

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE.  
MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE SCIENTIFIQUE.

UNIVERSITE MOULOUD MAMMARI, TIZI-OUZOU.  
FACULTE : DES SCIENCES  
DEPARTEMENT : MATHÉMATIQUES



## MEMOIRE DE MASTER

en  
Mathématiques appliquées

Option: Processus aléatoires et statistique de la décision

Thème

# Etude des modèles GARCH et application

Présenté par

TAHENNI Mona

Devant le jury d'examen composé de :

<i>M<sup>r</sup></i> BOUDIBA Mohand Arezki	Maître de conférence A	U.M.M.T.O	Président
<i>M<sup>r</sup></i> Berkoun Youcef	Maître de conférence A	U.M.M.T.O	Rapporteur
<i>M<sup>r</sup></i> Douki Hamid	Chargé de cours	U.M.M.T.O	Examineur

Soutenu le 03 / 10 / 2013

## *Remerciements*

*En préambule de ce mémoire, je tiens à remercier mes parents, ma soeur, mes frères walid et fares qui m'ont toujours soutenue, encouragé et stimulé tout au long de mes études.*

*Je tiens à exprimer toute ma reconnaissance à M<sup>r</sup> **Berkoun Youcef** pour l'honneur qu'il m'a fait en assurant la direction et le suivi scientifique et technique du présent mémoire. Je le remercie pour sa grande contribution à l'aboutissement de ce travail, et pour sa disponibilité.*

*Je remercie vivement M<sup>r</sup> **Boudiba Mohand Arezki** pour l'honneur qu'il me fait en acceptant de présider le jury de ce mémoire.*

*Mes remerciements chaleureux s'adressent également à M<sup>r</sup> **Douki Hamid** pour avoir accepté d'examiner ce travail.*

*Je remercie tous ceux qui ont contribué de près ou de loin à la réalisation de ce modeste mémoire.*

*Je remercie tous mes amis ainsi que tous ceux qui ont contribué de près ou de loin à la réalisation de ce modeste mémoire.*

# Table des matières

<b>Introduction Générale</b>	<b>3</b>
<b>1 Outils préliminaires</b>	<b>5</b>
1.1 Processus stochastiques : . . . . .	5
1.2 Processus stationnaires . . . . .	5
1.2.1 Processus stationnaires du second ordre . . . . .	5
1.2.2 La stationnarité stricte . . . . .	6
1.3 Séries chronologiques . . . . .	6
1.4 Processus stationnaire ergodique . . . . .	7
1.5 Équation aux récurrences stochastiques . . . . .	7
1.6 Solution stationnaire . . . . .	9
1.7 Moments de la loi stationnaire . . . . .	14
1.8 Modèles à volatilité stochastique . . . . .	15
<b>2 Processus conditionnellement hétéroscédastiques</b>	<b>17</b>
2.1 Présentation du modèle GARCH . . . . .	17
2.2 Stationnarité des modèles GARCH : . . . . .	19
2.2.1 Etude d'un modèle GARCH (1,1) . . . . .	19
2.2.2 Modèle GARCH (p,q) . . . . .	23
2.3 Moments d'ordre supérieurs . . . . .	29
2.4 Estimation des paramètres . . . . .	30
2.4.1 Méthode du maximum de vraisemblance . . . . .	33
2.4.2 Quasi-vraisemblance conditionnelle . . . . .	36
<b>3 Notion de Mesures de Risque</b>	<b>39</b>
3.1 Mesure de risque . . . . .	39
3.2 Caractéristiques d'une mesure de risque . . . . .	39
3.3 Présentation de la Value at Risk . . . . .	40
3.4 Calcul de la Value at Risk Par le modèle GARCH . . . . .	42
3.5 Prévission de la VaR . . . . .	43
<b>Conclusion</b>	<b>46</b>



# Introduction Générale

L'analyse des séries chronologiques a pour but, la mise en oeuvre et l'évaluation des modèles mathématiques adéquats pour la description des phénomènes aléatoires évolutifs dans le temps. On comprend du modèle mathématique une représentation simplifiée d'un phénomène aléatoire sous étude; cette représentation soumise à des hypothèse qui simplifient l'étude de la série chronologique pour un objectif préalablement fixé. L'élaboration d'un modèle mathématique est basé sur des outils mathématiques, c'est-à-dire des méthodes et des techniques qui sont fondées sur l'observation du phénomène, la mise en oeuvre du modèle correspondant et en terminant par la validation de ce modèle. Une fois ce modèle mis en oeuvre, on peut l'exploiter en vue de résoudre différentes problématiques posées au préalable. Le modèle mathématique le plus utilisé pour représenter une série chronologique est le processus aléatoire qui s'agit d'une famille de variables aléatoires caractérisées par une structure de dépendance stochastique et provenant d'une distribution spécifique. La qualité de la représentation du phénomène aléatoire par un modèle mathématique dépend de ce que la structure de dépendance du processus puisse le mieux prendre de dépendance du phénomène observé.

La modélisation des séries financières est un problème complexe. Cette complexité tient surtout à l'existence de régularités statistiques communes à un très grand nombre de séries financières et difficiles à reproduire artificiellement à partir des modèles stochastique. La plupart des séries financières (action, obligation, taux de change, ...) présentent généralement des volatilités; les modèles GARCH introduit par Bollerslev (1986), ont été proposé pour prendre en compte des variances conditionnelles dépendant du temps. Le principe général consiste donc à remettre en cause la propriété d'homoscédasticité que l'on retient généralement dans le cadre du modèle linéaire.

La modélisation GARCH est devenue un outil incontournable en finance, particulièrement utile pour prévoir et analyser la volatilité.

Ces modèles rendent compte des faits stylisés connus tels que la variation de la volatilité (hétéroscédasticité) et de clustering (les période de forte volatilité alternent avec des périodes de faible volatilité).

Les modèles GARCH presentent un intérêt pour s'adapter aux changements de regime assez important. Ces modèles decrivent finalement la nature aléatoire de certains phénomènes.

L'objectif de ce mémoire est d'étudier le processus GARCH.

Ce mémoire comporte 3 chapitres;

Le premier chapitre présente les concepts théorique de base; on présente quelques notions sur le processus stochastique à savoir, la stationnarité et l'ergodicité. La notion d'équations aux réccurrences stochastiques, de solution non-anticipative et l'exposant de Lyapounov est décrite. On donne les différentes conditions nécessaires et suffisantes assurant l'existence et l'unicité de la solution non-anticipative strictement stationnaire et ergodique; par la suite, on introduit les modèles à volatilité stochastique.

Le dixième chapitre est consacré aux processus GARCH; on donne les définitions précises du processus GARCH ainsi que ses propriétés. Le processus GARCH peut être défini à partir d'une équation aux récurrences stochastique. Par la suite, les paramètres du modèle GARCH sont estimées par la méthode du maximum de vraisemblance et la méthode du quasi-maximum de vraisemblance.

Dans le dernier chapitre, nous aborderons l'estimation d'une mesure de risque par la modélisation GARCH; on s'intéressera plus particulièrement à la valeur au risque (VaR). L'avantage d'une estimation de la VaR pour la modélisation GARCH réside dans:

- les violations de la VaR sont moins importantes qu'avec les autres modèles.
- les moindres écarts entre la prediction de la pert et la perte.

# Chapitre 1

## Outils préliminaires

Pour une meilleure compréhension de notre travail, on donne quelques outils qui seront nécessaires par la suite.

### 1.1 Processus stochastiques :

**Définition 1.1.1.** *soit  $(\Omega, \mathcal{A}, \mathbb{P})$  un espace probabilisé ou*

$$\begin{aligned} X : \Omega \times T &\rightarrow E \\ (w, t) &\rightarrow X(w, t) = X_t(w) \end{aligned}$$

*Un processus stochastique  $X$  est une famille de variables aléatoires réelles  $(X_t)_{t \in T}$ , où  $T \subset \mathbb{R}$  est appelé l'espace des temps et  $E$  l'espace des états. Si  $T \subset \mathbb{Z}$ , le processus est dit à temps discret et si  $T$  est un intervalle de  $\mathbb{R}$ , le processus est dit à temps continu.*

### 1.2 Processus stationnaires

#### 1.2.1 Processus stationnaires du second ordre

Dans de très nombreux cas, on ne peut pas renouveler la suite de mesures dans des conditions parfaitement identiques (météo, économie...). Pour que le modèle déduit à partir d'une suite d'observations ait un sens, il faut que toute portion de trajectoire observée fournisse des informations sur la loi de  $X$  et que des portions différentes mais de même longueur fournissent les mêmes indications. C'est ce qui nous amène à définir la notion de stationnarité.

**Définition 1.2.1.** *Un processus  $X = (X_t)_{t \in \mathbb{Z}}$  est stationnaire du second ordre (ou faiblement stationnaire) si et seulement si*

*(i)  $\mathbb{E}(X_t) = \mu, \forall t \in \mathbb{Z}$ ;*

- (ii)  $X_t$  est de carré intégrable pour tout  $t \in \mathbb{Z}$ :  $\mathbb{E}(X_t^2) < \infty$ ;
- (iii)  $Cov(X_s, X_{s+t}) = Cov(X_{s-1}, X_{s-1+t}) = \dots = Cov(X_0, X_t)$ ,  $\forall t, s \in \mathbb{Z}$

Dans cette définition, la propriété (i) exprime la stationnarité en moyenne, (ii) assure que la variance de chaque variable est finie et (iii) précise ce que l'on entend par "invariance de la structure de covariance". Par cette propriété, on peut introduire la fonction

$$\gamma(h) = Cov(X_t, X_{t+h})$$

qui est indépendante de  $t$  et n'est définie que si le processus est stationnaire du second ordre (et vérifie donc (iii)).

**Propriétés 1.2.1 (1).** *Si  $X$  est un processus stationnaire,  $Var(X_t) = \gamma(0) = \sigma^2$  est indépendante de  $t$ .*

*On dit que le processus est **homoscédastique**.*

On peut également définir un concept de stationnarité à partir des lois jointes des variables aléatoires du processus  $X$

## 1.2.2 La stationnarité stricte

**Définition 1.2.2.** *Un processus  $X = (X_t)_{t \in \mathbb{Z}}$  est **strictement stationnaire** si et seulement si*

*$\forall k \in \mathbb{N}, t_1, \dots, t_k \in \mathbb{Z}$  et  $\forall h \in \mathbb{Z}$ , la conjointe de  $(X_{t_1+h}, \dots, X_{t_k+h})$  ne dépend pas de  $h$  i.e*

$$L(X_{t_1+h_1}, X_{t_2+h_2}, \dots, X_{t_k+h_k}) = L(X_{t_1}, X_{t_2}, \dots, X_{t_k}), \forall t, k(h_1, h_2, \dots, h_k).$$

Notons que si  $X$  est un processus strictement stationnaire, la loi de  $X_t$  ne dépend pas de  $t$ , les  $var X_t$ , pour  $t \in \mathbb{Z}$  sont identiquement distribuées.

Cette dernière définition de stationnarité est plus exigeante que le concept de stationnarité du second ordre puisque tout processus strictement stationnaire est stationnaire du second ordre. La réciproque est fautive. La stationnarité à l'ordre 2 est bien plus facile à étudier et à vérifier que la stationnarité stricte.

## 1.3 Séries chronologiques

**Définition 1.3.1.** *Une série chronologique (ou temporelle) n'est que des enregistrements des observations, habituellement fait à des intervalles de temps réguliers, d'une grandeur aléatoire liée à un phénomène :*

- *Economique (évolution d'un indice boursier, les rendements d'un actif financier, le taux de change de devises, ... etc.)*

- *Météorologique (température moyenne hebdomadaire ou mensuelle dans une ville, niveau moyen mensuel de la précipitation de pluie dans une région, niveau mensuel d'eau dans un barrage,...) et autres.*

Une série chronologique peut être, donc, considérée comme un cheminement ou une trajectoire prise par un certain processus stochastique  $\{X_t, t \in \mathbb{Z}\}$

## 1.4 Processus stationnaire ergodique

On dit qu'une suite stationnaire est ergodique si elle satisfait la loi forte des grands nombres (même si les conditions d'application usuelles ne sont pas satisfaites).

**Définition 1.4.1.** Un processus strictement stationnaire  $(Z_t)_{t \in \mathbb{Z}}$ , à valeurs réelles, est dit ergodique si et seulement si, pour tout borélien  $\mathfrak{B}$  et tout entier  $k$ ,

$$n^{-1} \sum_{t=1}^n I_{\mathfrak{B}}(Z_t, Z_{t+1}, \dots, Z_{t+k}) \longrightarrow P(Z_1, \dots, Z_{1+k}) \in \mathfrak{B}p.s$$

Le concept d'ergodicité est bien plus général. Il peut être étendu à des suites non stationnaires (voir e.g. Billingsley (1995) "Probability and Measure", Wiley, New York.)

Toute transformation fixe d'une suite stationnaire ergodique reste stationnaire et ergodique.

**Théorème 1.4.1 (7).** Si  $(Z_t)_{t \in \mathbb{Z}}$  est une suite strictement stationnaire ergodique et si  $(Y_t)_{t \in \mathbb{Z}}$  est définie par

$$Y_t = f(Z_t, Z_{t+1}, \dots; Z_{t-1}, Z_{t-2}, \dots),$$

où  $f$  est une fonction mesurable de  $\mathbb{R}^\infty$  dans  $\mathbb{R}$ , alors  $(Y_t)_{t \in \mathbb{Z}}$  est également une suite strictement stationnaire ergodique.

## 1.5 Équation aux récurrences stochastiques

### Définition du modèle

On considère un espace vectoriel  $\mathbb{R}^d$  muni de la norme euclidienne notée  $|\cdot|$  et  $M_d(\mathbb{R})$  l'espace vectoriel des matrices carrées d'ordre  $d$  muni d'une norme euclidienne matricielle notée  $\|\cdot\|$ .

Pour tout  $x \in \mathbb{R}^d$ , on a :

$$|x| = \left( \sum_{k=1}^d x_k^2 \right)^{\frac{1}{2}}$$

Pour toute matrice  $a \in M_d(\mathbb{R})$ , on a :

$$\|a\| = \sup\{|ax| : x \in \mathbb{R}^d, |x| = 1\}$$

Soient  $(\Omega, \mathcal{A}, \mathbb{P})$  un espace de probabilité et  $((a_n, b_n))_{n \in \mathbb{Z}}$  une suite de variables aléatoires *i.i.d.*, où  $a_n$  est une matrice aléatoire à valeurs dans  $M_d(\mathbb{R})$  et  $b_n$  un vecteur aléatoire à valeurs dans  $\mathbb{R}^d$ .

On définit maintenant l'équation aux récurrences stochastique.

**Définition 1.5.1.** On dit qu'un processus aléatoire  $(X_n)_{n \in \mathbb{Z}}$  défini sur un espace de probabilité  $(\Omega, \mathcal{A}, \mathbb{P})$  à valeur dans  $\mathbb{R}^d$  satisfait ou est une solution de l'équation aux récurrences stochastique, s'il existe une suite de variables aléatoires  $((a_n, b_n))_{n \in \mathbb{Z}}$  *i.i.d.* définie sur le même espace de probabilité à valeurs dans  $M_d(\mathbb{R}) \times \mathbb{R}^d$  telle que :

$$X_{n+1} = a_n X_n + b_n, \quad n \in \mathbb{Z} \tag{1.1}$$

Il est clair que l'équation (1,1) définit un processus autorégressif avec des coefficients qui sont des matrices aléatoires. Dans le cas particulier où les matrices aléatoires  $a_n$  sont déterministes,  $(X_n)_{n \in \mathbb{Z}}$  est un processus autorégressif d'ordre 1, AR(1). Par conséquent, les processus aléatoires satisfaisant l'équation aux récurrences stochastique sont dits processus autorégressif généralisés.

Il serait intéressant de voir pour quelles conditions sur  $((a_n, b_n))_{n \in \mathbb{Z}}$ , la solution non anticipative strictement stationnaires de l'équation (1,1) existe.

**Définition 1.5.2.** Une solution non-anticipative de l'équation aux récurrences stochastique est un processus aléatoire  $(X_n)_{n \in \mathbb{Z}}$  tel que

$$\forall h \in \mathbb{N}, X_n \text{ et } (a_{n+h}, b_{n+h}) \text{ sont indépendantes,}$$

ou encore si

$$X_n = f((a_{n-1}, b_{n-1}), (a_{n-2}, b_{n-2}), \dots).$$

Avant qu'on puisse formuler les résultats sur l'existence d'une telle solution, on définit la notion d'exposant de Lyapounov.

**Définition 1.5.3.** Pour une suite  $(a_n)_{n \in \mathbb{Z}}$  de matrices aléatoires *i.i.d.*, la constante

$$\gamma = \inf \left\{ \frac{1}{n} E \log \|a_1 \dots a_n\|, n \in \mathbb{N} \right\} \tag{1.2}$$

est dite exposant de Lyapounov.

**Remarque 1.1.**

1. L'exposant de Lyapounov ne dépend pas de la norme choisie.
2. Supposons que la partie positive du logarithme de  $a_0$  est intégrable,  $E \log^+ \|a_0\| < \infty$ , alors une application du théorème ergodique subadditif de Kingman (1973) [20] ou des résultats de Furstenberg et Kesten (1960)[16] donnent

$$\gamma = \lim_{n \rightarrow \infty} \frac{1}{n} \log \|a_1 \dots a_n\|, \text{ p.s} \quad (1.3)$$

Une autre définition de L'exposant de Lyapounov est basée sur l'ensemble des vecteurs positifs dans  $\mathbb{R}^d$  de la norme égale à 1 et l'ensemble des matrices positive dans  $M_d(\mathbb{R})$ . Soit  $\mathbb{S}^{d-1}$  une sphère dans  $\mathbb{R}^d$  de rayon 1 et  $\mathbb{S}_+$  l'ensemble de ses élément positifs.

$$\mathbb{S}^{d-1} = \{x \in \mathbb{R}^d : |x| = 1\}, \mathbb{S}_+ = \{x \in \mathbb{S}^{d-1} : x \geq 0\}$$

Si la condition  $E(\log^+ \|a_n\|) < \infty$  est satisfaite, alors

$$\gamma = \lim_{n \rightarrow \infty} \frac{1}{n} \log \|a_1 \dots a_n\| = \lim_{n \rightarrow \infty} \frac{1}{n} \log |a_n \dots a_1 x|$$

## 1.6 Solution stationnaire

Dans cette partie, on décrit des résultats généraux des équations aux récurrences stochastiques. Premièrement, on étudie les équations aux récurrences stochastiques multidimensionnelles. Après qu'on ait défini l'exposant de Lyapounov, on donne les conditions suffisantes sur  $((a_n, b_n))_{n \in \mathbb{Z}}$  pour l'existence d'une solution non-anticipative strictement stationnaire pour l'équation aux récurrences stochastiques .

**Théorème 1.6.1 (6).** *Supposons que  $E(\log^+ |b_0|) < \infty$ ,  $E(\log^+ \|a_0\|) < \infty$ . Si  $\gamma < 0$ , alors*

$$X_n = \sum_{m=0}^{\infty} \left( \prod_{j=1}^m a_{n-j} \right) b_{n-m-1} \quad (1.4)$$

*converge presque sûrement et le processus  $(X_n)_{n \in \mathbb{Z}}$  est l'unique solution non-anticipative strictement stationnaire et ergodique de (1,1).*

**Preuve.** Soit  $X$  une variable initiale à l'instant  $n=0$ . Après  $n$  itérations dans l'équation (1,1), on obtient

$$X_n(X) = \sum_{m=0}^{n-1} \left( \prod_{j=1}^m a_{n-j} \right) b_{n-m-1} + \left( \prod_{m=1}^n a_{n-m} \right) X.$$

Considérons la norme du terme général de la série du côté droit de (1,4).

$$|a_{n-1} \dots a_{n-m} b_{n-m-1}| = \exp\left(m \frac{1}{m} \log |a_{n-1} \dots a_{n-m} b_{n-m-1}|\right)$$

Comme  $\log |a_{n-1} \dots a_{n-m} b_{n-m-1}| \leq \log |a_{n-1} \dots a_{n-m}| + \log |b_{n-m-1}|$ , alors

$$|a_{n-1} \dots a_{n-m} b_{n-m-1}| \leq \exp\left(m \left(\frac{1}{m} \log |a_{n-1} \dots a_{n-m}| + \log^+ |b_{n-m-1}|\right)\right)$$

En utilisant la condition  $E \log^+ |b_n| < \infty$  et la loi forte des grands nombres, il est clair que

$$E \log^+ |b_{n-m-1}| / m \xrightarrow{p.s} 0, \text{ si } m \rightarrow \infty$$

De plus, par la loi forte des grands nombres, on a

$$\frac{1}{m} \log |a_{n-1} \dots a_{n-m}| \leq \frac{1}{m} \sum_{j=1}^m \log |a_{n-j}| \xrightarrow{p.s} \gamma, \text{ si } m \rightarrow \infty.$$

Par conséquent,

$$\exp\left(m \left(\frac{1}{m} \sum_{j=1}^m \log |a_{n-j}| + \frac{1}{m} \log^+ |b_{n-m-1}|\right)\right) \xrightarrow{p.s} 0 \text{ si } m \rightarrow \infty$$

donc  $|a_{n-1} \dots a_{n-m} b_{n-m-1}|$  converge presque sûrement vers zéro et par conséquent la série (1,4) converge absolument presque sûrement.

Maintenant, on montre que  $|X_n(X) - X_n|$  converge presque sûrement vers zéro. En effet,

$$\begin{aligned} |X_n(X) - X_n| &= \left| \sum_{m=0}^{n-1} \left( \prod_{j=1}^m a_{n-j} \right) b_{n-m-1} + \left( \prod_{m=1}^n a_{n-m} \right) X - \sum_{m=0}^{\infty} \left( \prod_{j=1}^m a_{n-j} \right) b_{n-m-1} \right| \\ &= \left| - \sum_{m=n}^{\infty} \left( \prod_{j=1}^m a_{n-j} \right) b_{n-m-1} + \left( \sum_{m=1}^n a_{n-m} \right) X \right| \\ &= \left| \left( \prod_{m=1}^n a_{n-m} \right) (-X_0 + X) \right| \leq \left( \prod_{m=1}^n |a_{n-m}| \right) (|X_0| + |X|). \end{aligned} \quad (1.5)$$

Par la loi forte des grands nombres, on déduit que

$$\prod_{m=1}^n |a_{n-m}| = \exp\left(n \left(\frac{1}{n} \sum_{m=1}^n \log |a_{n-m}|\right)\right) \xrightarrow{p.s} 0 \text{ si } n \rightarrow \infty$$

Par conséquent,  $|X_n(X) - X_n|$  converge presque sûrement vers zéro.

La solution (1,4) est strictement stationnaire. En effet,  $(a_n, b_n) \stackrel{D}{=} (a_{n+i}, b_{n+i})$ ,  $\forall i \in \mathbb{Z}$  où  $\stackrel{D}{=}$  désigne l'égalité en distribution, alors le décalage dans les indices de  $a_n$  et  $b_n$  de la série infinie (1,4) implique le décalage dans  $X_n$ , ce qui signifie que  $X_{n+1} \stackrel{D}{=} X_n$ .

Considérons l'équation aux récurrences stochastiques (1,1) pour  $n \in \mathbb{N}$ , c'est-à-dire

$$X_{n+1} = a_n X_n + b_n, \quad n \in \mathbb{N}, \quad X_n \in \mathbb{R}^d. \quad (1.6)$$

La solution unique étant donné une variable initiale  $X_0$ , s'écrit

$$X_n = \left( \prod_{m=0}^{n-1} a_m \right) X_0 + \sum_{m=0}^{n-1} \left( \prod_{j=m+1}^{n-1} a_j \right) b_m. \quad (1.7)$$

Notons que pour tout  $n \in \mathbb{N}$ ,

$$(X_0, (a_k, b_k)_{0 \leq k \leq n-1}) \stackrel{D}{=} (X_0, (a_{n-k-1}, b_{n-k-1})_{0 \leq k \leq n-1}),$$

ceci implique que

$$X_n \stackrel{D}{=} \left( \prod_{m=0}^{n-1} a_m \right) X_0 + \sum_{m=0}^{n-1} \left( \prod_{j=1}^m a_{j-1} \right) b_m.$$

Les propriétés de la solution (1,7) sont données par le résultat suivant :

**Théorème 1.6.2 (24).** *Soit  $(X_n)_{n \in \mathbb{N}}$  définie par (1,7). Supposons que  $E(\log^+ |b_n|) < \infty$ ,  $E(\log^+ ||a_0||) < \infty$  et  $\gamma < 0$ , alors*

1.  $X_n \xrightarrow{D} X$  si  $n \rightarrow \infty$ ,  $X$  satisfait l'équation stochastique

$$X \stackrel{D}{=} aX + b, \quad X \text{ et } (a, b) \text{ sont indépendantes} \quad (1.8)$$

2.  $X$  est donnée par :

$$X = \sum_{m=0}^{\infty} \left( \prod_{j=1}^m a_{j-1} \right) b_m, \quad (1.9)$$

où le côté droit converge presque sûrement.

3. Si on choisit  $X_0 \stackrel{D}{=} X$  comme dans (1,9), alors le processus  $(X_n)_{n \in \mathbb{N}}$  défini par (1,7) est strictement stationnaire.

Supposons que  $(X_n)_{n \in \mathbb{N}}$  est une solution non-anticipative strictement stationnaire de (1,6), alors il est clair que  $(X_n, (a_n, b_n))_{n \in \mathbb{N}}$  est strictement stationnaire. Ce processus peut être prolongé au processus strictement stationnaire  $(X_n, (a_n, b_n))_{n \in \mathbb{Z}}$  qui vérifie l'équation aux récurrences stochastiques

$$X_{n+1} = a_n X_n + b_n, \quad \text{p.s. } \forall n \in \mathbb{Z},$$

et le processus  $(X_n)_{n \in \mathbb{Z}}$  est l'unique solution stationnaire de cette équation.

**Remarque 1.2.** On peut remplacer la condition  $E(\log^+ |b_0|) < \infty$  par une condition plus faible  $E(|b_0|^\beta) < \infty$  pour  $\beta > 0$ . Notons que l'exposant de Lyapounov de  $(a_n)_{n \in \mathbb{Z}}$  est inférieur à zéro dès que  $E(\log ||a_0||) < 0$  (prenant  $n = 1$  dans la définition 1.5.3).

**Définition 1.6.1.** Soit  $a = (a_{ij})_{1 \leq i, j \leq d}$  et  $b = (b_{ij})_{1 \leq i, j \leq d}$  deux matrices de  $M_d(\mathbb{R})$ , alors le produit de Kronecker noté  $\otimes$  de a et b est une matrice carrée d'ordre  $d^2$  qui s'écrit comme suit :

$$a \otimes b = \begin{pmatrix} a_{11}b & \cdots & a_{1d}b \\ \vdots & \ddots & \vdots \\ a_{d1}b & \cdots & a_{dd}b \end{pmatrix}$$

**Définition 1.6.2.** Soit  $a \in M_d(\mathbb{R})$  une matrice inversible. On appelle valeur propre de a tout  $\lambda \in \mathbb{C}$  tel qu'il existe  $x \in \mathbb{C}$ ,  $x \neq 0$  tel que  $ax = \lambda x$ . On appelle rayon spectral de a la quantité  $\rho(a) = \max\{|\lambda| : \lambda \in \mathbb{C}, \lambda \text{ valeur propre de } a\}$

On a aussi une autre condition assurant la négativité de l'exposant de Lyapounov.

**Proposition 1.6.1 (11).** *Si le rayon spectral de la matrice  $E(a_1 \otimes^t a_1)$  est de module strictement inférieur à 1, alors l'exposant de Lyapounov est négatif.*

**Preuve.** De la remarque 1.1,  $\gamma$  ne dépend pas de la norme matricielle. On considère les matrices  $a_1, \dots, a_n$  comme des matrices constantes dans  $\mathbb{R}^{d \times d}$  muni de la norme euclidienne  $\|\cdot\|_{d \times d}$ . On a alors

$$\|a\|_{d \times d}^2 = \sum_{j=1}^d \sum_{i=1}^d a_{ij}^2 = \text{trace}(a^t a)$$

On pose  $S_n = a_1 \times \dots \times a_n$  et  $F_n = E(S_n^t S_n)$ . On a alors

$$\begin{aligned} F_{n+1} &= E(S_{n+1}^t S_{n+1}) \\ &= E(a_{n+1} S_n^t S_n a_{n+1}) \\ &= E(a_{n+1} E(S_n^t S_n)^t a_{n+1}) \end{aligned}$$

Grace à l'indépendance de  $S_n$  et  $a_{n+1}$ . Ainsi, on a la relation suivante

$$F_{n+1} = T F_n$$

où T est définie de  $\mathbb{R}^{d \times d}$  dans  $\mathbb{R}^{d \times d}$  qui associe à chaque matrice A la matrice  $E(a_1 A^t a_1)$ . En particulier, on a

$$\begin{aligned} E\|a_1 \dots a_n\|_{d \times d}^2 &= E\|a_n \dots a_1\|_{d \times d}^2 \\ &= \text{trace}(F_n) = \text{trace}(T^n I) \\ &\leq d^2 \rho(T)^n, \end{aligned}$$

où  $I$  est la matrice identité. Il est clair que dès que  $\rho(T)$  est strictement inférieur à 1, l'exposant de Lyapounov est négatif. Soit maintenant  $E_{ij}$  la matrice dont tout ses sont nuls sauf l'élément  $(e_{ij})$  qui vaut 1, alors dans la base  $(E_{11}, E_{12}, \dots, E_{21}, E_{22}, \dots, E_{dd})$  de  $\mathbb{R}^{d \times d}$ ,  $T$  s'écrit comme  $E(a_1 \otimes^t a_1)$ . Cette condition ainsi que  $E(\log^+ \|b_0\|) < \infty$  assurent l'existence d'une solution non-anticipative strictement stationnaire.

Les conditions du théorème (1.6.1) sont simples dans le cas d'une équation aux récurrence stochastiques unidimensionnelle puisque

$$\gamma = \frac{1}{n} E(\log |a_1 \dots a_n|) = \frac{1}{n} \sum_{j=1}^n E(\log |a_j|) = E(\log |a_1|)$$

On donne maintenant les conditions suffisante pour l'existence d'une solution non-anticipative strictement stationnaire pour l'équation aux récurrences stochastiques suivante

$$X_{n+1} = a_n X_n + b_n, \quad n \in \mathbb{Z}, \quad X_n \in \mathbb{R} \quad (1.10)$$

où  $((a_n, b_n))_{n \in \mathbb{Z}}$  est une suite de couples de variables aléatoires *i.i.d* à valeurs dans  $\mathbb{R}^2$ .

**Théorème 1.6.3 (6).** *Si une des conditions,  $-\infty \leq E(\log |a_0|) < 0$  et  $E(\log^+ |b_0|) < \infty$  ou  $P(a_0 = 0) > 0$  est vérifiée, alors la série infinie (1,4) converge presque sûrement et le processus  $(X_n)_{n \in \mathbb{Z}}$  est l'unique solution non-anticipative strictement stationnaire de (1,10).*

**Preuve.** Pour la première condition, la preuve de la convergence de la série (1,4) est similaire à celle dans à celle dans le cas  $d > 1$ . Si  $P(a_0 = 0) > 0$ , la série (1,4) a seulement un nombre fini de termes non nuls, ainsi la série converge presque sûrement. De plus

$\prod_{m=1}^n |a_{m-1}| = 0$  presque sûrement lorsque  $n$  est suffisamment grand. Ceci et l'inégalité dans

(1,5) impliquent que  $|X_n(X) - X_n| \xrightarrow{p.s} 0$  si  $n \rightarrow \infty$

On montre que maintenant que  $P(a_0 = 0) > 0$  est satisfaite, le processus  $(X_n)_{n \in \mathbb{Z}}$  défini par (1,4) est l'unique solution non-anticipative strictement stationnaire pour l'équation (1,10). En effet, si  $(X'_n)_{n \in \mathbb{Z}}$  est autre solution pour (1,10), alors par l'inégalité triangulaire, on en

déduit que

$$\begin{aligned}
|X'_n - X_n| &= |a_{n-1}| \cdot |X'_{n-1} - X_{n-1}| \\
&\vdots \\
&= \left| \prod_{j=1}^k a_{n-j} \right| \cdot |X'_{n-k} - X_{n-k}| \\
&\leq \left| \prod_{j=1}^k a_{n-j} \right| \cdot |X'_{n-k}| + \left| \prod_{j=1}^k a_{n-j} \right| \cdot |X_{n-k}|
\end{aligned}$$

Du fait que  $P(a_0 = 0) > 0$  et la série (1,4) contient un nombre fini de termes non nuls, on en déduit que  $\prod_{j=1}^k |a_{n-j}| \xrightarrow{p.s} 0$  si  $k \rightarrow \infty$ . De plus, comme  $(X'_n)_{n \in \mathbb{Z}}$  et  $(X_n)_{n \in \mathbb{Z}}$  sont strictement stationnaire, alors

$$\left| \prod_{j=1}^k a_{n-j} \right| \cdot |X'_{n-k}| \xrightarrow{p.s} 0 \text{ et } \left| \prod_{j=1}^k a_{n-j} \right| \cdot |X_{n-k}| \xrightarrow{p.s} 0 \text{ si } k \rightarrow \infty$$

Par conséquent,  $\forall n \in \mathbb{Z}$ ,  $X'_n = X_n$  p.s., c'est-à-dire  $(X_n)_{n \in \mathbb{Z}}$  est l'unique solution non-anticipative strictement stationnaire de (1,10).

## 1.7 Moments de la loi stationnaire

Dans cette partie, on s'intéresse aux moments d'ordre supérieur de la solution non-anticipative strictement stationnaire de l'équation aux récurrences stochastique. Le lemme suivant est utile pour les prochains résultats.

**Lemme 1.7.1. ( Inégalité de Jensen )**[21] *Soit  $f : \mathbb{R} \rightarrow \mathbb{R}_+$  une fonction convexe, alors pour toute variable aléatoire  $X$  telle que  $E(X) < \infty$ ,*

$$E(f(X)) \geq f(E(X)).$$

Le résultat suivant donne les moments d'ordre supérieurs de la loi stationnaire de l'équation aux recurrences stochastique suivante

$$X_n = a_{n-1}X_{n-1} + b_{n-1}, \quad n \in \mathbb{N} \quad X \in \mathbb{R}$$

**Théorème 1.7.1 (24).** *Supposons qu'il existe  $p \in [1, +\infty[$  tel que  $E(|a_0|^p) < 1$  et  $E(|b_0|^p) < \infty$ , alors  $E(|X|^p) < \infty$  et la série infinie (1.9) converge en moment d'ordre*

$p$ . Si la variable initiale  $X_0$  est telle que  $E(|X_0|^p) < \infty$ , alors  $(X_n)_{n \in \mathbb{N}}$  définie par (1.7) converge vers  $X$  en moment d'ordre  $p$ , en particulier  $E(|X_n|^p) \rightarrow E(|X|^p)$  si  $n \rightarrow \infty$ . Les moments  $E(X^m)$  sont déterminés par :

$$E(X^m) = \sum_{k=0}^m \binom{m}{k} E(a^k b^{m-k}) E(X^k), \quad m = 1, 2, \dots, [p] \quad (1.11)$$

où  $[p]$  est la partie entière de  $p$ .

**Preuve.** Voir Vervaat, W. (1979)

## 1.8 Modèles à volatilité stochastique

La plupart des modèles pour les rendements financiers sont de la forme :

$$\epsilon_t = \sigma_t \eta_t, \quad t \in \mathbb{Z}$$

Où  $(\eta_t)_t$  est une suite des variables aléatoires symétriques *i.i.d* et  $(\sigma_t)_t$  est un processus stochastique non négatif tel que pour un  $t$  fixé  $\eta_t$  et  $\sigma_t$  sont indépendantes.

La présence de volatilité stochastique implique que les rendements ne sont pas nécessairement indépendantes au fil du temps. L'hypothèse standard pour le bruit  $\eta_t \sim i.i.dN(1,0,1)$  est indépendante du processus de déviation standard  $\sigma_t$ .

La volatilité, comme aspect des séries financières est constatée depuis longtemps. Cependant il y a peu de temps, depuis le fameux papier de Engle, que sa modélisation a connu un développement satisfaisant et remarquable. Dès lors, plusieurs approches ont été conçues pour d'écrire et modéliser l'évolution dans le temps de la volatilité d'une série chronologique financière. La modélisation à volatilité stochastique en constitue une.

Ce modèle peut capturer et d'écrire le phénomène de volatilité d'une manière plus flexible que les modèles de volatilité déterministe.

Le modèle est tel que :

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

Ce modèle est appelé le Hétéroscédasticité conditionnelle autorégressive (ARCH processus). Le processus dépend de son passé. Le terme "hétéroscédasticité conditionnelle" signifie que la variance (conditionnelle à l'information disponible) est variable et dépend des valeurs anciennes du processus.

Il est assez naturel de définir  $\sigma_t^2$ , non seulement en tant que moyenne pondérée des  $\epsilon_t^2$ , mais aussi de passé de  $\sigma_t^2$ . Les données empiriques montrent que l'ordre de ARCH élevé doit être choisi afin de capturer la dynamique de la variance conditionnelle. Cela a conduit à

l'introduction du modèle ARCH généralisée (GARCH) en 1986 par Bollerslev  
Le processus de volatilité :

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2$$

Où les  $\alpha_i$  et les  $\beta_j$  sont des paramètres non négatifs. Ce modèle réduit le nombre de paramètres à estimer

# Chapitre 2

## Processus conditionnellement hétéroscédastiques

Dans ce chapitre, nous introduisons une classe importante de modèles de l'hétéroscédasticité conditionnelle qui est le modèle GARCH.

### 2.1 Présentation du modèle GARCH

Le modèle GARCH (Generalized Autoregressive Conditional Heteroskedasticity) à été introduit en 1986 par Bollerslev. Ce modèle consiste à écrire la variance conditionnelle du processus GARCH en fonction de son passé.

Soit un modèle linéaire autorégressif exprimé sous la forme suivante :

$$Y_t = E[Y_t/Y_{t-1}] + \epsilon_t$$

où  $\epsilon$  est un bruit blanc faible, tel que  $E[\epsilon_t] = 0$ ,  $E[\epsilon_t \epsilon_s] = 0$  si  $t \neq s$  et  $E[\epsilon_t / \epsilon_{t-1}] = 0$

**Définition 2.1.1.** Le processus  $(\epsilon_t)_t$  satisfait un modèle GARCH(p,q) si

$$\epsilon_t = \sigma_t \eta_t \tag{2.1}$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2 \tag{2.2}$$

$$\text{avec } \alpha_0 > 0, \alpha_i \geq 0, \beta_j \geq 0, i = 1, \dots, q, j = 1, \dots, p \tag{2.3}$$

où  $(\eta_t)_t$  est une suite de variables aléatoires *i.i.d* :  $\mathbb{E}(\eta_t) = 0$  ,  $\mathbb{E}(\eta_t^2) = 1$

L'équation (1,2) peut être écrite sous la forme symbolique suivante :

$$\mathfrak{B}(B)\sigma_t^2 = \alpha_0 + \mathfrak{A}(B)\epsilon_t^2, t \in \mathbb{Z} \quad (2.4)$$

où B est l'opérateur de retard ( $B^i\epsilon_t^2 = \epsilon_{t-i}^2$  et  $B^i\sigma_t^2 = \sigma_{t-i}^2, \forall i \in \mathbb{N}$ ),  $\mathfrak{A}$  et  $\mathfrak{B}$  sont deux polynômes de degrés p et q, respectivement :

$$\mathfrak{A}(B) = \sum_{i=1}^q \alpha_i B^i \quad (2.5)$$

$$\mathfrak{B}(B) = 1 - \sum_{j=1}^p \beta_j B^j \quad (2.6)$$

Si  $\mathfrak{B}(z) = 1$ , alors on a :

$$\sigma_t^2 = \alpha_0 + \alpha_1\epsilon_{t-1}^2 + \dots + \alpha_q\epsilon_{t-q}^2$$

et le processus  $(\epsilon_t)_{t \in \mathbb{Z}}$  est dit processus ARCH(q).

Le processus GARCH est un modèle ARCH généralisé

En pratique, le modèle GARCH permet d'avoir moins de paramètres à estimer que le modèle ARCH.

### Exemples 2.1.1.

Si P=1, q=0 c'est un modèle ARCH(1)

Si p=0, q=1 c'est un modèle GARCH(0,1)

Si p=1, q=1 c'est un modèle GARCH(1,1)

Si p=2, q=1 c'est un modèle GARCH(2,1)...etc

**Propriétés 2.1.1 (3).** Si  $(\epsilon_t)_t$  est un processus GARCH(p,q), alors

$$\begin{aligned} \mathbb{E}(\epsilon_t) &= 0 \\ \mathbb{E}(\epsilon_t \setminus \mathfrak{F}_{t-1}) &= 0 \\ Cov(\epsilon_t, \epsilon_{t+h}) &= \sigma_h, \quad \forall h > 0 \\ Cov(\epsilon_t, \epsilon_{t+h} \setminus \mathfrak{F}_{t-1}) &= 0 \end{aligned}$$

**Remarque 2.1.** Les inégalités dans (2,3) sont imposées pour garantir que

$$\sigma_t^2 > 0, p.s. \quad (2.7)$$

Dans le cas du modèle GARCH(1,1) avec :

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

les conditions suffisantes, mais pas nécessaires pour (2,7) sont:

$$\alpha_0 > 0, \alpha_i \geq 0, \beta_j \geq 0. \quad (2.8)$$

En d'autres termes, par substitution on montre que :

$$\begin{aligned} \sigma_t^2 &= \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 (\alpha_0 + \alpha_1 \epsilon_{t-2}^2 + \beta_1 \sigma_{t-2}^2) \\ &\vdots \\ &= \alpha_0 (1 - \beta_1)^{-1} + \alpha_1 \epsilon_{t-1}^2 + \alpha_1 \beta_1 \sum_{j=0}^{\infty} \beta_1^j \epsilon_{t-j-2}^2 \end{aligned}$$

avec  $\alpha_0 > 0, \alpha_1 \geq 0, 0 \leq \beta \leq 1$  sont des conditions nécessaires et suffisantes pour (2,7)

en supposant que la somme  $\sum_{j=0}^{\infty} \beta_1^j \epsilon_{t-j-2}^2$  est convergente.

## 2.2 Stationarité des modèles GARCH :

La définition du modèle GARCH est simple mais son étude mathématique est difficile notamment le problème de la stationnarité;

La condition nécessaire et suffisante pour l'existence d'une unique solution stationnaire est due à Nelson [18] pour le modèle GARCH(1,1) et à Bougerol et Picard [5] pour le modèle GARCH (p,q).

On considère d'abord le modèle GARCH (1,1) qui peut être étudié d'une manière explicite que le cas général.

### 2.2.1 Etude d'un modèle GARCH (1,1)

Soit un modèle GARCH (1,1) défini par :

$$\epsilon_t = \sigma_t \eta_t \quad (2.9)$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \quad (2.10)$$

avec  $\alpha_0 > 0, \alpha_1 \geq 0, \beta_1 \geq 0$  et  $(\eta_t)_{t \in \mathbb{Z}}$  est une suite de variables aléatoires *i.i.d*  $N(0,1)$ .

Le modèle GARCH (1,1) peut être écrit comme équation aux récurrences stochastiques :

$$X_t = a_{t-1} X_{t-1} + b_{t-1}, \quad t \in \mathbb{Z}$$

où  $(a_t, b_t)_{t \in \mathbb{Z}}$  est une suite de variables aléatoires *i.i.d.*  $(X_t)_{t \in \mathbb{Z}}$  est une suite de vecteurs aléatoires,  $(a_t)_{t \in \mathbb{Z}}$  est une suite de matrices aléatoires *i.i.d.* et  $(b_t)_{t \in \mathbb{Z}}$  une suite de vecteurs aléatoires *i.i.d.* définis par

$$X_t = \begin{pmatrix} \epsilon_t^2 \\ \sigma_t^2 \end{pmatrix}, a_{t-1} = \begin{pmatrix} \alpha_1 \eta_t^2 & \beta_1 \eta_t^2 \\ \alpha_1 & \beta_1 \end{pmatrix}, b_{t-1} = \begin{pmatrix} \alpha_0 \eta_t^2 \\ \alpha_0 \end{pmatrix}$$

Nelson (1990)[23] a donné une condition suffisante pour que le modèle GARCH (1,1) admette une unique solution non-anticipative strictement stationnaire et ergodique.

**Théorème 2.2.1 (23).** *Soit un modèle GARCH(1,1) supposons que*

$$-\infty \leq E(\log(\alpha_1 \eta_0^2 + \beta_1)) < 0 \quad (2.11)$$

alors

1. La série infinie suivante :

$$\sigma_t^2 = \alpha_0 \sum_{m=0}^{\infty} \prod_{j=1}^m (\alpha_1 \eta_{t-j}^2 + \beta_1) \quad (2.12)$$

converge presque sûrement. Le processus  $(\sigma_t^2)_{t \in \mathbb{Z}}$  est strictement stationnaire et ergodique et le processus  $(\epsilon_t)_{t \in \mathbb{Z}}$  défini par  $\epsilon_t = \sigma_t \eta_t$  est l'unique solution non-anticipative strictement stationnaire et ergodique .

2. Supposons que  $(\sigma_t^2)_{t \in \mathbb{Z}}$  est strictement stationnaire . Soit  $p \in [1, +\infty]$  tel que

$$E(\alpha_1 \eta_0^2 + \beta_1)^p < 1 \quad (2.13)$$

alors  $\exists m : 1 \leq m \leq [p]$  tel que  $E(\sigma^{2m}) < \infty$  et

$$E(\sigma^{2m}) = (1 - (\alpha_1 \eta_0^2 + \beta_1)^m)^{-1} \sum_{k=0}^{m-1} \binom{m}{k} E(\alpha_1 \eta_0^2 + \beta_1)^k \alpha_0^{m-k} E(\sigma^{2k}) \quad (2.14)$$

**Preuve.** De (2,10), on a :

$$\begin{aligned} \sigma_t^2 &= \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \\ &= \alpha_0 + (\alpha_1 \eta_{t-1}^2 + \beta_1) \sigma_{t-1}^2 \end{aligned}$$

On a aussi la représentation de  $\sigma_t^2$  sous l'équation aux récurrences stochastique :

$$X_t = a_{t-1} X_{t-1} + b_{t-1}, \quad t \in \mathbb{Z} \quad (2.15)$$

avec  $X_t = \sigma_t^2$ ,  $a_{t-1} = \alpha_1 \eta_{t-1}^2 + \beta_1$  et  $b_{t-1} = \alpha_0$ .

On a  $E(\log^+ |b_0|) < \infty$  et par hypothèse  $E(\log |a_0|) = E(\log(\alpha_1 \eta_0^2 + \beta_1)) < 0$ .

Ainsi, par le théorème 1.6.3 , la série (2,12) converge presque sûrement et  $(\sigma_t^2)_{t \in \mathbb{Z}}$  est l'unique solution non-anticipative strictement stationnaire et ergodique de l'équation aux

réurrences stochastique (2,15). Par conséquent, le processus  $(\epsilon_t)_{t \in \mathbb{Z}}$  est défini par :

$$\epsilon_t = (\alpha_0 \sum_{m=0}^{\infty} \prod_{j=1}^m (\alpha_1 \eta_{t-j}^2 + \beta_1))^{\frac{1}{2}} \eta_t \quad (2.16)$$

est l'unique solution non-anticipative strictement stationnaire et ergodique du modèle GARCH (1,1).

Supposons que  $E(|a_0|^p) = E(\alpha_1 \eta_0^2 + \beta_1)^p < 1$ , avec  $p \in [1, \infty]$ , alors par le théorème 1.7.1, on déduit que

$$E(\sigma^{2m}) = (1 - E((\alpha_1 \eta_0^2 + \beta_1)^m))^{-1} \sum_{k=0}^{m-1} \binom{m}{k} E(\alpha_1 \eta_0^2 + \beta_1)^k \alpha_0^{m-k} E(\sigma^{2k}) \quad (2.17)$$

avec  $1 \leq m \leq [p]$ .

Le resultat suivant montre la non-stationnarité du processus GARCH(1,1).

**Corollaire 2.2.1 (14).** *Soit un modèle GARCH (1,1) avec  $t \geq 1$ .*

*Si  $E(\log(\alpha_1 \eta_0^2 + \beta_1)) > 0$ , alors :*

$$\lim_{t \rightarrow \infty} \sigma_t^2 = +\infty, \quad p.s. \quad (2.18)$$

*Si, en plus,  $E(|\log(\eta_t^2)|) < \infty$ , alors :*

$$\lim_{t \rightarrow \infty} \epsilon_t^2 = +\infty, \quad p.s. \quad (2.19)$$

**Preuve.** On a:

$$\sigma_t^2 \geq \alpha_0 \sum_{m=1}^t \prod_{j=1}^{m-1} (\alpha_1 \eta_{t-j}^2 + \beta_1) \geq \alpha_0 (\alpha_1 \eta_{t-1}^2 + \beta_1) \cdots (\alpha_1 \eta_1^2 + \beta_1),$$

alors

$$\liminf_{t \rightarrow \infty} \frac{1}{t} \log \sigma_t^2 \geq \liminf_{t \rightarrow \infty} \frac{1}{t} \sum_{m=1}^{t-1} \log(\alpha_1 \eta_{t-j}^2 + \beta_1) = E(\log(\alpha_1 \eta_0^2 + \beta_1))$$

Par conséquent  $\log \sigma_t^2 \rightarrow +\infty$  et  $\sigma_t^2 \rightarrow +\infty$  p.s

De la même manière, on a:

$$\begin{aligned} \liminf_{t \rightarrow \infty} \frac{1}{t} \log \epsilon_t^2 &= \liminf_{t \rightarrow \infty} \frac{1}{t} (\log \sigma_t^2 + \log \eta_t^2) \\ &\geq E(\log(\alpha_1 \eta_0^2 + \beta_1)) + \liminf_{t \rightarrow \infty} \frac{1}{t} \log \eta_t^2 = E(\log(\alpha_1 \eta_0^2 + \beta_1)), \end{aligned}$$

alors  $\log \epsilon_t^2 \rightarrow +\infty$  et  $\epsilon_t^2 \rightarrow +\infty, p.s$

Bolleslev (1986)[4] a établi une condition nécessaire et suffisante pour l'existence d'une solution stationnaire au second ordre du modèle GARCH(1,1).

**Théorème 2.2.2 (4).** *Le modèle GARCH (1,1) admet une solution non-anticipative stationnaire au second ordre donnée par (2.16), avec  $E(\epsilon_t) = 0$ ,  $var(\epsilon_t) = \alpha_0 + (1 - \alpha_1 - \beta_1)^{-1}$  et  $cov(\epsilon_t, \epsilon_s) = 0$  pour  $t \neq s$ , ssi*

$$\alpha_1 + \beta_1 < 1 \tag{2.20}$$

**Preuve.** Si  $(\epsilon_t)_{t \in \mathbb{Z}}$  au second ordre, alors on a :

$$E(\epsilon_t^2) = E(E(\epsilon_t^2 | \mathfrak{F}_{t-1})) = E(\sigma_t^2) = \alpha_0 + (\alpha_1 + \beta_1)E(\epsilon_{t-1}^2)$$

Par conséquent,

$$(1 - \alpha_1 - \beta_1)E(\epsilon_t^2) = \alpha_0$$

ce qui donne pour  $1 - \alpha_1 - \beta_1 > 0$ ,  $E(\epsilon_t^2) = \frac{\alpha_0}{1 - \alpha_1 - \beta_1}$

Inversement, supposons que  $1 - \alpha_1 - \beta_1 < 1$ . Par l'inégalité de Jensen, on a :

$$E(\log(\alpha_1 \eta_0^2 + \beta_1)) \leq \log(E(\alpha_1 \eta_0^2 + \beta_1)) = \log(\alpha_1 + \beta_1) < 0$$

c'est-à-dire la condition (2,11) est satisfaite. Il suffit alors de montrer que la solution  $(\epsilon_t)_{t \in \mathbb{Z}}$  définie par (2,16) admet une variance finie. En effet, Notons que :

$$\prod_{j=1}^{m+1} (\alpha_1 \eta_{t-j}^2 + \beta_1) = (\alpha_1 \eta_{t-1}^2 + \beta_1) \prod_{j=1}^m (\alpha_1 \eta_{t-j-1}^2 + \beta_1)$$

et puisque  $(\eta_t)_{t \in \mathbb{Z}}$  est une suite *i.i.d*, l'espérance de  $\prod_{j=1}^{m+1} (\alpha_1 \eta_{t-j}^2 + \beta_1)$  ne dépend pas de t. Par conséquent,

$$\begin{aligned} E\left(\prod_{j=1}^{m+1} (\alpha_1 \eta_{t-j}^2 + \beta_1)\right) &= (\alpha_1 + \beta_1) E\left(\prod_{j=1}^m (\alpha_1 \eta_{t-j}^2 + \beta_1)\right) \\ &\vdots \\ &= (\alpha_1 + \beta_1)^m. \end{aligned}$$

On a ainsi;

$$\begin{aligned}
E(\epsilon_t^2) = E(\eta_t^2).E(\sigma_t^2) &= E(\alpha_0 \sum_{m=0}^{\infty} \prod_{j=1}^m (\alpha_1 \eta_{t-j}^2 + \beta_1)) \\
&= \alpha_0 \sum_{m=0}^{\infty} E(\prod_{j=1}^m (\alpha_1 \eta_{t-j}^2 + \beta_1)) \\
&= \alpha_0 \sum_{m=0}^{\infty} (\alpha_1 + \beta_1)^m \\
&= \alpha_0 (1 - \alpha_1 - \beta_1)^{-1}.
\end{aligned}$$

Ceci montre que la solution du modèle GARCH (1,1) est stationnaire au second ordre. De plus cette solution est un bruit blanc parce que  $E(\epsilon_t) = E(E(\epsilon_t | F_{t-1})) = 0$  et  $\forall h > 0$ ,

$$cov(\epsilon_t, \epsilon_{t-h}) = E(\epsilon_t \epsilon_{t-h}) = E(E(\epsilon_t \epsilon_{t-h} | \mathfrak{F}_{t-1})) = E(\epsilon_{t-h} E(\epsilon_t | \mathfrak{F}_{t-1})) = 0$$

Notons que depuis les conditions du théorème 2.2.1, la stationnaire stricte de  $(\sigma_t^2)_{t \in \mathbb{Z}}$  implique la stationnarité stricte du processus  $(\epsilon_t^2, \sigma_t^2)_{t \in \mathbb{Z}} = (\sigma_t^2(\eta_t^2, 1))_{t \in \mathbb{Z}}$ . Par la constuction du processus  $(\epsilon_t)_{t \in \mathbb{Z}}$ , le processus  $(\epsilon_t, \sigma_t)_{t \in \mathbb{Z}}$  est strictement stationnaire.

## 2.2.2 Modèle GARCH (p,q)

Dans le cas général d'un modèle GARCH(p,q), on a la représentation vectorielle (équation aux récurences stochastique multivariée) suivante :

$$X_t = a_{t-1} X_{t-1} + b_{t-1}, \quad t \in \mathbb{Z} \quad (2.21)$$

avec

$$X_t = (\epsilon_t^2, \dots, \epsilon_{t-q+1}^2, \sigma_t^2, \dots, \sigma_{t-p+1}^2)' \in \mathbb{R}^{p+q}, b_{t-1} = (\alpha_0 \eta_t^2, 0, \dots, \alpha_0, 0, \dots, 0)' \in \mathbb{R}^{p+q}$$

$$\alpha_{t-1} = \begin{pmatrix} \alpha_1 \eta_t^2 & \alpha_2 \eta_t^2 & \dots & \alpha_{q-1} \eta_t^2 & \alpha_q \eta_t^2 & \beta_1 \eta_t^2 & \beta_2 \eta_t^2 & \dots & \beta_{p-1} \eta_t^2 & \beta_p \eta_t^2 \\ 1 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 & 0 & & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 & 0 & 0 & \dots & 0 & 0 \\ \alpha_1 & \alpha_2 & \dots & \alpha_{q-1} & \alpha_q & \beta_1 & \beta_2 & \dots & \beta_{p-1} & \beta_p \\ 0 & 0 & \dots & 0 & 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 1 & 0 \end{pmatrix} \quad (2.22)$$

telle que  $a_t \in \mathbb{R}^{p+q} x \in \mathbb{R}^{p+q}$ . Dans le cas d'un modèle ARCH(q), le vecteur  $X_t$  se réduit à  $\epsilon_t^2$  et ses (q-1) premières valeurs passées, la matrice  $a_t$  se réduit à une sous-matrice de la matrice définie par (2,22).

Le résultat suivant donne les conditions suffisantes pour que le modèle défini par (2,1) et (2,2) admet une unique solution non-anticipative strictement stationnaire et ergodique.

**Théorème 2.2.3 (14).** *L'équation aux récurrences stochastique admet une unique solution non-anticipative, strictement stationnaire et ergodique définie par*

$$X_t = \sum_{k=0}^{\infty} \left( \prod_{j=1}^k a_{t-j} \right) b_{t-k-1} \quad (2.23)$$

si l'exposant de lyapounov  $\gamma$  de la suite  $(a_t)_{t \in \mathbb{Z}}$  est strictement négatif.

Par conséquent le modèle (2,1) et (2,2) admet une unique solution non-anticipative strictement stationnaire et ergodique.

**Preuve.** On utilise la norme matricielle définie  $\|a\| = \sum_{i,j} |a_{ij}|$ . Notons que la suite  $(a_t)_{t \in \mathbb{Z}}$  vérifie la condition  $E(\log^+ \|a_t\|) < \infty$ . En effet, puisque  $E(\eta_t^2) = 1$ , alors les composantes de la matrice  $a_t$  sont toutes intégrables et par conséquent, on a :

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

De plus, la suite  $(b_t)_{t \in \mathbb{Z}}$  vérifie la condition  $E(\log^+ |b_t|) < \infty$ . Les conditions du théorème 1.6.1 sont alors satisfaites.

Les conséquences de la condition de stationnarité stricte sont données dans la proposition suivante :

**Proposition 2.2.1 (13).** *Si l'exposant de lyapounov  $\gamma$  de la suite  $(a_t)_{t \in \mathbb{Z}}$  est strictement négatif, alors les propriétés suivantes sont équivalentes :*

1.  $\sum_{i=1}^p \beta_i < 1$ .
2. Les racines de

$$1 - \beta_1 z - \dots - \beta_p z^p = 0$$

sont en dehors du disque unité.

3.  $\rho(B_0) < 1$ , avec  $\rho(B_0)$  est le rayon spectral de la matrice

$$B_0 = \begin{pmatrix} \beta_1 & \beta_2 & \dots & \beta_{p-1} & \beta_p \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{pmatrix}$$

**Preuve.** Comme les éléments de la matrice  $a_t$  sont positifs, alors il est clair que l'exposant de lyapounov de la suite  $(a_t)_{t \in \mathbb{Z}}$  est plus grand que l'exposant de lyapounov de la suite des matrices constantes obtenues en remplaçant par zéro les composantes des  $q$  premières lignes et les  $q$  premières colonnes de la matrice  $a_t$ . La matrice obtenue a les mêmes valeurs propres non nulles et le même rayon spectral que  $B_0$ . Pour une matrice constante  $A$ , on a

$$\lim_{t \rightarrow \infty} \frac{1}{t} \log \|A^t\| = \log \rho(A) \quad (2.24)$$

alors  $\log \rho(B_0) \leq \gamma < 0$ , ce qui implique que  $\rho(B_0) < 1$ .

En calculant le déterminant de  $(\lambda I_p - B_0)$  par rapport à la dernière colonne, on obtient

$$\det(\lambda I_p - B_0) = \lambda^p - \lambda^{p-1} \beta_1 - \dots - \beta_p = \lambda^p \mathfrak{B}\left(\frac{1}{\lambda}\right)$$

avec  $\mathfrak{B}(z) = 1 - \beta_1 z - \dots - \beta_p z^p$ .

L'équivalence entre (2) et (3) est évidente.

On montre maintenant que (1)  $\Leftrightarrow$  (2).

On a  $\mathfrak{B}(0) = 1$  et  $\mathfrak{B}(1) = 1 - \sum_{i=1}^p \beta_i$ . Par conséquent, si  $\sum_{i=1}^p \beta_i \geq 1$ , alors  $\mathfrak{B}(1) \leq 0$  et par le théorème de Rolle,  $\exists z_0 \in ]0, 1[$  t.q  $\mathfrak{B}(z_0) = 0$ . On a montré alors (2)  $\Rightarrow$  (1)

Inversement, si  $\sum_{i=1}^p \beta_i < 1$  et  $\mathfrak{B}(z_0) = 0$  avec  $|z_0| \leq 1$ , alors

$$1 = \sum_{i=1}^p \beta_i z_0^i = \left| \sum_{i=1}^p \beta_i z_0^i \right| \leq \sum_{i=1}^p \beta_i |z_0|^i \leq \sum_{i=1}^p \beta_i < 1$$

d'où la contradiction. On a montré alors (1)  $\Rightarrow$  (2).

**Corollaire 2.2.2 (14).** *Si l'exposant de lyapounov  $\gamma$  de la suite  $(a_t)_{t \in \mathbb{Z}}$  définie par (2,24) est strictement négatif, alors*

$$\exists s > 0 \text{ tel que } E(\sigma_t^{2s}) < \infty \text{ et } E(\epsilon_t^{2s}) < \infty.$$

Le résultat suivant donne une autre caractérisation de la stationnarité stricte des modèles GARCH.

**Lemme 2.2.1 (14).** *Soit  $(a_t)_{t \in \mathbb{Z}}$  une suite de matrices aléatoire i.i.d, alors l'exposant de lyapounov  $\gamma$  de la suite  $(a_t)_{t \in \mathbb{Z}}$  est strictement négatif, ssi*

$$\exists s > 0, \exists k_0 \geq 1 \text{ tels que } \delta = E(\|a_{k_0} a_{k_0-1} \dots a_1\|^s) < 1.$$

**Preuve.** Supposant que  $\gamma < 0$ . On a :

$$\gamma = \inf_k \frac{1}{k} E(\log \|a_k a_{k-1} \cdots a_1\|),$$

alors  $\exists k_0 \geq 1$  tel que  $E(\log \|a_k a_{k-1} \cdots a_1\|) < 0$ .

De plus, en utilisant la norme  $\|A\| = \sum_{i,j} |A_{ij}|$ , on obtient

$$\begin{aligned} E(\|a_{k_0} a_{k_0-1} \cdots a_1\|) &= \|E(a_{k_0} a_{k_0-1} \cdots a_1)\| \\ &= \|E((a_1)^{k_0})\| \\ &\leq (E\|a_1\|)^{k_0} \end{aligned}$$

Rappelons que si  $\exists r > 0$  tel que  $E(X^r) < \infty$  et  $E(\log X) < 0$ , avec  $X$  est une variable aléatoire positive, alors  $\exists s > 0$  tel que  $E(X^s) < 1$ . en utilisant ce dernier résultat, on conclut que

$$\gamma < 0 \implies \exists s > 0, \exists k_0 \geq 1 \text{ tel que } E(\|a_{k_0} a_{k_0-1} \cdots a_1\|^s) < 1$$

si  $\exists s > 0, \exists k_0 \geq 1$  tel que  $\delta < 1$ , alors par l'inégalité de Jensen, on a :

$$\gamma \leq \frac{1}{k_0} E(\|a_{k_0} a_{k_0-1} \cdots a_1\|) \leq \frac{1}{s k_0} \log \delta < 0.$$

**Preuve du corollaire 2.2.2** La preuve du lemme 2.2.1 montre que le nombre  $s$  peut être inférieure à 1. Soit  $s \in ]0,1]$ , alors la solution strictement stationnaire définie par (2,23) satisfait

$$\begin{aligned} E(\|X_t\|^s) &\leq \sum_{k=0}^{\infty} E(\|a_{t-1} \cdots a_{t-k}\|^s) E(\|b_{t-k-1}\|^s) = E(\|b_0\|^s) \sum_{i=0}^{\infty} E(\|a_{t-1} \cdots a_{t-k}\|^s) \\ &= E(\|b_0\|^s) \left(1 + \sum_{k=1}^{k_0} E(\|a_{t-1} \cdots a_{t-k}\|^s) + \sum_{k=k_0+1}^{\infty} E(\|a_{t-1} \cdots a_{t-k}\|^s) E(\|b_{t-k-1}\|^s)\right) \\ &\leq E(\|b_0\|^s) \left(1 + \sum_{k=1}^{k_0} (E\|a_0\|^s)^k\right) + \sum_{k=k_0+1}^{2k_0} E(\|a_{t-1} \cdots a_{t-k}\|^s) + \sum_{k=2k_0+1}^{\infty} E(\|a_{t-1} \cdots a_{t-k}\|^s) \\ &\leq E(\|b_0\|^s) \left(1 + \sum_{k=1}^{k_0} (E\|a_0\|^s)^k + \delta \sum_{k=1}^{k_0} (E\|a_0\|^s)^k + \sum_{k=2k_0+1}^{\infty} E(\|a_{t-1} \cdots a_{t-k}\|^s)\right) \\ &\vdots \\ &\leq E(\|b_0\|^s) \left(1 + \sum_{j=0}^{\infty} \delta^j \sum_{k=1}^{k_0} (E\|a_0\|^s)^k\right) < \infty. \end{aligned}$$

On déduit que  $\sigma_t^{2s} \leq \|X_t\|^s$  et  $\epsilon_t^{2s} \leq \|X_t\|^s$  et par conséquent  $E(\sigma_t^{2s}) < \infty$  et  $E(\epsilon_t^{2s}) < \infty$ . Le théorème suivant donne les conditions nécessaires et suffisantes de stationnarité du second ordre des processus GARCH(p,q).

**Théorème 2.2.4 (4).** *Si un processus GARCH (p,q) est stationnaire au second ordre, alors*

$$\sum_{i=1}^q \alpha_i + \sum_{j=1}^p \beta_j < 1 \quad (2.25)$$

*Inversement, si la condition (2,25) est satisfaite, alors l'unique solution non-anticipative strictement stationnaire du modèle GARCH(p,q) est un bruit blanc faible (stationnaire au second ordre).*

**Preuve.** On montre d'abord que la condition (2,25) est nécessaire. Soit  $(\epsilon_t)_{t \in \mathbb{Z}}$  un processus GARCH (p,q) stationnaire au second ordre, alors

$$E(\epsilon_t^2) = E(E(\epsilon_t^2 \mathfrak{F}_{t-1})) = E(\sigma_t^2)$$

ne dépend pas de t. En prenant l'espérance mathématique des deux cotés de (2,2), on obtient

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

c'est-à-dire

$$(1 - \sum_{i=1}^q \alpha_i - \sum_{j=1}^p \beta_j) E(\epsilon_t^2) = \alpha_0,$$

ce qui montre (2,25).

Maintenant, on montre la condition suffisante. Supposons que la condition (2,25) est satisfaite et on construit une solution stationnaire satisfaisant un modèle GARCH (p,q).

Pour  $t, k \in \mathbb{Z}$ , on définit le vecteur  $X_k(t)$  comme suit

$$X_k(t) = \begin{cases} 0 & \text{si } k < 0 \\ a_{t-1}X_{k-1}(t-1) + b_{t-1} & \text{si } k \geq 0 \end{cases}$$

et  $X_k(t) - X_{k-1}(t) = \begin{cases} 0 & \text{si } k < 0 \\ b_{t-1} & \text{si } k = 0 \\ a_{t-1}(X_{k-1}(t-1) - X_{k-2}(t-1)) & \text{si } k > 0 \end{cases}$

Par itération, on obtient pour  $k > 0$ ,

$$X_k(t) - X_{k-1}(t) = a_{t-1}a_{t-2} \cdots a_{t-k}b_{t-k-1}.$$

Par conséquent,

$$E(\|X_k(t) - X_{k-1}(t)\|) = \|E(a_{t-1}a_{t-2} \cdots a_{t-k}b_{t-k-1})\|$$

Notons que tous les termes de  $a_{t-1}a_{t-2} \cdots a_{t-k}b_{t-k-1}$  sont indépendants parce que  $(\eta_t)_{t \in \mathbb{Z}}$  est une suite de variables aléatoires *i.i.d.* En posant  $a = E(a_t)$  et  $b = E(b_t)$ , on obtient pour  $k > 0$ ,

$$E(\|X_k(t) - X_{k-1}(t)\|) = \|a^k b\| = (1, 1, \dots, 1) a^k b.$$

La condition (2,25) implique que les valeurs propres de  $a$  sont strictement inférieures à 1 en valeur absolue. En effet, on peut vérifier que

$$\det(\lambda I_{p+q} - a) = \lambda^{p+q} \left( 1 - \sum_{i=1}^q \alpha_i \lambda^{-i} - \sum_{j=1}^p \beta_j \lambda^{-j} \right).$$

Si  $|\lambda| \geq 1$ , alors

$$|\det(\lambda I_{p+q} - a)| \geq \left| 1 - \sum_{i=1}^q \alpha_i \lambda^{-i} - \sum_{j=1}^p \beta_j \lambda^{-j} \right| \geq 1 - \sum_{i=1}^q \alpha_i - \sum_{j=1}^p \beta_j > 0,$$

donc  $\rho(a) < 1$ . En utilisant le résultat (2.24), on déduit que  $a^k \rightarrow 0$  si  $k \rightarrow \infty$ . De plus, pour  $t$  fixé,  $X_k(t)$  converge presque sûrement lorsque  $k \rightarrow \infty$ . Soit  $X_t$  la limite de  $(X_k(t))_k$ . Pour  $k$  fixé, le processus  $(X_k(t))_{t \in \mathbb{Z}}$  est strictement stationnaire. Le processus limite  $(X_t)_{t \in \mathbb{Z}}$  est strictement stationnaire et une solution de l'équation de récurrences stochastique (2,21).

### Remarque

1. Sous la condition (2,25) du théorème 2.2.4, l'unique solution du modèle défini par (2,1) et (2,2) est un bruit blanc de variance

$$\text{var}(\epsilon_t) = \alpha_0 \left( 1 - \sum_{i=1}^q \alpha_i - \sum_{j=1}^p \beta_j \right)^{-1}$$

2. Sous la condition (2,25) du théorème 2.2.4, on a nécessairement

$$\left( 1 - \sum_{i=1}^q \alpha_i - \sum_{j=1}^p \beta_j < 1 \right) \Rightarrow \gamma < 0.$$

parce que la solution stationnaire au second ordre donnée par le théorème 2.2.4 est aussi strictement stationnaire. En effet, si la condition (2,27) est satisfaite, alors le rayon spectral  $\rho(E(a_t))$  est strictement inférieur à 1. De plus, on a :

$$\begin{aligned} \frac{1}{k} E(\log \|a_k \cdots a_1\|) &\leq \frac{1}{k} \log E(\|a_k \cdots a_1\|) \\ &= \frac{1}{k} \log \|(E(a_1))^k\| \rightarrow \log \rho(E(a_1)) < 0, \end{aligned}$$

ce qui montre que  $\gamma < 0$

## 2.3 Moments d'ordre supérieurs

Dans cette partie, on s'intéresse aux moments d'ordres supérieurs de la solution strictement stationnaire du modèle GARCH(p,q). Le résultat suivant donne les conditions assurant l'existence des moments d'ordre  $2m$ , où  $m$  est un entier positif.

**Théorème 2.3.1 (12).** *Soit  $a^{(m)} = E(a_t^{\otimes m})$ , avec  $a_t$  définie par (2.22). Supposons que  $E(\eta_t^{2m}) < \infty$  et  $\rho(a^{(m)}) < 1$ , la série (2.23) converge dans  $L^m$  et le processus  $(\epsilon_t^2)_{t \in \mathbb{Z}}$  est strictement stationnaire admettant des moments jusqu'à l'ordre  $m$ .*

**Preuve.** Soit  $k > 0$ , alors après  $k$  itérations dans l'équation aux récurrences (2,23), on obtient :

$$X_t = \prod_{i=1}^k a_{t-i} X_{t-k} + \sum_{j=0}^{k-1} \left( \prod_{i=1}^j a_{t-i} \right) b_{t-j-1}$$

On pose  $a_{t-1,k} = \prod_{i=1}^k a_{t-i}$  et  $X_{t,k} = a_{t-1,k} b_{t-k-1}$ , avec  $a_{t-1,0} = I_{p+q}$  et  $X_{t,0} = b_{t-1}$ . Notons que

$$X_t = \sum_{k=0}^{\infty} \left( \prod_{i=1}^k a_{t-i} \right) b_{t-j-1} = \sum_{k=0}^{\infty} X_{t,k} \quad (2.26)$$

En utilisant la norme  $\|A\| = \sum_{i,j} |A_{i,j}|$  et l'égalité  $\|A\| \cdot \|B\| = \|A \otimes B\|$ , on obtient :

$$E(\|X_{t,k}\|^m) = E(\|a_{t-1,k} b_{t-k-1} \otimes \dots \otimes a_{t-1,k} b_{t-k-1}\|) = \|E(a_{t-1,k} b_{t-k-1} \otimes \dots \otimes a_{t-1,k} b_{t-k-1})\|.$$

Soit  $x \in \mathbb{R}^{p+q}$ , alors  $(ax)^{\otimes m} = a^{\otimes m} \times x^{\otimes m}$ . En utilisant ce résultat, on obtient :

$$E(\|X_{t,k}\|^m) = \|E(a_{t-1,k}^{\otimes m} b_{t-k-1}^{\otimes m})\| = \|E(a_{t-1}^{\otimes m} \dots a_{t-k}^{\otimes m} b_{t-k-1}^{\otimes m})\| \quad (2.27)$$

On pose  $b^{(m)} = E(b_t^{\otimes m})$ . De (2,27) et du fait que les matrices  $a_{t-i}$  sont indépendantes, on obtient :

$$E(\|X_{t,k}\|^m) = \|(a^{(m)})^k b^{(m)}\|.$$

En utilisant ce dernier résultat et la relation (2,26), on obtient :

$$\begin{aligned} \|X_t\|_m &= (E(\|X_t\|^m))^{\frac{1}{m}} \leq \sum_{k=0}^{\infty} \|X_{t,k}\|_m \\ &\leq \|b^{(m)}\|_m^{\frac{1}{m}} \left( \sum_{k=0}^{\infty} \|(a^{(m)})^k\|_m^{\frac{1}{m}} \right). \end{aligned}$$

Si  $\rho(a^{(m)}) < 1$ , alors  $\|(a^{(m)})^k\|$  converge vers zéro. Dans ce cas,  $X_t$  est finie presque sûrement, le processus  $(X_t)_{t \in \mathbb{Z}}$  est la solution strictement stationnaire de (2,21) et il appartient à  $L^m$ . Il est clair que  $\|\epsilon_t^2\|_m \leq \|X_t\|_m$ . Par conséquent, la condition suffisante pour l'existence de  $E(\epsilon_t^{2m})$  est alors  $\rho(a^{(m)}) < 1$ .

## 2.4 Estimation des paramètres

Les paramètres du modèle GARCH peuvent être estimés selon différentes méthodes : maximum de vraisemblance, pseudo maximum de vraisemblance, méthode des moments etc. Les méthodes généralement retenues sont celles du maximum de vraisemblance (MV) ou du pseudo-maximum de vraisemblance (PMV). L'avantage du PMV réside dans le fait que l'estimateur obtenu converge malgré une mauvaise spécification de la distribution conditionnelle des résidus, à condition que la loi spécifiée appartienne à la famille exponentielle.

### Estimateurs du MV sous l'hypothèse de normalité et Estimateurs du PMV

On reprend ici la présentation de Gouriéroux (1992). Soit un modèle tel que :

$$E(Y_t \setminus Y_{t-1}, X_t) = m_t(Y_{t-1}, X_t, \theta) = m_t(\theta) \quad (2.28)$$

$$V(Y_t \setminus Y_{t-1}, X_t) = \sigma_t(Y_{t-1}, X_t, \theta) = \sigma_t(\theta) \quad (2.29)$$

On note  $\theta$  l'ensemble des paramètres intervenant à la fois dans l'expression de la moyenne conditionnelle et de la variance conditionnelle. Nous commencerons par présenter les méthodes du MV et PMV.

### Maximum et Pseudo Maximum de Vraisemblance appliqués aux modèles GARCH

Nous présenterons de façon parallèle la méthode d'estimation du MV sous l'hypothèse de normalité de la distribution conditionnelle des résidus et la méthode d'estimation du PMV. En effet, l'idée générale des estimateurs du PMV consiste à démontrer que si l'on commet une erreur sur la distribution conditionnelle des résidus en utilisant à tort une log-vraisemblance fondée sur une loi normale, l'estimateur du MV ainsi obtenu peut tout de même être convergent si la vraie loi des résidus appartient à la même classe de loi que la loi normale (Gourieroux, Montfort, 1989). L'estimateur sera (i) asymptotiquement convergent et (ii) asymptotiquement normal. Par conséquent la fonction de vraisemblance définissant l'estimateur du MV sous l'hypothèse de normalité et la fonction de pseudo-vraisemblance de l'estimateur du PMV sont les mêmes.

**Définition 2.4.1.** La fonction de log-vraisemblance associée à un échantillon de T observations  $(y_1, y_2, \dots, y_T)$  de  $Y_t$  sous l'hypothèse de normalité de la loi conditionnelle de  $Y_t$  sachant  $Y_{t-1}$  et  $X_t$  s'écrit :

$$\log L(\theta) = -\frac{T}{2} \log(2\pi) - \frac{1}{2} \sum_{t=1}^T \log[\sigma_t(\theta)] - \frac{1}{2} \sum_{t=1}^T \frac{[y_t - m_t(\theta)]^2}{\sigma_t(\theta)} \quad (2.30)$$

Dans ces notations la fonction  $\sigma_t(\theta)$  désigne la variance conditionnelle et donc il n'y a pas lieu de l'élever au carré.

**Exemple** Appliquons cette formule au cas d'un modèle de régression linéaire avec erreur ARCH(q) .

$$\begin{aligned} Y_t &= X_t \beta + \epsilon_t \\ \epsilon_t &= \sigma_t \eta_t(\theta) \text{ avec } \eta_t \text{ N.i.d}(0,1) \\ E(\epsilon_t | \epsilon_{t-1}) &= 0 \quad V(\epsilon_t | \epsilon_{t-1}) = \alpha_0 + \sum_{i=1}^q \alpha_i \epsilon_{t-i}^2 \end{aligned}$$

Dans ce cas on donc :

$$\begin{aligned} E(Y_t | Y_{t-1}, X_t) &= m_t(\theta) = X_t \beta \\ V(Y_t | Y_{t-1}, X_t) &= \sigma_t(\theta) = \alpha_0 + \sum_{i=1}^q \alpha_i (Y_{t-i} - \beta X_{t-i})^2 \end{aligned}$$

où  $\theta = (\beta, \alpha_0, \alpha_1, \dots, \alpha_q) \in \mathbb{R}^{q+2}$ . La log-vraisemblance s'écrit par conséquent :

$$\log L(\theta) = -\frac{T}{2} \log(2\pi) - \frac{1}{2} \sum_{t=1}^T \log[\alpha_0 + \sum_{i=1}^q \alpha_i (Y_{t-i} - \beta X_{t-i})^2] - \frac{1}{2} \sum_{t=1}^T (y_t - X_t \beta)^2 \times [\alpha_0 + \sum_{i=1}^q \alpha_i (Y_{t-i} - \beta X_{t-i})^2]^{-1}$$

On peut déduire assez facilement l'écriture des CPO qui définissent l'estimateur du MV ou du PMV selon les cas.

**Proposition 2.4.1 (17).** *Les estimateurs du MV sous l'hypothèse de normalité ou du PMV, notés  $\hat{\theta}$ , des paramètres  $\theta \in \mathbb{R}^k$ , satisfont un système non linéaire à  $k$  équations : ( $k = p + q + 2 =$  le nombre de paramètres ) :*

$$\frac{\partial \log L(\theta)}{\partial \theta} \Big|_{\theta=\hat{\theta}} = 0 \quad (2.31)$$

avec

$$\frac{\partial \log L(\theta)}{\partial \theta} \Big|_{\theta=\hat{\theta}} = -\frac{1}{2} \sum_{t=1}^T \frac{1}{\sigma_t(\hat{\theta})} \frac{\partial \sigma_t(\theta)}{\partial \theta} \Big|_{\theta=\hat{\theta}} + \frac{1}{2} \sum_{t=1}^T \frac{[y_t - m_t(\hat{\theta})]^2}{\sigma_t(\hat{\theta})^2} \frac{\partial \sigma_t(\theta)}{\partial \theta} \Big|_{\theta=\hat{\theta}} + \sum_{t=1}^T \left[ \frac{y_t - m_t(\hat{\theta})}{\sigma_t(\hat{\theta})} \right] \frac{\partial m_t(\theta)}{\partial \theta} \Big|_{\theta=\hat{\theta}}$$

**Remarque 2.4.1.** Ce système peut se décomposer en deux sous systèmes lorsque les paramètres  $\theta$  interviennent de façon séparée dans l'écriture de l'espérance et de la variance conditionnelle. Ainsi, si l'on a  $\theta = (\alpha\beta)'$  où  $\alpha$  n'apparaît que dans l'espérance conditionnelle et  $\beta$  dans la variance conditionnelle, on peut décomposer ce système en deux sous systèmes puisque :

$$\begin{aligned}\frac{\partial \log L(\alpha)}{\partial \alpha} \Big|_{\theta=\hat{\theta}} &= \sum_{t=1}^T \left[ \frac{y_t - m_t(\hat{\alpha})}{\sigma_t(\hat{\beta})} \right] \frac{\partial m_t(\alpha)}{\partial \alpha} \Big|_{\alpha=\hat{\alpha}} \\ \frac{\partial \log L(\beta)}{\partial \beta} \Big|_{\theta=\hat{\theta}} &= -\frac{1}{2} \sum_{t=1}^T \frac{1}{\sigma_t(\hat{\beta})} \frac{\partial \sigma_t(\beta)}{\partial \beta} \Big|_{\theta=\hat{\theta}} + \frac{1}{2} \sum_{t=1}^T \frac{[y_t - m_t(\hat{\alpha})]^2}{\sigma_t(\hat{\beta})^2} \frac{\partial \sigma_t(\beta)}{\partial \beta} \Big|_{\beta=\hat{\beta}}\end{aligned}$$

Dans le cas général du PMV, on sait que l'estimateur  $\hat{\theta}$  est asymptotiquement normal et que sa matrice de variance-covariance est définie par la formule suivante:

**Théorème 2.4.1.** *Sous certaines conditions de régularité, l'estimateur du PMV est asymptotiquement convergent et normal.*

$$\sqrt{T}(\hat{\theta} - \theta) \xrightarrow[T \rightarrow \infty]{d} N(0, J^{-1} I J^{-1})$$

où la matrice de variance-covariance asymptotique de l'estimateur du PMV est calculée à partir de :

$$J = E_0 \left[ -\frac{\partial^2 \log L(\theta)}{\partial \theta \partial \theta'} \right] \quad I = E_0 \left[ \frac{\partial \log L(\theta)}{\partial \theta} \frac{\partial \log L(\theta)}{\partial \theta'} \right]$$

où  $E_0$  désigne l'espérance prise par rapport à la vraie loi.

Naturellement dans la pratique les matrices  $I$  et  $J$  sont directement estimées en remplaçant l'espérance  $E_0$  par la moyenne empirique et le paramètre inconnu  $\theta$  par son estimateur convergent  $\hat{\theta}$ . Ainsi, on utilise :

$$\begin{aligned}\hat{I} &= \frac{1}{T} \sum_{t=1}^T \frac{\partial \log L(\theta)}{\partial \theta} \Big|_{\theta=\hat{\theta}} \frac{\partial \log L(\theta)}{\partial \theta'} \Big|_{\theta=\hat{\theta}} \\ \hat{J} &= -\frac{1}{T} \sum_{t=1}^T \frac{\partial^2 \log L(\theta)}{\partial \theta \partial \theta'} \Big|_{\theta=\hat{\theta}}\end{aligned}$$

et la variance estimée de  $\hat{\theta}$  vérifie alors

$$Var[\sqrt{T}(\hat{\theta} - \theta)] = \hat{J}^{-1} \hat{I} \hat{J}^{-1}$$

### Remarque 2.4.2.

1. Dans le cas où la vraie loi sous jacente est normale (Maximum de Vraisemblance), la matrice de variance-covariance asymptotique se réduit à :

$$Var[\sqrt{T}(\hat{\theta} - \theta)] = J^{-1}$$

puisque  $J = I$ .

2. Dans le cas du MV lorsque l'on peut séparer les paramètres de l'espérance conditionnelle et de la variance conditionnelle, on montre que :

$$Var[\sqrt{T}(\hat{\beta} - \beta)] = \left[ \frac{1}{T} \sum_{t=1}^T \frac{1}{2\sigma_t(\hat{\beta})^2} \frac{\partial \sigma_t(\beta)}{\partial \beta} \Big|_{\beta=\hat{\beta}} \frac{\partial \sigma_t(\beta)}{\partial \beta'} \Big|_{\beta=\hat{\beta}} \right]^{-1}$$

### 2.4.1 Méthode du maximum de vraisemblance

Étant donné un échantillon  $\{x_1, x_2, \dots, x_n\}$  de  $n$  observation *i.i.d* qui provient d'une distribution  $f(x)$  avec des paramètres inconnus  $\theta$ , alors la fonction de densité conjointe est

$$\log L(\theta \setminus x_1, x_2, \dots, x_n) = f(x_1, x_2, \dots, x_n \setminus \theta) = f(x_1 \setminus \theta) \times f(x_2 \setminus \theta) \times \dots \times f(x_n \setminus \theta)$$

Considérons les valeurs observées  $\{x_1, x_2, \dots, x_n\}$  fixées et  $\theta$  la fonction variable qui veut varier librement

Dans la pratique, il est souvent plus commode de travailler avec le logarithme de la fonction de vraisemblance et on l'appelle log-vraisemblance:

$$\log L(\theta \setminus x_1, x_2, \dots, x_n) = \sum_{i=1}^n \log f(x_i \setminus \theta)$$

Supposons que les observations  $\{x_1, x_2, \dots, x_n\}$  suivent une distribution normale avec des paramètres inconnus  $\theta = \{\mu, \sigma^2\}$ , alors

$$\begin{aligned} \log L(\theta \setminus x_1, x_2, \dots, x_n) &= \sum_{i=1}^n \log \left\{ \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x_i - \mu}{2\sigma^2}} \right\} \\ &= \sum_{i=1}^n \left( -\frac{1}{2} \log 2\pi - \log \sigma - \frac{(x_i - \mu)^2}{2\sigma^2} \right) \\ &= -\frac{n}{2} \log 2\pi - n \log \sigma - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2 \end{aligned}$$

Pour estimer les paramètres du modèle GARCH(p,q) avec k donné, on a

$$\begin{aligned} y_t &= C + \sum_{i=1}^k a_i y_{t-i} + \epsilon_t \\ \epsilon_t &= \sigma_t \eta_t \\ \sigma_t^2 &= \alpha_0 + \sum_{i=1}^q \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2, \quad \forall t \in \mathbb{Z} \end{aligned}$$

où  $\eta_t$  est un bruit blanc alors  $\epsilon_t$  est une distribution normale avec une moyenne 0 et de variance conditionnelle  $\sigma_t$

$$p(\epsilon_t | \epsilon_{t-1}, \dots, \epsilon_0) = \frac{1}{\sqrt{2\pi\sigma_t}} e^{-\frac{\epsilon_t^2}{2\sigma_t}}$$

La fonction log-vraisemblance du vecteur de paramètre  $\theta = (\alpha_0, \alpha_1, \dots, \alpha_q, \beta_1, \dots, \beta_p)^T$  est

$$L(\theta) = \sum_{t=q+1}^n l_t(\theta) = \sum_{t=q+1}^n \left( -\frac{1}{2} \log 2\pi - \frac{1}{2} \log \sigma_t - \frac{\epsilon_t^2}{2\sigma_t} \right)$$

Par conséquent, on a

$$\begin{aligned} \frac{\partial l_t(\theta)}{\partial \theta} &= \left( \frac{\epsilon_t^2}{2\sigma_t^2} - \frac{1}{2\sigma_t} \right) \frac{\partial \sigma_t}{\partial \theta} \\ \frac{\partial^2 l_t(\theta)}{\partial \theta \partial \theta^T} &= \left( \frac{\epsilon_t^2}{2\sigma_t} - \frac{1}{2\sigma_t} \right) \frac{\partial^2 \sigma_t}{\partial \theta \partial \theta^T} + \left( \frac{1}{2\sigma_t^2} - \frac{\epsilon_t^2}{\sigma_t^3} \right) \frac{\partial \sigma_t}{\partial \theta} \frac{\partial \sigma_t}{\partial \theta^T} \\ \frac{\partial \sigma_t}{\partial \theta} &= (1, \epsilon_{t-1}^2, \dots, \epsilon_{t-q}^2, \sigma_{t-1}, \dots, \sigma_{t-p})^T + \sum_{i=1}^p \beta_i \frac{\partial \sigma_{t-i}}{\partial \theta} \end{aligned}$$

Ainsi, le gradient est

$$\nabla L(\theta) = \frac{1}{2} \sum_{t=q+1}^n \left( \frac{\epsilon_t^2}{\sigma_t^2} - \frac{1}{\sigma_t} \right) \frac{\partial \sigma_t}{\partial \theta}$$

et la matrice d'information de Fisher est

$$\begin{aligned} J &= \sum_{t=q+1}^n E \left[ \left( \frac{\epsilon_t^2}{2\sigma_t^2} - \frac{1}{2\sigma_t} \right) \frac{\partial^2 \sigma_t}{\partial \theta \partial \theta^T} + \left( \frac{1}{2\sigma_t^2} - \frac{\epsilon_t^2}{\sigma_t^3} \right) \frac{\partial \sigma_t}{\partial \theta} \frac{\partial \sigma_t}{\partial \theta^T} \right] \\ &= -\frac{1}{2} \sum_{t=q+1}^n E \left( \frac{1}{\sigma_t^2} \frac{\partial \sigma_t}{\partial \theta} \frac{\partial \sigma_t}{\partial \theta^T} \right) \end{aligned}$$

**Exemples 2.4.1.** On considère un GARCH(1,1) alors

$$\begin{aligned}\epsilon_t &= \eta_t \sqrt{\sigma_t} \\ \sigma_t^2 &= \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2\end{aligned}$$

Pour estimer les coefficients inconnus  $\theta = (\alpha_0, \alpha_1, \beta_1)^T$ , il est

$$\nabla L(\theta) = \frac{1}{2} \sum_{t=2}^n \left( \frac{\epsilon_t^2}{\sigma_t^2} - \frac{1}{\sigma_t} \right) \frac{\partial \sigma_t}{\partial \theta}, \quad J = -\frac{1}{2} \sum_{t=2}^n E \left( \frac{1}{\sigma_t^2} \frac{\partial \sigma_t}{\partial \theta} \frac{\partial \sigma_t}{\partial \theta^T} \right)$$

où

$$\frac{\partial \sigma_t}{\partial \theta} = (1, \epsilon_{t-1}^2, \sigma_{t-1})^T + \beta_1 \frac{\partial \sigma_{t-1}}{\partial \theta}$$

### Algorithme

- Compte tenu des observations  $\{y_t\}_{t=1}^n$ , on obtient  $C, \hat{a}_1, \dots, \hat{a}_n$  du modèle prévisionnel AR(k) et  $y_t = \hat{C} + \sum_{i=1}^k \hat{a}_i y_{t-1} + \hat{\epsilon}_t$
- Puisque  $\eta_t$  est un bruit blanc et  $\epsilon_t = \eta_t \sqrt{\sigma_t}$  est une distribution normale avec une moyenne nulle et de variance  $\sigma_t$ , ainsi, la fonction de vraisemblance de  $\epsilon_1, \epsilon_2, \dots, \epsilon_n$  est:

$$L = \prod_{t=q+1}^n \left( \frac{1}{\sqrt{2\pi\sigma_t}} \right)$$

La fonction de Log-vraisemblance est

$$\log L(\theta) = \sum_{t=q+1}^n \left( -\frac{1}{2} \ln 2\pi - \frac{1}{2} \ln \sigma_t - \frac{\epsilon_t^2}{2\sigma_t} \right)$$

- Soit une valeur initiale  $\theta_0 = (\alpha_0, \alpha_1, \dots, \alpha_q, \beta_1, \dots, \beta_q)^T$  donnée et une estimation de  $\frac{\partial \sigma_1}{\partial \theta}, \dots, \frac{\partial \sigma_q}{\partial \theta}$ , nous avons

$$\theta_{k+1} = \theta_k - J^{-1}(\theta_k) \nabla L(\theta_k)$$

où

$$\nabla L(\theta_k) = \frac{1}{2} \sum_{t=q+1}^n \left( \frac{\epsilon_t^2}{\sigma_t^2} - \frac{1}{\sigma_t} \right) \frac{\partial \sigma_t}{\partial \theta}$$

$$\frac{\partial \sigma_t}{\partial \theta} = (1, \epsilon_{t-1}^2, \dots, \epsilon_{t-q}^2, \sigma_{t-1}, \dots, \sigma_{t-p})^T + \sum_{i=1}^p \beta_i \frac{\partial \sigma_{t-i}}{\partial \theta}$$

et

$$J = -\frac{1}{2} \sum_{t=q+1}^n E \left( \frac{1}{\sigma_t^2} \frac{\partial \sigma_t}{\partial \theta} \frac{\partial \sigma_t}{\partial \theta^T} \right)$$

$\{\epsilon_t\}$  la dérivées de la première étape .

– Répétez le processus d'itération jusqu'à ce que nous obtenons la convergence vers  $\theta$ .

## 2.4.2 Quasi-vraisemblance conditionnelle

Nous étudions en détail la méthode du quasi-maximum de vraisemblance conditionnelle (à des valeurs initiales), nous allons présenter une procédure itérative de calcul de la log-vraisemblance gaussienne, conditionnellement à des valeurs initiales fixes ou aléatoires. Cette vraisemblance est écrite comme si la loi des variables  $\eta_t$  était normale centrée réduite (on parle de pseudo ou quasi-vraisemblance).

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

$$\begin{aligned}\epsilon_t &= \sigma_t \eta_t \\ \sigma_t^2 &= \alpha_0 + \sum_{i=1}^q \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2, \quad \forall t \in \mathbb{Z}\end{aligned}$$

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

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

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

appartient à un espace de paramètres  $E \subset ]0, +\infty[ \times ]0, \infty[^{p+q}$ . La vraie valeur du paramètre est inconnue et est notée

$$\theta_0 = (\alpha_{00}, \alpha_{01}, \dots, \alpha_{0q}, \beta_{01}, \dots, \beta_{0p})'.$$

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

La spécification d'une distribution gaussienne pour les variables  $\eta_t$  ne permet pas d'en déduire simplement la loi de l'échantillon. On travaille avec la vraisemblance de  $\epsilon_1, \dots, \epsilon_n$  conditionnellement à certaines valeurs initiales.

Etant données des valeurs initiales  $\epsilon_0, \dots, \epsilon_{1-q}, \tilde{\sigma}_0^2, \dots, \tilde{\sigma}_{1-p}^2$  que nous allons préciser, la vraisemblance conditionnelle gaussienne  $L_n(\theta)$  s'écrit

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

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

$$\tilde{\sigma}_t^2 = \tilde{\sigma}_t^2(\theta) = \alpha_0 + \sum_{i=1}^q \alpha_i \epsilon_{t-1}^2 + \sum_{j=1}^p \beta_j \tilde{\sigma}_{t-j}^2 \quad (2.32)$$

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

$$\epsilon_0^2 = \dots = \epsilon_{1-q}^2 = \sigma_0^2 = \dots = \sigma_{1-p}^2 = \frac{\alpha_0}{1 - \sum_{i=1}^q \alpha_i - \sum_{j=1}^p \beta_j}$$

On peut prendre alors comme valeurs initiales

$$\epsilon_0^2 = \dots = \epsilon_{1-q}^2 = \tilde{\sigma}_0^2 = \dots = \tilde{\sigma}_{1-p}^2 = \alpha_0$$

ou encore

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

Un estimateur du QMV de  $\theta$  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  $\theta$

$$\tilde{I}_n(\theta) = n^{-1} \sum_{t=1}^n \tilde{\ell}_t, \quad \text{où} \quad \tilde{\ell}_t = \tilde{\ell}_t(\theta) = \frac{\epsilon_t^2}{\tilde{\sigma}_t^2} + \log \tilde{\sigma}_t^2 \quad (2.33)$$

et  $\sigma_t^2$  est définie en (3,31). Un estimateur du quasi-maximum de vraisemblance est donc une solution mesurable de l'équation

$$\hat{\theta}_n = \arg \min_{\theta \in \Theta} \tilde{I}_n(\theta) \quad (2.34)$$

Nous montrerons que le choix des valeurs initiales n'a pas d'importance pour les propriétés asymptotiques de l'estimateur du QMV. En pratique ce choix peut cependant avoir de l'importance. Notons que d'autres méthodes sont possibles  $\tilde{\sigma}_t^2$ , par exemple en prenant  $\tilde{\sigma}_t^2 = c_0(\theta) + \sum_{i=1}^{t-1} c_i(\theta) \epsilon_t^2$  où les  $c_i(\theta)$  sont calculés récursivement (voir Berkes et al., 2003).

Remarquons que pour calculer  $\tilde{I}_n(\theta)$ , cette procédure nécessite un nombre d'opérations de l'ordre de  $n^2$ , tandis que celle que nous proposons ne requiert qu'un ordre de  $n$  opérations.

## Équations de vraisemblance

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

$$\frac{1}{n} \sum_{t=1}^n \{\epsilon_t^2 - \tilde{\sigma}_t^2\} \frac{1}{\tilde{\sigma}_t^4} \frac{\partial \tilde{\sigma}_t^2}{\partial \theta} = 0 \quad (2.35)$$

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

$$\frac{1}{n} \sum_{t=1}^n \{\epsilon_t^2 - \sigma_t^2\} \frac{1}{\sigma_t^4} \frac{\partial \sigma_t^2}{\partial \theta} = 0 \quad (2.36)$$

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

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

car  $\frac{1}{\sigma_t^4(\theta_0)} \frac{\partial \sigma_t^2(\theta_0)}{\partial \theta}$  est une fonction mesurable des  $\epsilon_{t-i}$ ,  $i > 0$ . Ce résultat n'est autre que la version asymptotique de (2.34) en  $\theta_0$ , en utilisant le théorème ergodique.

# Chapitre 3

## Notion de Mesures de Risque

Nous allons aborder dans ce chapitre le concept de mesure de risque. Ensuite, nous allons donner une estimation de la VaR en utilisant la modélisation GARCH(p,q).

### 3.1 Mesure de risque

Soit  $\nu$  l'ensemble des variables aléatoires à valeurs réelles.

**Définition 3.1.1.** Mesure de risque

Une mesure de risque  $\rho$  est une fonction définie sur l'espace des variables aléatoires, et prenant ses valeurs dans  $\mathbb{R}$ .

$$\rho : \nu \rightarrow \overline{\mathbb{R}}$$

### 3.2 Caractéristiques d'une mesure de risque

La définition d'une mesure de risque est très générale puisque toute fonctionnelle réelle positive d'une variable aléatoire peut être considérée comme étant une mesure de risque. Aussi, en pratique, on exige de telles mesures qu'elles disposent de propriétés mathématiques dont la transcription conceptuelle permette de les juger. En pratique, on exige fréquemment qu'une mesure de risque  $\rho$  possède une partie des caractéristiques suivantes :

1. **Invariance en loi :**

$$\forall X, Y \in \mathcal{V}, \quad X \stackrel{\ell}{=} Y \Rightarrow \rho(X) = \rho(Y)$$

2. **Invariance par translation:**

$$\forall X \in \mathcal{V}, \forall \alpha \in \mathbb{R}, \quad \rho(X + \alpha) = \rho(X) + \alpha$$

3. **Sous-additivité:**

$$\forall X, Y \in \mathcal{V}, \quad \rho(X + Y) \leq \rho(X) + \rho(Y)$$

4. **Homogénéité positive:**

$$\forall \lambda \geq 0, \forall X \in \mathcal{V}, \quad \rho(\lambda X) = \lambda \rho(X)$$

5. **Monotonie:**

$$\forall X, Y \in \mathcal{V} \quad X \leq Y \Rightarrow \rho(X) \leq \rho(Y)$$

6. **Convexité:**

$$\forall \beta \in [0,1], \forall X, Y \in \mathcal{V}, \quad \rho(\beta X + (1 - \beta)Y) \leq \beta \rho(X) + (1 - \beta) \rho(Y)$$

**Remarque 3.2.1.**

- La monotonie signifie que si une variable aléatoire génère moins de pertes qu'une autre, alors son critère de risque est inférieur.
- L'invariance par translation signifie que les éléments déterministes additifs ne sont pas sujets à la mesure de risque. La propriété implique :  $\rho(X - \rho(X)) = 0$  Ainsi,  $\rho(X)$  est ce qu'il faut retirer à X pour annuler son risque sous la mesure  $\rho(\Delta)$ .
- L'homogénéité positive signifie que les éléments déterministes multiplicatifs n'influencent pas la mesure de risque.  
L'homogénéité implique que  $\rho(0)=0$

### 3.3 Présentation de la Value at Risk

Cette notion est originaire du secteur de l'assurance. Elle a été importée à la fin des années 1980 sur les marchés financiers aux États-Unis par la banque Bankers Trust et popularisée par la banque JP Morgan en 1993 et son service (gratuit et public) Riskmetrics puis adoptée sous une forme embryonnaire par le Comité de Bâle (Bâle II) pour les banques et la Solvabilité pour les assurances.

La VaR d'un portefeuille dépend essentiellement de trois paramètres :

- la distribution des résultats des portefeuilles. Souvent cette distribution est supposée normale, mais beaucoup d'acteurs financiers utilisent des distributions historiques. La difficulté réside dans la taille de l'échantillon historique : s'il est trop petit, les probabilités de pertes élevées sont peu précises, et s'il est trop grand, la cohérence temporelle des résultats est perdue (on compare des résultats non comparables) ;
- le niveau de confiance choisi (95 ou 99 en général). C'est la probabilité que les pertes éventuelles du portefeuille ou de l'actif ne dépassent pas la Value at Risk, par définition ;
- l'horizon temporel choisi. Ce paramètre est très important car plus l'horizon est lointain plus les pertes peuvent être importantes. Par exemple, pour une distribution normale des rendements, il faut multiplier la Value at Risk à un jour par pour avoir la Value at Risk sur jours.

D'une manière générale, la VaR donne une estimation des pertes qui ne devrait pas être dépassée sauf événement extrême sur un portefeuille pouvant être composé de différentes classes d'actifs.

**Définition 3.3.1.** (L'inverse généralisé )

Soit  $X$  une variable aléatoire de fonction de répartition  $F$  connue; on définit l'inverse généralisé de  $F$ , noté  $F^{-1}$  par :

$$\begin{aligned} F_X^{-1}(\alpha) &= \inf\{x \in \mathbb{R}/F(x) \geq \alpha\}; 0 < \alpha < 1 \\ &= \sup\{x \in \mathbb{R}/F(x) \leq \alpha\} \end{aligned}$$

La VaR acronyme désignant la value at risk, est l'un des derniers nés de mesures de risque et des plus largement utilisés de nos jours; c'est une mesure très simple à appréhender puisqu'elle définit le risque par une valeur numérique unique.

Selon Esch et al en 1997 et Jorion en 2000, la VaR est définie par

**Définition 3.3.2.** (La VaR)

La VaR au niveau  $\alpha$  est définie par le quantile de niveau  $\alpha$  ( $\alpha$  – *quantile*) :

$$P[X \leq VaR_\alpha] = \alpha \Leftrightarrow F_X(VaR_\alpha) = \alpha$$

D'où

$$VaR_\alpha(X) = F_X^{-1}(\alpha)$$

**Remarque 3.3.1.**

$$P[X \leq VaR_\alpha] = \alpha \Rightarrow P\left[\frac{X - E(X)}{\sigma(X)} \leq \frac{VaR_\alpha - E(X)}{\sigma(X)}\right] = \alpha$$

Posons :

$$Z_\alpha = \frac{VaR_\alpha - E(X)}{\sigma(X)}$$

On déduit que :

$$VaR_\alpha = E(X) + Z_\alpha \sigma(X).$$

### 3.4 Calcul de la Value at Risk Par le modèle GARCH

On dénombre trois grandes classes de méthode de calcul de la VaR

1. Méthode non-paramétrique (la VaR historique, simulation de monte Carlo, ...)
2. Méthode semi-paramétrique (théorie des extrêmes, ...)
3. Méthode paramétrique (ARCH, GARCH univarié)

Parmi les méthodes paramétriques l'ensemble des méthodes de calcul et de prévision de la VaR fondées sur des modèles GARCH univariés (Engle, 2001). Ces modèles permettent de modéliser et de prévoir la variance conditionnelle de la distribution de pertes et profits, ce qui permet dans un second temps d'en déduire une modélisation ou une prévision de la Value-at-Risk sous un certain nombre d'hypothèses sur la distribution conditionnelle des rendements.

Soit  $p_t$  la valeur d'un actif à la date  $t$

$$\begin{aligned}\Delta p_t &= p_t - p_{t-1} \\ R_t &= \log \frac{p_t}{p_{t-1}} = \log p_t - \log p_{t-1}\end{aligned}$$

Soit  $f_{R_t}(r/\mathfrak{F}_t)$  la densité conditionnelle à un ensemble d'information disponible à la date  $t$ , notée  $\mathfrak{F}_t$ .

La VaR à la date  $t$ , conditionnellement à l'ensemble d'information  $\mathfrak{F}_t$  s'écrit sous la forme:

$$P(p_t \leq VaR_t(\alpha)/\mathfrak{F}_t) = \alpha$$

*i.e*

$$VaR_t(\alpha) = F_{R_t}^{-1}(\alpha/\mathfrak{F}_t)$$

On suppose que  $R_t$  satisfait le modèle:

$$\begin{aligned}R_t &= c + \epsilon_t \\ \epsilon_t &= \eta_t \sqrt{\sigma_t} \\ \sigma_t^2 &= \alpha_0 + \sum_{t=1}^q \alpha_t \epsilon_{t-1}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2\end{aligned}$$

$(\eta_t)_t$  *i.i.d*

$$\sigma_t^2 = V(R_t/\mathfrak{F}_{t-1}) = V(\epsilon_t/\mathfrak{F}_{t-1})$$

Notons  $\hat{\alpha}_i, \hat{\beta}_i, \hat{c}$ , les estimateurs convergents donnés par la méthode du maximum de vraisemblance.

Soit  $\sigma_1$  une condition initiale sur le processus de variance conditionnelle  $\sigma_t$  (généralement fixée de façon arbitraire).

Par définition, la VaR de niveau  $\alpha$  est:

$$\begin{aligned} P(R_t \leq VaR_t(\alpha)/\mathfrak{F}_t) &= \alpha \\ \Rightarrow P(\eta_t \leq \frac{VaR_t(\alpha) - c}{\sqrt{\sigma_t}}/\mathfrak{F}_t) &= \alpha \end{aligned}$$

En remplaçant  $c$  par sa valeur estimée  $\hat{c}$  et  $\sigma_t^2$  par sa valeur anticipée  $\hat{\sigma}_t^2$ , on obtient :

$$\begin{aligned} P(\eta_t \leq \frac{VaR_t(\alpha) - \hat{c}}{\sqrt{\hat{\sigma}_t}}/\mathfrak{F}_t) &= \alpha \\ \Rightarrow VaR_t(\alpha) &= \sqrt{\hat{\sigma}_t} F_t^{-1}(\alpha) + \hat{c} \end{aligned}$$

$F_t$  est la fonction de répartition de  $\eta_t$

$$\hat{\sigma}_t = \hat{\alpha}_0 + \sum_{i=1}^q \hat{\alpha}_i \epsilon_{t-1}^2 + \sum_{j=1}^p \hat{\beta}_j \sigma_{t-j}^2$$

### 3.5 Prévision de la VaR

Comment obtenir une prévision de la VaR à partir d'un modèle GARCH? la démarche est indirecte: dans un premier temps, on fait, on fait une hypothèse sur la distribution de l'échantillon, puis l'on estime les paramètres du modèle GARCH sur les observations de la période de 1 à T, généralement par une procédure de type maximum de vraisemblance.

Dans une second étape, on déduit du modèle GARCH estimé, une prévision de la variance conditionnelle, qui couplée à l'hypothèse retenue sur la distribution de l'échantillon, permet de construire une prévision de la distribution de l'échantillon valable pour (T+1), ce qui permet dans un second temps d'en déduire une modélisation ou une prévision de la VaR.

#### Prévision à la date (T+1)

Considérons l'exemple d'un modèle GARCH(1,1); sous l'hypothèse de la distribution de la loi normale

$$Y_t = m + \epsilon_t, \quad m = E[Y_t/Y_t - 1]$$

$$\begin{aligned}\epsilon_t &= \sigma_t \eta_t \\ \sigma_t^2 &= \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2\end{aligned}$$

avec  $\eta_t \rightarrow N(0,1)$

Dans ce cas  $\theta = (\alpha_0, \alpha_1, \beta_1, m)$  et notons  $\hat{\theta} = (\hat{\alpha}_0, \hat{\alpha}_1, \hat{\beta}_1, \hat{m})$  l'estimateur du maximum de vraisemblance.

Une prévision de la variance conditionnelle pour la date (T+1) est donnée par,

$$\hat{\sigma}_{T+1}^2 = \hat{\alpha}_0 + \hat{\alpha}_1 \hat{\epsilon}_T^2 + \hat{\beta}_1 \hat{\sigma}_T^2 \text{ avec } \sigma_1^2 \text{ donné}$$

soit encore

$$\hat{\sigma}_{T+1}^2 = \hat{\alpha}_0 + \hat{\alpha}_1 (Y_t - \hat{m})^2 + \hat{\beta}_1 \hat{\sigma}_T^2$$

On note  $VaR_{T+1/T}(\alpha)$  la prévision de la Value-at-Risque de niveau  $\alpha$  pour la date (T+1) conditionnellement à l'information disponible à la date T.

Par définition de la VaR on a :

$$P[Y_{T+1} \leq VaR_{T+1/T}(\alpha) / \mathfrak{F}_T] = \alpha$$

Où  $\mathfrak{F}_T$  est l'ensemble d'information disponible à la date T.

On déduit immédiatement que

$$P[\eta_{T+1} \leq \frac{VaR_{T+1/T}(\alpha) - m}{\sigma_{T+1}} / \mathfrak{F}_T] = \alpha$$

En remplaçant le paramètre m par sa valeur estimée et la variance conditionnelle  $\sigma_{T+1}$  par sa valeur anticipée  $\hat{\sigma}_{T+1}$ , il vient :

$$P[\eta_{T+1} \leq \frac{VaR_{T+1/T}(\alpha) - \hat{m}}{\hat{\sigma}_{T+1}} / \mathfrak{F}_T] = \alpha$$

Si l'on suppose que la distribution de  $\eta_t$  est une loi normale  $N(0,1)$ , on peut en déduire immédiatement une prévision de la distribution valable pour la date (T+1), c'est-à-dire la VaR.

$$VaR_{T+1/T}(\alpha) = \hat{\sigma}_{T+1} \phi^{-1}(\alpha) + \hat{m} \quad (3.1)$$

Où  $\phi^{-1}$  est la fonction inverse de la fonction de répartition de la loi normale centrée réduite.

### Prévision à la date (T+k)

Considérons le modèle GARCH(p,q) présenté dans la définition (2,1,1) sous l'hypothèse de la loi normale  $\eta_t \rightarrow N(0,1)$

Soit  $\hat{\theta} = (\hat{\alpha}_0, \hat{\alpha}_i (i = 1, \dots, p), \hat{\beta}_i (i = 1, \dots, q), \hat{m})$  l'estimateur du vecteur des paramètres

$\theta = (\alpha_0, \alpha_i (i = 1, \dots, p), \beta_i (i = 1, \dots, q), m)$  obtenu par la méthode de maximum de vraisemblance.

Une prévision de la variance conditionnelle pour la date  $(T+k)$  est donnée par,

$$\widehat{\sigma}_{T+k}^2 = \widehat{\alpha}_0 + \sum_{i=1}^p \widehat{\alpha}_i \widehat{\epsilon}_{T+k-i}^2 + \sum_{i=1}^q \widehat{\beta}_i \widehat{\sigma}_{T+k-i}^2, \text{ avec } \sigma_1^2 \text{ donné} \quad (3.2)$$

On note  $VaR_{T+k/T}(\alpha)$  la prévision de la Value-at-Risque de niveau  $\alpha$  pour la date  $(T+k)$  conditionnellement à l'information disponible à la date  $T$ .

Par définition de la VaR on a :

$$P[Y_{T+k} \leq VaR_{T+k/T}(\alpha) / \mathfrak{F}_T] = \alpha$$

Où  $\mathfrak{F}_T$  est l'ensemble d'information disponible à la date  $T$ .

Par suite

$$P[\eta_{T+k} \leq \frac{VaR_{T+k/T}(\alpha) - \widehat{m}}{\widehat{\sigma}_{T+k}} / \mathfrak{F}_T] = \alpha$$

Si l'on suppose que la distribution de  $\eta_t$  est une loi normale  $N(0,1)$ , on peut en déduire immédiatement une prévision de la distribution valable pour la date  $(T+k)$ , c'est-à-dire la VaR.

$$VaR_{T+k/T}(\alpha) = \widehat{\sigma}_{T+k} \phi^{-1}(\alpha) + \widehat{m} \quad (3.3)$$

Où  $\phi^{-1}$  est la fonction inverse de la fonction de répartition de la loi normale centrée réduite.

# Conclusion

Dans ce mémoire, nous avons étudié la modélisation par la volatilité stochastique d'une série financière par la modélisation GARCH. Nous avons donné les estimations des paramètres de ce modèle par deux méthodes à savoir méthode du maximum de vraisemblance et quasi-maximum de vraisemblance. Les propriétés probabilistes de ces estimateurs ont été données.

Nous avons donné aussi une application en finance qui est le calcul d'une mesure de risque à savoir la VaR par le modèle GARCH.

Il serait intéressant d'étendre cette étude à d'autres types de modèles GARCH, comme par exemple : le modèle EGARCH, TGARCH, etc. . . , et aussi le modèle GARCH périodique.

De plus une comparaison du calcul de la VaR avec les autres méthodes existantes est envisagée.

# Bibliographie

- [1 ] Agnès Lagnoux (2010-2011) *Renforcement Statistique Séries chronologiques: Université de toulouse Le Mirail* ISMAG MASTER 1 -MIS243Y  
(<http://www.math.univ-toulouse.fr/>)
- [2 ] Alfonsi (2006), *Mesure de risque, Cours De Master Recherche : Probabilité Et Application*. ([http://cermic.enpc.fr / alfonsi /mrf-risk.pdf](http://cermic.enpc.fr/alfonsi/mrf-risk.pdf) )
- [3 ] Arthur Charpentier, A.(2010)*Mesure de risque:Université Rennes1 France*.
- [4 ] Bollerslev, T.(1986) *Generalized Athoregressive Conditional Heteroskedasticity*. J.Econometrics. Vol. 31, No. 3, 307-327.
- [5 ] Bougerol, P., and Picard, N. (1992) *Strict stationarity of generalized autoregressive processes*. Ann. Probab. Vol. 20, No. 4, 1714-1730
- [6 ] Brantdt, A. (1986) *The stochastic equation  $Y_{n+1} = A_n Y_n + B_n$  with stationary coefficients*. Adv. in App. 18, 211-220.
- [7 ] Christian Francq *Modèles GARCH et à volatilité stochastique* CREST-ENSAE Université Lille 3.
- [8 ] Engle, R.F. (1982) *Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation*. Econometrica, Vol. 50, No. 4,987-1008.
- [9 ] Engle, R. F.(2001) *Value at Risk Models in Finance*. European Central Bank, Working Paper No. 75, ISSN 1651 1810.
- [10 ] Esch, K. et al en (1997) *Value at Risk: vers un Risque Management moderne*. DeBoeck, Bruxelles.
- [11 ] Saporta, B. (2004) *Étude de la solution stationnaire de l'équation  $Y_{n+1} = A_n Y_n + B_n$  à coefficient aléatoire*. PhD thesis, Université de Renne 1.
- [12 ] Francq, C., and Zakoïan, J, M. (2008)*A tour in the asymptotic theory of GARCH estimation*. In T.G. Andersen, R.A. Davis, J.P. Kreiss and T. Mikosch (eds), *Handbook of Financial Time Series*. Springer-Verlag, Berlin.
- [13 ] Francq, C., and Zakoïan, J, M. (2004)*Maximum likelihood estimation of pure GARCH and ARMA-GARCH processes*. Bernoulli 10(4), 605-637.
- [14 ] Francq, C., and Zakoïan, J, M. (2010)*GARCH Models, Structure, Statistical inference and financial application*. John Wiley and Sons, Ltd, Publication.
- [15 ] Francq, C., and Zakoïan, J, M.(2007)*Quasi-maximum likelihood estimation in GARCH processes when some coefficients are equal to zero*. Stochastic processes and their Applications 117, 1265-1284.

- [16] Furstenberg, H., and Kesten, H. (1960) *Products of random matrices*. Ann. Math. Statist. Vol. 31, No. 2, 457-469.
- [17] Gouriéroux, C. (1992) *Modèles ARCH et Applications Financières*. Collection ENSAE, Economica.
- [18] Hurlin, C., 2006-2007, *Econométrie pour la finance: Modèles ARCH-GARCH, Applications à la VaR*. Master économétrie et Statistique appliquée, Université d'Orléans.
- [19] Jorion, P., (2000), *The Value at Risk Fieldbook: The Complete Guide to Implementing Var*. McGraw-Hill Companies.
- [20] Kingman, J. F. C. (1973) *Subadditive ergodic theory*. Ann. Probab. Vol. 1, No. 6, 883-909
- [21] Koralov, L. B., and Sinai, Y.G. (2000) *Theory of probability and random processes, second edition*. Springer-Verlag, New York.
- [22] Michel Prenat, Christine Keribin, Raphaël Rossignol (2010-2011) *Séries chronologiques Volume 1: cours et exercices Université Paris-Sud*.
- [23] Nelson, D. B. (1990) *Stationarity and Persistence in the GARCH(1,1) Model*. Econometric Theory. Vol. 6, 318-334.
- [24] Vervaat, W. (1979) *On a stochastic difference equation and a representation of non-negative infinitely divisible random variables*. Adv. in Appl. Probab. 11, 750-783
- [25] Yiyang Yang, (2012) *Parameter Estimation of GARCH Model: State University of New York at Stony Brook*.