1, 1.5)$ correspond à ARCH(1) de variance infinie. En d'autres termes, on constate que les erreurs d'estimation sont symétriques, ce qui est mis en évidence par l'écart interquartile étalé d'une façon symétrique (homogène et bien réparti).

Par contre, les figures du bas obtenus pour un ARCH explosif (cas non stationnaire, $\theta_0 = (1, 4)$), mettent en évidence, une dissymétrie des erreurs d'estimation du paramètre ω_0. Plus précisément, on constate que les valeurs sont hétérogènes et mal réparties, ce qui confirme le bien fondé de la théorie. Ce résultat est d'autant plus clair lorsque la taille de l'échantillon augmente.

Aussi, ce résultat illustre l'impossibilité d'estimer efficacement le paramètre ω_0 avec une précision raisonnable sous l'hypothèse de non stationnarité.

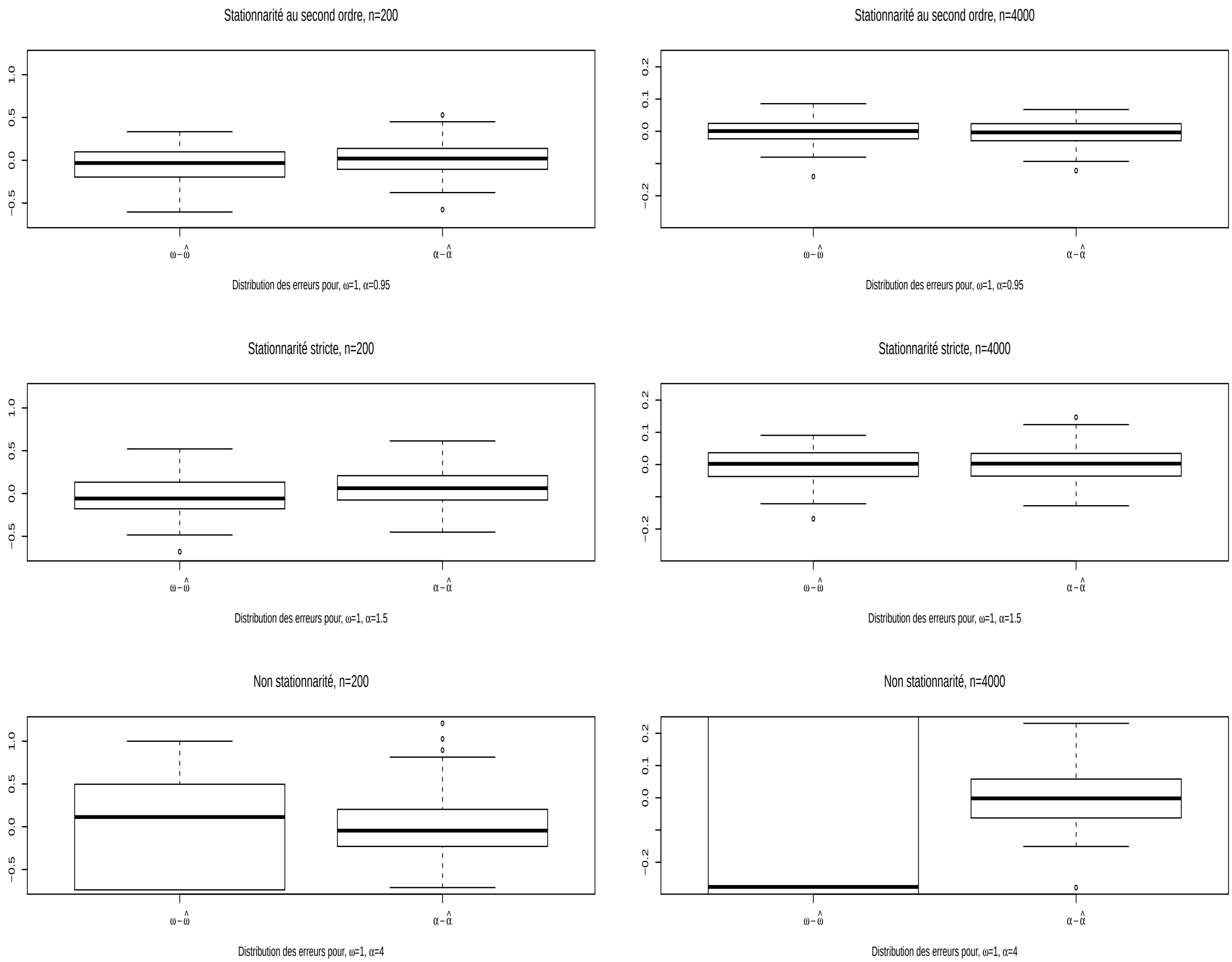

FIG. 3.1 – *Boîtes-à-moustaches des erreurs d'estimation par Q.M.V des paramètres ω_0 et α_0 d'un ARCH(1), avec $\eta_t \sim \mathcal{N}(0,1)$.*

Conclusion

Le but de ce travail a été de mettre en évidence les modèles non linéaires et l'hétéroscédasticité conditionnelle qui possède, maintenant, des outils puissants d'analyse. Cette modélisation est fondée sur des bases théoriques solides nécessaire pour les séries chronologiques présentant une dynamique non linéaire. Le concept de variance conditionnelle a commencé a jouer un grand rôle au début des années quatre vingt avec l'article fondateur de Engle (1982). Il caractérise les modèles venus élargir la classe des modèles classiques fondés essentiellement sur une structure de dépendance linéaire entre une variable a un instant t et ses valeurs passées et celles d'un bruit blanc et de ses valeurs passées.
Par ailleurs, il est important de noter que les séries financières sont aussi caractérisées par une volatilité non stationnaire et par des phénomènes d'asymétrie qui ne peuvent pas être pris en compte par les modélisations classique. C'est dans ce contexte que nous nous sommes intéressés aux modèles ARCH. Après présentation des différentes approches qui s'offrent à nous pour estimer les paramètres du modèle ARCH stationnaires, entre autres la méthode de Bose et Mukherjee (2003) et le maximum de vraisemblance. Par suite, un intérêt particulier a été accordé à l'estimation du modèle ARCH(1) non stationnaire. La consistance et la normalité asymptotique du Q.M.V est établie lorsque les paramètres se retrouvent en dehors de la région de stationnarité. Nous nous sommes aussi basés sur des résultats de simulation afin d'illustrer les résultats théoriques.
Ce travail ouvre de très nombreuses voies de recherche en particulier dans l'estimation des modèles ARCH(P) non stationnaires.

Annexe

3.4 Martingale

Définition 3.1 (Martingale). Soit $(\mathrm{X}_t)_{t\in\mathbb{N}}$ une suite de variables aléatoires réelles (v.a.r) sur un espace probabilisé $(\Omega, \mathcal{F}, \mathrm{P})$, et $(\mathcal{F}_t)_{t\in\mathbb{N}}$ est une suite de tribus. La suite $\{(\mathrm{X}_t, \mathcal{F}_t) : t = 1, 2, \cdots\}$ 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. $\mathbb{E}(|\mathrm{X}_t|) < \infty$;
4. $\mathbb{E}(\mathrm{X}_{t+1} \mid \mathcal{F}_t) = \mathrm{X}_t$.

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

Définition 3.2 (Différence de martingale). Soient $(X_t)_{t\in\mathbb{N}}$ une suite de variables aléatoires réelles (v.a.r), et $(\mathcal{F}_t)_{t\in\mathbb{N}}$ est une suite de tribus. La suite $\{(\mathrm{X}_t, \mathcal{F}_t) : t = 1, 2, \cdots\}$ 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. $\mathbb{E}(|\mathrm{X}_t|) < \infty$;
4. $\mathbb{E}(\mathrm{X}_{t+1} \mid \mathcal{F}_t) = 0$.

3.5 Ergodicité

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

Théorème 3.5.1. *Si $(Z_t)_{t\in\mathbb{Z}}$, est une suite fortement stationnaire et ergodique.*

et si $f : \mathbb{R}^{\infty} \longrightarrow \mathbb{R}$ *est une fonction mesurable.*
et soit $Y_t = f(\cdots, Z_{t-1}, Z_t, Z_{t+1}, \cdots)$, *alors* $(Y_t)_{t\in\mathbb{Z}}$ *reste une suite fortement stationnaire et ergodique.*

Théorème 3.5.2 (d'ergodicité). *Si* $(Z_t)_{t\in\mathbb{Z}}$ *est strictement stationnaire et ergodique, si* f *est mesurable et si* $\mathbb{E}\{|f(\cdots, Z_{t-1}, Z_t, Z_{t+1}, \cdots)|\} < \infty$, *alors:*
$\frac{1}{n}\sum_{t=1}^{n} f(\cdots, Z_{t-1}, Z_t, Z_{t+1}, \cdots) \longrightarrow \mathbb{E}\{f(\cdots, Z_{t-1}, Z_t, Z_{t+1}, \cdots)\} \quad p.s.$

3.6 Théorème de Chung-Fuchs

Théorème 3.6.1 (sur les marches aléatoires). *Si* $X_1, \cdots, X_n$ *est une suite iid telle que* $\mathbb{E}(X_1) = 0$ *et* $E|X_1| > 0$, *alors p.s.*

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

et

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

3.7 Critère de Cauchy

Définition 3.7.1 (Critère de Cauchy pour la convergence d'une suite de terme $a_n \geq 0$ **).** Soit $\lambda = \limsup a_n^{\frac{1}{n}}$.

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

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

3.8 La Régression Linéaire Multiple

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

Définition 3.8.1. De manière générale, le modèle linéaire peut s'écrire de la manière suivante :

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

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

3.8.1 Notations

On rencontre principalement trois types de notations :

Notation simple

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

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

Notation vectorielle

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

$$y_i = x_i^{'}\beta + \mu_i.$$

Notation matricielle

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

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

avec :

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

3.9 théorème de Wold-Cramer

Théorème 3.9.1. *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 $\lambda' Z_n \xrightarrow{\mathcal{L}} \lambda' Z$.*

3.10 théorème central limite (T.C.L) pour différence de martingale stationnaire

Théorème 3.10.1. *Si $(\nu_t, \mathcal{F}_t)$ 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 $\mathbf{V}(\nu_t) = \sigma_\nu^2 \neq 0$, alors*

$$\frac{1}{\sqrt{n}} \sum_{t=1}^{n} \nu \xrightarrow{\mathcal{L}} \mathcal{N}(0,\, \sigma_\nu^2).$$

3.11 T.C.L de Lindeberg

Théorème 3.11.1. *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 $\sigma_{nk}^2 = \mathbb{E}(\eta_{nk}^2 \mid \mathcal{F}_{n(k-1)})$. Si*

$$\sum_{k=1}^{n} \sigma_{nk}^2 \xrightarrow{p} \sigma_0^2 \quad quand \quad n \to \infty, \tag{3.38}$$

où σ_0^2 est une constante strictement positive, et

$$\sum_{k=1}^{n} \mathbb{E}(\eta_{nk}^2 \mathbb{I}_{\{|\eta_{nk}| \geq \epsilon\}}) \to 0 \quad quand \quad n \to \infty, \tag{3.39}$$

pour chaque réel positif ϵ, alors $\sum_{k=1}^{n} \eta_{nk} \xrightarrow{\mathcal{L}} \mathcal{N}(0,\, \sigma_0^2)$.

3.12 Lemme de Slutsky

Lemme 3.8. *Soit $\{X_n\}_{n=0,\cdots,\infty}$ une suite de variables aléatoires qui tend en loi vers X et soit $\{Y_n\}_{n=0,\cdots,\infty}$ 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_nY_n tend en loi vers cX.*

3.13 Lemme de Césaro

Lemme 3.9. *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 Cesàro converge également vers* ℓ *:*

$$C_n = \sum_{k=1}^{n} a_k \to \ell, \quad \textit{quand} \quad n \to \infty.$$

3.14 Lemme de Borel-Cantelli

Lemme 3.10. *Soit* (A_n) $(n \geq 1)$ *une suite d'évènements ; 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'évènements* A_n *se réalisent.*
b) Supposons les évènements A_n *indépendants 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'évènements* A_n *se réalisent.*

3.15 Inégalité de Hölder

Définition 3.15.1. Soit $1 \leq p \leq \infty$, $\dfrac{1}{p} + \dfrac{1}{q} = 1$, 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

$$\mathbb{E}(|XY|) \leq (\mathbb{E}|X|^p)^{1/p}(\mathbb{E}|Y|^q)^{1/q}.$$

3.16 Inégalité de Minkowski

Soit $1 \leq p \leq \infty$, et soit X et Y deux variables aléatoires alors

$$\mathbb{E}(|X+Y|) \leq (\mathbb{E}|X|^p)^{1/p} + (\mathbb{E}|Y|^q)^{1/q}.$$

Définition 3.16.1.

3.17 Cauchy-Schwarz

Définition 3.17.1. Si les variables X et Y sont de carré intégrable, alors

$$|\mathbb{E}(XY)| \leq \sqrt{\mathbb{E}(X^2)}\sqrt{\mathbb{E}(Y^2)}.$$

3.18 Inégalité de Markov

Définition 3.18.1. Soit Z une variable aléatoire réelle définie sur un espace probabilisé $(\Omega,\mathcal{A},\mathcal{P})$ et supposée $p.s$ positive ou nulle alors : $\forall\ a > 0,\ P(Z \geq a) \leq \mathbb{E}(Z)/a$.

3.19 Inégalité de Jensen

Définition 3.19.1. Soit X une v.a.r. et f une application convexe de $\mathbb{R}$ dans $\mathbb{R}$. On suppose X et $f(X)$ intégrables. Alors $f(\mathbb{E}(X)) \leq \mathbb{E}(f(X))$.

3.20 Convergence en probabilité

Soit $\{a_n,\ n = 1,\ 2,\ \cdots\}$ une suite de nombres réelles strictement positifs et soit $\{X_n,\ n = 1,\ 2,\ \cdots\}$ une suite de variables aléatoires dans le même espace probabilisé.

Définition 3.20.1 (Convergence en probabilité vers zero). On dit que X_n converge vers zero en probabilité, on écrit $X_n = o_p(1)$ ou $X_n \xrightarrow{p} 0$, si pour tout $\epsilon > 0$,

$$\mathbb{P}(|X_n| > \epsilon) \longrightarrow 0 \quad quand \quad n \longrightarrow \infty.$$

Définition 3.20.2 (Bornitude en probabilité). On dit que $\{X_n\}$ est borné en probabilité, on écrit $X_n = O_p(1)$, si pour tout $\epsilon > 0$, $\exists\ \delta(\epsilon) \in\ ]0, \infty[$,

$$\mathbb{P}(|X_n| > \delta(\epsilon)) < \epsilon \quad \forall \quad n.$$

Définition 3.20.3 (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 3.20.1. *Si X_n et Y_n, $n = 1,\ 2,\ \cdots$, des variables aléatoires dans un même espace probabilisé et $a_n > 0$, $b_n > 0$, $n = 1,\ 2,\ \cdots$, alors*
(i) si $X_n = o_p(a_n)$ et $Y_n = o_p(b_n)$, on aura
- $X_n Y_n = o_p(a_n b_n)$,
- $X_n + Y_n = o_p(\max(a_n, b_n))$,
- $|X_n|^r = o_p(a_n^r)$ *pour* $r > 0$.

(ii) si $X_n = o_p(a_n)$ et $Y_n = O_p(b_n)$, on aura :

$X_nY_n = o_p(a_nb_n)$.

Remarque 3.20.1. (i) reste vraie même si on remplace o_p par O_p.

Bibliographie

[1] Bose, A. and Mukherjee, K. (2003) Estimating the ARCH parameters by solving linear equations. Journal of Time Series Analysis 24, 127–136.

[2] Christophe Hurlin (2004) Econométrie pour la Finance, Modèles ARCH - GARCH.

[3] Francq , C. and Zakoian, J.-M. (2008) Can one really estimate nonstationary GARCH models?

[4] Francq , C. and Zakoian, J.-M. (2010) GARCH models : Structure, Statistical Inference and Financial Applications, 1st edition. WILEY.

[5] Jensen, S.T. and Rahbek, A. (2004a) Asymptotic normality of the QMLE estimator of ARCH in the nonstationary case. Econometrica 72, 641–646.

[6] Zakoian, J.-M. (1992) Modèles ARCH : une revue de la littérature. Journal de la société statistique de Paris, tome 133, $n°$ 1-2(1992), p.40-57.