

MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE SCIENTIFIQUE
UNIVERSITE MOULOU D MAMMERI, TIZI-OUZOU.



FACULTE: DES SCIENCES
DEPARTEMENT : MATHEMATIQUES

Memoire de Master

FILIERE : MATHEMATIQUES
SPECIALITE : Modélisation Mathématique

Présentée par:
LALLALI Karima

Thème:

Estimation des modèles ARCH en présence de données manquantes

Devant le jury d'examen composé de :

Khellas	Fazia	Professeur	UMMTO	Présidente
Hamaz	Abdelghani	MCB	UMMTO	Rapporteur
Zougab	Nabil	MCA	UMMTO	Examinateur
Mellah	Omar	MCB	UMMTO	Examinateur

2015/2016

Table des matières

Introduction générale	2
1 Estimation et interpolation des données manquantes	5
1.1 Introduction	5
1.2 Stationnarité de processus $\{X_t, t \in \mathbb{Z}\}$ et théorème de Wold	6
1.3 Prédiction linéaire de processus stationnaire à un pas	8
1.3.1 Equation de prédiction	8
1.3.2 Calcul des coefficients des prédicteurs	10
1.3.3 Algorithme de Durbin-Levinson	10
1.4 Interpolation de données manquantes	12
1.4.1 Préliminaires	12
1.4.2 Interpolation d'une seule valeur manquante	14
1.4.3 Interpolation de plusieurs valeurs manquantes	17
1.4.4 Algorithme EM	18
2 Estimations des modèles ARCH en présence de données manquantes	21
2.1 Introduction	21
2.2 Modèles ARCH-GARCH	22
2.2.1 Stationnarité forte d'un GARCH	23
2.2.2 Stationnarité faible d'un GARCH	25
2.3 Estimation des modèles ARCH par la méthode des moindres carrés	27
2.4 Estimation des modèles ARCH(q) par la méthode des (M.C.O) en présence des données manquantes	31
2.4.1 Construction de l'estimateur préliminaire $\hat{\alpha}_{pr}$	31
2.4.2 Comportement asymptotique de l'estimateur préliminaire $\hat{\alpha}_{pr}$	32
2.4.3 Construction de l'estimateur d'intérêt $\hat{\alpha}$	34

Table des matières	2
3 Illustration numérique	39
Conclusion générale	55
Bibliographie	55

Introduction générale

Le problème de données manquantes d'une série chronologique est un problème concret et toujours embarrassant lorsqu'il s'agit de données réelles. En effet, on a souvent affaire à des séries que l'on n'observe que pendant quelques heures de la journée, voire quelques jours de la semaine, à cause des rythmes de la vie sociale ou économique. En général, les données manquantes peuvent être attribuées à des origines diverses, soit par nature comme

1. Donnée manquante totale: C'est à dire que toute l'observation manque.
2. Donnée manquante partielle: C'est à dire que l'observation est présente mais il manque certaines données de cette observation,

soit parce que des valeurs aberrantes ont été détectées et donc éliminées.

On peut citer des exemples des séries avec données manquantes plus systématique par exemple les données manquantes peut être considérées que les vacances ou les week-ends, des jours où il n'y a pas de mesures de la variable à l'étude, bien que l'activité économique continue en tant que produit des phénomènes politiques et sociaux qui ont une influence directe sur la prochaine valeur de l'indice étudié. Par conséquent, de grandes différences peuvent être observées dans les valeurs de certains indices économiques après un week-end ou un jour férié. En revanche, sur les marchés financiers, les données transaction par transaction sont des exemples de données de haute fréquence qui se produisent à des heures irrégulières.

Le problème de la gestion des données manquantes est un vaste sujet. En effet, les observations ayant des valeurs manquantes représentent un défi important car les procédures de modélisation classique éliminent tout simplement ces observation de l'analyses, pour éviter de les supprimer, des procédures d'imputations existent dans la littérature des séries chronologiques.

L'objectif principale de ce mémoire est de généraliser la procédure d'estimation des paramètres du modèle ARCH de Bose et Mukherjee (2003), au cas de données manquantes

et d'étudier la consistance et la normalité asymptotique de cet estimateur d'intérêt.

Ce travail est organisé comme suit:

Le premier chapitre sera consacré à l'étude de l'estimation et l'interpolation de données manquantes, après avoir rappeler la procédure de prédiction de processus stationnaire. Le second chapitre fera l'objet de l'étude du comportement asymptotique de l'estimateur de P. Bondon (2012) obtenue par la méthode moindre carré ordinaire (M.C.O) en deux étapes. L'accent sera mis sur l'étude de la consistance et de la normalité asymptotique des estimateur préliminaire et d'intérêt noté, respectivement, $\hat{\alpha}_{pr}$ et $\hat{\alpha}$. Quant au troisième chapitre, nous allons effectuer des simulations intensives dont le but d'illustrer numériquement les résultats du chapitre 2 en utilisant le langage de programmation R.

Chapitre 1

Estimation et interpolation des données manquantes

Tout au long de ce mémoire nous noterons (Ω, \mathcal{F}, P) un espace probabilisé, sur lequel nous considérerons une suite de variables aléatoires réelles $\{X_t, t \in \mathbb{Z}\}$. Une telle suite est appelée série temporelle et constitue un exemple de processus stochastique à temps discret.

1.1 Introduction

Ce chapitre porte sur la prédiction de série chronologique et le traitement des données manquantes. Il s'agit de développer différentes approches et algorithmes permettant d'atteindre l'objectif en question. L'intérêt de ce problème, est justifié par la rencontre fréquente de séries avec données manquantes dans la réalité des problèmes pratiques.

Nous traitons le problème de la prédiction linéaire à un pas de $\{X_t\}$ qui consiste à trouver une fonction linéaire de $X_{t-1}, X_{t-2}, \dots, X_{t-n}$ qui approxime $\{X_t\}$ au sens de l'erreur quadratique moyenne, à partir des quelles on construit le prédicteur $\hat{X}_{t,n}$ sous la forme $\hat{X}_{t,n} = \sum_{k=1}^n \phi_{k,n} X_{t-k}$. Cette forme représente un modèle autorégressif (AR), tel que $\phi_{k,n}$ s'appellent les paramètres autorégressifs. L'accent sera mis sur le calcul de coefficient.

En suite la seconde partie du chapitre sera consacrée à l'étude des problèmes d'interpolation d'une seule valeur manquante. Une généralisation du résultat obtenu précédemment est présentée pour le cas de plusieurs valeurs manquantes. La troisième partie concerne l'étude de l'algorithme d'espérance maximisation (EM) qui est une méthode itérative efficace pour calculer l'estimation de vraisemblance maximale en présence de données manquantes.

1.2 Stationnarité de processus $\{X_t, t \in \mathbb{Z}\}$ et théorème de Wold

Rappelons qu'un processus aléatoire X indexé par un ensemble \mathbf{T} (a priori quelconque) est une famille $\{X_t, \text{ ou } t \in \mathbf{T}\}$, de vecteurs aléatoires à valeurs dans l'espace d'états $\mathbf{E} = \mathbb{R}^k$ ou $\mathbf{E} = \mathbb{C}^k$. Si $\mathbf{E} = \mathbb{R}$ ou $\mathbf{E} = \mathbb{C}$ (i.e $k = 1$), le processus est dit unidimensionnel (ou univarié).

Si $\mathbf{T} = \mathbb{Z}$ ou $\mathbf{T} = \mathbb{N}$, X est une série chronologique (ou temporelle) univariée, car les éléments de \mathbf{T} correspondent implicitement à des instants.

Etude de la stationnarité

La stationnarité joue un rôle majeur en séries temporelles car elle remplace de manière naturelle l'hypothèse d'observations *iid* (indépendante identiquement distribuée) en statistique standard.

Définition 1.1 (stationnarité forte). Le processus $\{X_t, t \in \mathbb{Z}\}$ est dit strictement stationnaire, ou fortement stationnaire, ou stationnaire au premier ordre si, quel que soit le n -uplet du temps $t_1 < t_2 < \dots < t_n$, tels que $t_i \in \mathbb{Z}$ et pour tout temps $h \in \mathbb{Z}$ avec $t_{i+h} \in \mathbb{Z}, \forall i = 1, \dots, n$, la suite $(X_{t_{i+h}}, \dots, X_{n+h})$ a la même loi que la suite $(X_{t_1}, \dots, X_{t_n})$.

Cependant cette notion est rarement vérifiée, alors que la notion suivante peut sembler moins exigeante car elle n'impose de contraintes qu'aux deux premiers moments des variables $\{X_t, t \in \mathbb{Z}\}$.

Définition 1.2 (stationnarité faible). . Un processus $\{X_t, t \in \mathbb{Z}\}$ est dit stationnaire au second ordre, ou stationnaire au sens faible, ou stationnaire d'ordre deux si les trois conditions sont satisfaites:

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

La première condition $\mathbb{E}(X_t^2) < \infty$ garantit tout simplement l'existence (ou la convergence) des moments d'ordre deux. La seconde condition $\forall t \in \mathbb{Z}, \mathbb{E}(X_t) = m$ porte sur les moments d'ordre un, et signifie tout simplement que les variables aléatoires $\{X_t, t \in \mathbb{Z}\}$ doivent avoir la même espérance quelle que soit la date t . Autrement dit, l'espérance du processus $\{X_t, t \in \mathbb{Z}\}$ doit être indépendante du temps. Enfin, la troisième condition d'autocovariance, implique que ces moments doivent être indépendants de la date considérée, et ne doivent

dépendre uniquement que de l'ordre des retards. Autrement dit la fonction d'autocovariance des processus $\{X_t, t \in \mathbb{Z}\}$ doit être indépendant du temps. En résumé,

un processus est stationnaire au second ordre si l'ensemble de ses moments sont indépendants du temps. Par conséquent, il convient de noter que la stationnarité implique que la variance $\gamma(0)$ du processus $\{X_t, t \in \mathbb{Z}\}$ est constante au cours du temps.

Remarque 1.1. La stationnarité forte implique les deux dernières conditions de la stationnarité faible. Autrement dit, la stationnarité forte implique la stationnarité faible si le moment d'ordre deux est fini.

Théorème de wold (Décomposition de Wold)

Wold (1939) montre que tout processus stationnaire au sens faible peut s'écrire comme une somme pondérée de chocs aléatoires orthogonaux, centrés et de même variances. Ce théorème joue un rôle central dans la modélisation des processus stochastiques linéaires.

Définition 1.3. Tout processus stochastique stationnaire du second ordre $\{X_t\}$, possède une décomposition unique donnée par

$$X_t = C_t + Y_t$$

Où $\{C_t\}$ et $\{Y_t\}$ sont deux processus orthogonaux. Le processus $\{C_t\}$ est pûrment déterminable (ou encore processus singulier), et $\{Y_t\}$ un processus indéterminable (régulier).

X_t peut être représenté par une combinaison linéaire du présent et de passé d'un processus bruit blanc, converge en moyenne quadratique de la forme $Y_t = \sum_{j=0}^{\infty} b_j \epsilon_{t-j}$ où $\{\epsilon_t\}$ est un processus bruit blanc du second ordre du processus. Les paramètres b_j satisfont $b_0 = 1$, $b_j \in \mathbb{R}$, $\forall j \in \mathbb{N}^*$, $\sum_{j=0}^{\infty} b_j^2 < \infty$. Cette dernière condition est très importante elle assure l'existence des moments d'ordre deux.

si $C_t \equiv 0$, alors $\{X_t\}$ est dit pûrment non déterminable.

Remarques 1.1. 1- On appelle innovation d'un processus stationnaire $\{X_t\}$, la variable

$$\epsilon_t = X_t - \hat{X}_t = X_t - \mathbb{E}(X_t / X_{t-1}, X_{t-2}, \dots)$$

\hat{X}_t est la prévision optimale de X_t fonction de son passé (meilleure approximation linéaire) et l'erreur de prévision correspondante à 1 pas est appelée innovation.

Ainsi, l'innovation est la partie de $\{X_t\}$ non corrélée au passé de la série.

2- L'implication forte de ce théorème est que, si nous connaissons les pondération b_j nous pouvons proposer une représentation de n'importe quel processus stationnaire. Cette représentation est qualifiée de moyenne mobile $MA(\infty)$.

1.3 Prédiction linéaire de processus stationnaire à un pas

Soit $\{X_t, t \in \mathbb{Z}\}$ un processus stationnaire au second ordre à valeurs réelles, d'espérance nulle et de fonction d'autocovariance $\gamma(h) = \text{cov}(X_t, X_{t+h})$.

Dans ce qui suit la notations $\langle \cdot, \cdot \rangle$ correspondant au produit scalaire dans $L^2(\Omega, \mathcal{F}, P)$ défini pour X et Y dans $L^2(\Omega, \mathcal{F}, P)$ par $\langle X, Y \rangle = \mathbb{E}(XY)$.

1.3.1 Equation de prédiction

Pour prédire la valeur du processus à la date t à partir d'une combinaison linéaire des n dernières échantillons du passé X_{t-1}, \dots, X_{t-n} . La meilleure combinaison linéaire (i.e le prédicteur linéaire optimale) est la projection orthogonale de X_{n+1} sur $\mathcal{H}_{k,l} = \overline{\text{sp}}\{X_j, k \leq j \leq l\}$ où $\overline{\text{sp}}\{A\}$ est le sous espace engendré par les éléments de A , qui est un sous espace fermé de $L^2(\Omega, \mathcal{F}, P)$ note $Pr(X_{n+1}/\mathcal{H}_{1,n})$ où

$$Pr(X_{n+1}/\mathcal{H}_{1,n}) = \text{vect}\{X_{t-1}, \dots, X_{t-n}\}$$

Ainsi: $\mathcal{H}_{1,n}$ est le sous-espace vectoriel engendré par les n observations précédant X_{t-1} à savoir X_{t-1}, \dots, X_{t-n} . D'après le théorème de projection,

$$\hat{X}_{n+1} = Pr(X_{n+1}/\mathcal{H}_{1,n}) = \sum_{j=1}^n \phi_{nj} X_{n+1-j}, \quad n \geq 1 \quad (1.1)$$

Où les coefficients $\phi_{n,j_{1 \leq j \leq n}}$ vérifient les équations de prédiction suivantes:

$$\left\langle X_{n+1} - \sum_{i=1}^n \phi_{ni} X_{n+1-i}, X_{n+1-j} \right\rangle = 0, \quad j = 1, \dots, n. \quad (1.2)$$

Cette équation se réécrit encore sous la forme

$$\left\langle X_{n+1}, X_{n+1-j} \right\rangle = \left\langle \sum_{i=1}^n \phi_{ni} X_{n+1-i}, X_{n+1-j} \right\rangle, \quad j = 1, \dots, n \quad (1.3)$$

En utilisant la linéarité de produit scalaire, l'équation 1.3 est équivalente à:

$$\sum_{i=1}^n \phi_{ni} \gamma(i-j) = \gamma(j), \quad j = 1, \dots, n. \quad (1.4)$$

En posant Γ_n la matrice de covariance du vecteur $(X_{t-1}, X_{t-2}, \dots, X_{t-n})$ définie par:

$$\Gamma_n = \begin{pmatrix} \gamma(0) & \gamma(1) & \dots & \gamma(n-1) \\ \gamma(1) & \gamma(0) & \gamma(1) & \vdots \\ \vdots & \ddots & \ddots & \ddots \\ \vdots & & & \gamma(1) \\ \gamma(n-1) & \gamma(n-2) & \dots & \gamma(1) & \gamma(0) \end{pmatrix}$$

ou de manière plus compacte, on peut réécrire 1.4 comme suit:

$$\Gamma_n \Phi_n = \gamma_n \tag{1.5}$$

où $\Phi_n = (\phi_{n,1}, \dots, \phi_{n,n})'$, $\gamma_n = (\gamma(1), \gamma(2), \dots, \gamma(n))'$, $\Gamma_n = [\gamma(i-j)]_{i,j=1, \dots, n}$

Définition 1.4. Nous appelons dans la suite erreur de prédiction direct d'ordre n ou innovation partielle d'ordre n le processus

$$\epsilon_n = X_{n+1} - \hat{X}_{n+1} = X_{n+1} - \sum_{j=1}^n \phi_{nj} X_{n+1-j} \tag{1.6}$$

La variance de l'erreur de prédiction direct d'ordre n est notée ν_n et définie par

$$\nu_n = \| X_{n+1} - \hat{X}_{n+1} \|^2 = \mathbb{E}(X_{n+1} - \hat{X}_{n+1})^2.$$

D'après 1.1, la variance de l'erreur de prédiction directe d'ordre n a pour expression

$$\nu_n = \langle X_{n+1}, X_{n+1} - \hat{X}_{n+1} \rangle = \gamma(0) - \sum_{j=1}^n \phi_{nj} \gamma(n) = \gamma(0) - \Phi_n' \gamma_n. \tag{1.7}$$

Les équations 1.5 et 1.7 sont appelées les équation de Yule-Walker. Notons que 1.5 a une unique solution si et seulement si la matrice Γ_n est inversible auquel cas la solution vout:

$$\Phi_n = \Gamma_n^{-1} \gamma_n \tag{1.8}$$

La propositin suivante fournit les condition suffisantes assurant que Γ_n est inversible pour tout n . On a ainsi des conditions sous lesquelles on peut calculer le prédicteur de X_{n+1} à partir de $X_{t-1}, X_{t-2}, \dots, X_{t-n}$.

Proposition 1.1. Si $\gamma(0) > 0$ et $\gamma(h) \rightarrow 0$ lorsque $h \rightarrow \infty$ alors la matrice de covariance Γ_n est inversible pour tout n .

Corollaire 1.1. *Sous les hypothèses de la proposition précédente le meilleur prédicteur linéaire \hat{X}_{n+1} de X_{n+1} en fonction de X_1, \dots, X_n est*

$$\hat{X}_{n+1} = \sum_{i=1}^n \phi_{ni} X_{n+1-i}, \quad n = 1, 2, \dots \quad (1.9)$$

ou $\Phi_n = (\phi_{n1}, \dots, \phi_{nn}) = \Gamma_n^{-1} \gamma_n$, $\gamma_n = (\gamma(1), \dots, \gamma(n))'$ et $\Gamma_n = [\gamma(i-j)], \dots, i, j = 1, \dots, n$ de plus, l'erreur quadratique moyenne est $\nu_n = \gamma(0) - \gamma_n \Gamma_n^{-1} \gamma_n$.

1.3.2 Calcul des coefficients des prédicteurs

L'équation 1.8 donne l'expression du prédicteur mais nécessite le calcul de l'inverse de la matrice Γ_n . Dans ce qui suit, nous décrivons un algorithme récursif simple et efficace qui permet d'obtenir l'expression du meilleur prédicteur linéaire \hat{X}_{n+1} défini dans 1.1, sans avoir à calculer explicitement l'inverse de Γ_n .

1.3.3 Algorithme de Durbin-Levinson

L'algorithme de Durbin-Levinson donne un schéma récursif pour calculer les coefficients $\phi_n = (\phi_{n1}, \dots, \phi_{nn})'$ du prédicteur \hat{X}_{n+1} définie dans 1.9 vérifiants $\phi_{11} = \gamma(1)/\gamma(0)$, $\nu_0 = \gamma(0)$,

Proposition 1.2. *Soit X_t un processus stationnaire, centré de fonction d'autocovariance $\gamma(\cdot)$ telle que $\gamma(0) > 0$ et $\gamma(h) \rightarrow 0$ quand $h \rightarrow \infty$, alors les coefficients ϕ_{nj} et l'erreur quadratique moyenne ν_n définie dans 1.9 et 1.7 vérifient $\phi_{11} = \gamma(1)/\gamma(0)$, $\nu_0 = \gamma(0)$,*

$$\phi_{nn} = \left[\gamma_n - \sum_{j=1}^{n-1} \phi_{n-1,j} \gamma(n-1) \right] \nu_{n-1}^{-1} \quad (1.10)$$

$$\begin{bmatrix} \phi_{n1} \\ \vdots \\ \phi_{n,n-1} \end{bmatrix} = \begin{bmatrix} \phi_{n-1,1} \\ \vdots \\ \phi_{n-1,n-1} \end{bmatrix} - \phi_{nn} \begin{bmatrix} \phi_{n-1,n-1} \\ \vdots \\ \phi_{n-1,1} \end{bmatrix} \quad (1.11)$$

et

$$\nu_n = \nu_{n-1} [1 - \phi_{nn}^2] \quad (1.12)$$

Preuve. *L'idée principale est de composer le sous espace $\mathcal{H}_{1,n}$ en deux sous espaces orthogonaux. par définition $\mathcal{H}_{2,n} = \overline{\text{sp}}\{X_2, \dots, X_n\}$ et $\mathcal{H}_{1,1}^\perp \{X_1 - Pr_{\mathcal{H}_{2,n}} X_1\}$ sont deux espaces orthogonaux de $\mathcal{H}_{1,n}$. Ce qui implique que $\forall Y \in L^2(\Omega, \mathcal{F}, P)$, $Pr_{1,n} Y = Pr_{2,n} Y + Pr_{1,1}^\perp Y$ où*

$Pr_{1,1}^\perp$ est la projecteur sur $\mathcal{H}_{1,1}^\perp\{X_1 - Pr_{2,n}X_1\}$.

Donc

$$\hat{X}_{n+1} = Pr_{2,n}X_{n+1} + Pr_{1,1}^\perp X_{n+1} = Pr_{2,n}X_{n+1} + a(X_1 - Pr_{2,n}X_1), \quad (1.13)$$

où

$$a = \langle X_{n+1}, X_1 - Pr_{2,n}X_1 \rangle / \| X_1 - Pr_{2,n}X_1 \|^2 \quad (1.14)$$

La stationnarité entraîne que les vecteurs $(X_1, \dots, X_n)'$ et $(X_2, \dots, X_{n+1})'$ ont une même matrice d'autocovariance et pour la même raison, les vecteurs $(X_n, X_{n-1}, \dots, X_1)'$ et $(X_2, \dots, X_{n+1})'$ ont également la même matrice d'autocovariance, par conséquent

$$Pr_{2,n}X_1 = \sum_{j=1}^{n-1} \phi_{n-1,j} X_{j+1} \quad (1.15)$$

$$Pr_{2,n}X_{n+1} = \sum_{j=1}^{n-1} \phi_{n-1,j} X_{n+1-j} \quad (1.16)$$

et

$$\| X_1 - Pr_{2,n}X_1 \|^2 = \| X_{n+1} - Pr_{2,n}X_{n+1} \|^2 = \| X_n - \hat{X}_n \|^2 = \nu_{n-1}. \quad (1.17)$$

En utilisant 1.13, 1.15 et 1.16, on obtient:

$$\hat{X}_{n+1} = aX_1 + \sum_{j=1}^{n-1} [\phi_{n-1,j} - a\phi_{n-1,n-j}] X_{n+1-j}. \quad (1.18)$$

De 1.13

$$\begin{aligned} a &= \left(\langle X_{n+1}, X_1 \rangle - \sum_{j=1}^{n-1} \phi_{n-1,j} \langle X_{n+1}, X_{j+1} \rangle \right) \nu_{n-1}^{-1} \\ &= \left[\gamma(n) - \sum_{j=1}^{n-1} \phi_{n-1,j} \gamma(n-j) \right] \nu_{n-1}^{-1} \end{aligned}$$

La proposition (1.1) nous assure l'unicité de représentation

$$\hat{X}_{n+1} = \sum_{j=1}^n \phi_{nj} X_{n+1-j}. \quad (1.19)$$

En identifiant les paramètres dans 1.18, 1.19, on déduit

$$\phi_{nn} = a, \quad (1.20)$$

et

$$\phi_{nj} = \phi_{n-1,j} - a\phi_{n-1,n-j}, \quad j = 1, \dots, n-1. \quad (1.21)$$

De 1.20 et 1.21 découle

$$\begin{aligned} \nu_n &= \| X_{n+1} - \hat{X}_{n+1} \|^2 \\ &= \| X_{n+1} - Pr_{2,n}X_{n+1} \|^2 + \| Pr_{1,1}^\perp X_{n+1} \|^2 - \langle X_{n+1} - Pr_{2,n}X_{n+1}, Pr_{1,1}^\perp X_{n+1} \rangle \\ &= \nu_{n-1} + a\nu_{n-1} - 2a\langle X_{n+1}, X_1 - Pr_{2,n}X_1 \rangle. \end{aligned}$$

En utilisant 1.17 et l'orthogonalité de $\mathcal{H}_{2,n}$ et $\mathcal{H}_{1,1}^\perp$, on obtient

$$\nu_n = \nu_{n-1}(1 - a^2) \quad (1.22)$$

1.4 Interpolation de données manquantes

1.4.1 Préliminaires

Soit $\{X_t\}$ une série chronologique centrée stationnaire de fonction d'autocovariance $\gamma(k)$. Notons $L^2(\Omega, \mathcal{F}, P)$ l'espace des variables aléatoires réelles définies sur (Ω, \mathcal{F}, P) pour lesquelles $\mathbb{E}(X^2) < \infty$, muni du produit scalaire $\langle X, Y \rangle = \mathbb{E}(XY)$.

Pour $t \in \mathbb{Z}$, nous définissons le sous espace

$$\mathcal{H}_t = \overline{\text{sp}}\{X_s, s \leq t\},$$

où $\overline{\text{sp}}\{A\}$ est le sous espace engendré par les éléments de A , qui est un sous espace fermé de $L^2(\Omega, \mathcal{F}, P)$.

On dit que le processus $\{X_t\}$ est non déterminable si $X_{t+1} \notin \mathcal{H}_t$.

Tout processus non déterminable peut se décomposer d'une manière unique. Comme d'après la décomposition de Wold

$$X_t = \sum_{k=0}^{\infty} b_k \epsilon_{t-k} + V_t, \quad b_0 = 1, \quad \sum_{k=0}^{\infty} b_k^2 < \infty \quad (1.23)$$

où (ϵ_t) est une suite de v.a non corrélées, avec $\text{var}(\epsilon_t) = \sigma^2$, ϵ_t et V_t sont non corrélées.

Lemme 1.1. *Supposons que $\{X_t\}$ un processus purment non déterminable de représentation $MA(\infty)$ donné par*

$$X_t = \sum_{i=0}^{\infty} c_i \epsilon_{t-i}$$

où $\epsilon_t = X_t - P_{t-1}X_t$, $c_0 = 1$, et $\sum_{i=0}^{\infty} c_i^2 < \infty$. La représentation autorégressive $AR(\infty)$ de $\{X_t\}$

$$X_t = \sum_{i=1}^{\infty} a_i X_{t-i}$$

Alors les paramètres de la représentation autorégressive et la représentation moyenne mobile sont liés par les relations.

$$\begin{aligned} a_0 &= 1 \\ a_i &= \sum_{j=0}^{\infty} a_j c_{i-j} \end{aligned}$$

Alors d'après le lemme précédent les coefficients de la représentation autorégressive d'un tel processus s'obtiennent par récurrence:

$$\begin{aligned} b_0 &= 1 \\ b_l &= \sum_{k=0}^{l-1} b_k a_{l-k} \quad (l = 1, 2, \dots). \end{aligned} \tag{1.24}$$

quand les coefficients (b_k) sont connues, les coefficients (a_k) sont obtenus par la résolution du système matriciel suivant

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ b_1 & 1 & 0 & 0 \\ b_2 & b_1 & 1 & 0 \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \cdot \\ \vdots \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \cdot \\ \vdots \end{bmatrix}$$

La décomposition de Wold engendre la décomposition utile de la matrice de covariance $\Gamma = (\gamma_{k-1})$ du processus $\{X_t\}$:

$$\Gamma = \sigma^2 \mathbf{T} \mathbf{T}' + \Gamma_V \tag{1.25}$$

où

$$\mathbf{T} = (b_{j-i})_{i=0, \infty, j=0, \infty} \tag{1.26}$$

Avec $b_0 = 1$, $b_k = 0$ pour $k < 0$, et Γ_V est la matrice de covariance du processus $\{X_t\}$.

Rappelons aussi que $\{X_t\}$ est purement non déterminable si $V_t = 0$ ou $\Gamma_V = 0$.

Pour $k > 0$, l'estimateur de l'observation X_k basé sur $\mathcal{H}_{-1} = \overline{sp}\{X_{-1}, X_{-2}, \dots\}$ est noté

par \hat{X}_k . Il est la projection orthogonale de X_k sur \mathcal{H}_{-1} . De 1.23, il suit

$$\begin{aligned}\hat{X}_k &= Pr_{-1}X_k = \sum_{i=k+1}^{\infty} b_i \epsilon_{k-i} + V_t \\ X_t - \hat{X}_k &= \sum_{i=0}^k b_i \epsilon_{k-i} \\ \text{var}(X_t - \hat{X}_k) &= \text{var}\left(\sum_{i=0}^k b_i \epsilon_{k-i}\right) = \sigma^2 \sum_{i=0}^k b_i^2 \\ \text{cov}(X_k - \hat{X}_k, X_l - \hat{X}_l) &= \sigma^2 \sum_{i=1}^{\min(k,l)} b_i^2 |l - k| \quad (l \geq 0).\end{aligned}\tag{1.27}$$

Il est important de signaler que $\text{cov}(X_k - \hat{X}_k, X_l - \hat{X}_l)$ est le (k, l) ème élément de la matrice $G = TT'$. Cette matrice joue un rôle important dans ce qui suit.

Il est bien connu que $\hat{X}_k, k \geq 0$, possède une représentation AR, en fonction des valeurs passées X_{-1}, X_{-2}, \dots :

$$\hat{X}_k \sim \sum_{i=1}^{\infty} a_{i,k} X_{-i} + V_k\tag{1.28}$$

et que les coefficients $(a_{i,k})$ de la représentation AR dans 1.28 satisfont

$$\begin{aligned}b_0 &= 1 \\ \sum_{k=0}^{l-1} b_k a_{l-k} &= b_{l+k} \quad (l = 1, 2, \dots).\end{aligned}\tag{1.29}$$

1.4.2 Interpolation d'une seule valeur manquante

Pour la clarté de la présentation, nous commençons par le cas où une seule valeur est manquante. Soit $X'_{0,r}$ la projection orthogonale de X_0 sur

$$\mathcal{H}'_{0,r} = \overline{\text{sp}}\{X_t; \quad t \leq r, t \neq 0\}.$$

Une difficulté essentielle dans la projection orthogonale sur le sous espace $\mathcal{H}'_{0,r}$, et que les sous espaces engendrés par le passé et le future ne sont pas orthogonaux. Il convient donc d'écrire le sous espace $\mathcal{H}'_{0,r}$ sous forme d'une somme directe de deux espaces orthogonaux.

Les deux sous espaces sont \mathcal{H}_{-1} et $S_r = \overline{\text{sp}}\{X_k - \hat{X}_k; \quad 1 \leq k \leq r\}$.

Lemme 1.2 (Pourahmadi, 1989). *Soit $\{X_t\}$ un processus stationnaire non déterminable.*

Alors, pour tout $0 \leq r < \infty$,

- (a) $X_0 \notin \mathcal{H}'_{0,r}$
- (b) \mathcal{H}_{-1} et S_r sont orthogonaux, et

$$\mathcal{H}'_{0,r} = \mathcal{H}_{-1} \oplus S_r.$$

Preuve. (a)

Supposons que $X_0 \in \mathcal{H}'_{0,r}$. Alors on peut l'écrire comme

$$X_0 = \sum_{k=0}^r c_r X_k + Y \tag{1.30}$$

avec $Y \in \mathcal{H}_{-1}$. Sans perte de généralité, nous supposons que $c_r \neq 0$. Ceci entraîne que

$$X_r = c_r^{-1} X_0 - \sum_{k=1}^{r-1} c_r^{-1} c_k X_k - c_r^{-1} Y \in \mathcal{H}_{r-1},$$

ce qui contredit que (X_t) est non déterminable. Par conséquent $X_0 \notin \mathcal{H}'_{0,r}$.

(b).

Tout $X \in \mathcal{H}'_{0,r}$ peut s'écrire sous la forme

$$X = Y + \sum_{k=1}^r c_k X_k = (Y + \sum_{k=1}^r c_r \hat{X}_k) + \sum_{k=1}^r (X_k - \hat{X}_k) \in \mathcal{H}_{-1} \oplus S_r.$$

Il est maintenant possible d'écrire

$$\hat{X}'_{0,r} = Pr_{\mathcal{H}'_{0,r}} X_0 = Pr_{\mathcal{H}_{-1}} X_0 + Pr_{S_r} X_0 \tag{1.31}$$

Une représentation explicite de $Pr_{S_r} X_0$ est fournie par le lemme suivant:

Lemme 1.3 (Pourahmadi, 1989.). *Si $\{X_t\}$ est un processus non déterminable alors*

$$Pr_{S_r} X_0 = \sum_{i=1}^r c_{i,r} (X_i - \hat{X}_i),$$

où le vecteur $c = (c_1, \dots, c_r)'$ vérifie

$$A_r c = b$$

et $b = (b_1, \dots, b_r)'$ représente les premières composantes de la représentation moyenne mobile de $\{X_t\}$,

$$A_r = T_r' T_r + b b'$$

avec

$$T_r = (b_{j-i})_{i,j=0,\dots,r-1}.$$

Le théorème qui suit donne l'expression explicite de la meilleure interpolation linéaire d'une seule valeur manquante.

Théorème 1.1 (Pourahmadi, 1989). *Soit $\{X_t\}$ un processus stationnaire non déterminable de représentation autorégressive (AR) de paramètre (a_k) , et soit $\hat{X}'_{0,r}$ la meilleure interpolation linéaire de X_0 basée sur $\{X_t, t \leq r, t \neq 0\}$, alors*

(a)

$$\hat{X}'_{0,r} = \hat{X}_0 + \sum_{k=1}^r c_{k,r} (X_k - \hat{X}_k), \quad (1.32)$$

où

$$c_{k,r} = (1 + \sum_{i=1}^r a_i^2)^{-1} (a_k - \sum_{i=1}^{r-k} a_i a_{i+k}) \quad (k = 1, 2, \dots, r)$$

(b)

$$\text{var}(X_0 - \hat{X}'_{0,r}) = \sigma^2 (1 + \sum_{i=1}^r a_i^2)^{-1} \quad (1.33)$$

Preuve. La démonstration de 1.33 découle du lemme 1.2.

Pour démontrer 1.32, remarquons d'abord que

$$X_0 - \hat{X}'_{0,r} = X_0 - \hat{X}_0 - \sum_{k=1}^r c_{k,r} (X_k - \hat{X}_k).$$

Comme $\hat{X}'_{0,r}$ est orthogonale à $(X_0 - \hat{X}'_{0,r})$, et

$$\text{var}(X_0 - \hat{X}'_{0,r}) = \sigma^2,$$

on déduit $\text{var}(X_0 - \hat{X}'_{0,r}) = \text{cov}(X_0 - \hat{X}'_{0,r}, X_0)$

$$\begin{aligned} &= \text{cov}(X_0 - \hat{X}_{0,r}, X_0) - \sum_{k=1}^r c_{k,r} \text{cov}(X_k - \hat{X}_k, X_0) \\ &= \text{var}(X_0 - \hat{X}_0) - \sigma^2 \sum_{k=1}^r c_{k,r} b_k \end{aligned}$$

$$\begin{aligned}
&= \sigma^2(1 - b' A_r^{-1} b) \\
&= \sigma^2(1 + a' a).
\end{aligned}$$

1.4.3 Interpolation de plusieurs valeurs manquantes

Le schéma précédent s'applique tout à fait à plusieurs valeurs manquantes et fournit une procédure d'estimation de ces dernières en usant de la même technique. Ainsi, nous résoudrons ce problème sans restriction sur le modèle et nous distinguons le cas particulier où les indices des valeurs manquantes sont successifs du cas où les indices sont quelconques. Le cas particulier a été l'objet des travaux de Brubaker et Wilson (1976) et Abraham (1981). Pourhadi (1989) s'est intéressé en suite à la généralisation de ces résultats, d'autant plus c'est le cas le plus fréquent dans la pratique. Soit

$$M = \{n_1, \dots, n_m\} \quad n_1 = 0 < n_2 < \dots < n_m$$

Les indices des valeurs manquantes du processus, et $m = \text{card}M$ est le nombre total des valeurs manquantes. Soit

$$K = \{k_1, \dots, k_r\}, \quad 0 < k_1 < k_2 < \dots < k_r < \infty$$

l'ensemble des indices des valeurs observées après l'instant $t = 0$. Nous utilisons le vecteur

$$X_M = (X_{n_1}, \dots, X_{n_m})'$$

Pour caractériser les données manquantes et le vecteur

$$X_K = (X_{k_1}, \dots, X_{k_r})'$$

Pour caractériser les données observées.

Il est à noter que le lemme 1.2 est démontré d'une manière analogue quand plusieurs valeurs sont manquantes, L'équivalent de la matrice A_r dans le lemme 1.2 est

$$A_K = \sigma^{-2} \{ \text{cov}(X_{k_i} - \hat{X}_{k_i}, X_{k_j} - \hat{X}_{k_j}) \}_{i,j=1,\dots,r}$$

L'équivalent du théorème 1.1 s'annonce alors:

Théorème 1.2 (pourhadi, (1989)). *Soit $\{X_t\}$ un processus stationnaire non déterminable dont les indices des valeurs manquantes sont données par M . Alors, avec les notations précédentes, nous avons:*

(a) Pour tout $1 \leq j \leq m$

$$\hat{X}'_{n_j} = \hat{X}_{n_j} + \sum_{i=1}^r c_{i,j} (X_{k_i} - \hat{X}_{k_i}),$$

où les vecteurs

$$c_j = (c_{1j}, \dots, c_{rj})'$$

sont solutions des équations

$$A_K c_j = b_j.$$

L'interpolation des valeurs manquantes peut s'écrire sous la forme vectorielle suivante

$$\hat{X}'_M = \hat{X}_M + C'(X_K - \hat{X}_K),$$

où C est une $r * m$ vérifiant

$$A_K C = B.$$

(b)

$$\text{cov}(X_M - \hat{X}'_M) = \text{cov}(X_M - \hat{X}_M) - \sigma^2 B' A_K^{-1} B.$$

1.4.4 Algorithme EM

L'algorithme d'espérance maximisation (EM) est une méthode itérative efficace pour calculer l'estimation de vraisemblance maximales en présence de données manquantes. Le but est d'estimer les paramètres du modèle pour lesquelles les données observées sont les plus vraisemblables tout en tenant compte de l'existence de données manquantes. Le problème de la maximisation de la vraisemblance consiste à trouver les paramètres qui incomplètes observées, on assume qu'il existe un ensemble complet de donnée $X = (X_K, X_M)$ où X_M représente l'ensemble des données manquantes. On peut donc écrire la fonction de densité jointe comme suit:

$$P(X/\theta) = P(X_K, X_M/\theta) = P(X_M/X_K, \theta).P(X_K, \theta)$$

et de même la log-vraisemblance de l'ensemble complet de données:

$$L(\theta/X) = L(\theta/X_K, X_M) = \ln(P(X_K, X_M/\theta)).$$

On doit donc trouver la valeur espérée de la log-vraisemblance de l'ensemble complet de donnée X par rapport aux données manquante X_M sachant les données observées X_K et les paramètres du modèle θ . Le problème est qu'on ne connaît pas les variables cachées, il faut donc utiliser les données et les paramètres d'une itération précédente X_K et $\theta^{(i-1)}$. La définition de cette espérance est:

$$Q(\theta, \theta^{(i-1)}) = \mathbb{E}_{X_M}[\ln(P(X_K, X_M/\theta)/X_K, \theta^{(i-1)})],$$

où \mathbb{E}_{X_M} est l'espérance par rapport à X_M , $\theta^{(i-1)}$ les paramètres utilisés pour évaluer l'espérance et θ les nouveaux paramètres à optimiser pour maximiser Q . C'est par le calcul de cette espérance conditionnelle qu'on mesure la log-vraisemblance. Dans cette expression, X_K et $\theta^{(i-1)}$ sont constants, θ est la variable ajustable et X_M est une variable aléatoire gouvernée par la distribution $f(X_M/X_K, \theta^{(i-1)})$. On peut alors réécrire le terme de droite comme:

$$\mathbb{E}_{X_M}[\ln(P(X_K, X_M/\theta)/X_K, \theta^{(i-1)})] = \int_{x_m} \ln(P(X_K, x_m/\theta))f(x_m/X_K, \theta^{(i-1)})d_{x_m}.$$

Cette fonction est déterminable et peut donc être maximisée. Une fois ces quantités définies, tout est prêt pour l'optimisation.

Les étapes de l'algorithme *EM* sont donc:

1. Étape E: Évaluation de l'espérance $Q(\theta, \theta^{(i-1)})$ selon les données observées et les paramètres à notre disposition.
2. Étape M: Maximisation de cette espérance $\theta^{(i)} = \max_{\theta} Q(\theta, \theta^{(i-1)})$ selon θ .

Ces étapes sont répétées autant de fois que nécessaire. Chaque itération fait augmenter la log-vraisemblance, donc l'algorithme converge directement vers un maximum local de la fonction de log-vraisemblance.

En utilisant les notations de la section 4, dans ce qui suit, nous appliquons l'algorithme *EM* pour estimer les paramètres du modèles et les valeurs manquantes en s'appuyant sur le théorème 2.2.

- Étape E: Sachant θ , l'estimateur des valeurs manquantes s'obtiennent par le théorème 2.2

$$\hat{X}'_M = \hat{X}_M + C'(X_K - \hat{X}_K),$$

où $\hat{X}_K = (\hat{X}_{n_1}, \dots, \hat{X}_{n_m})'$ et $\hat{X}_M = (\hat{X}_{k_1}, \dots, \hat{X}_{k_r})'$ sont les prédicteurs de X_M et X_K basés sur le passé fini $X_1, X_2, \dots, X_{n_1-m}$, $\hat{X}'_M = (\hat{X}'_{k_1}, \dots, \hat{X}'_{k_r})'$ est le vecteurs des interpolateurs des valeurs manquantes X_M basé sur $\{X_t; t \in M^c = \{1, 2, \dots, N\}/M\}$

- Etape M: Après l'estimation des données manquantes, les données "complètes" peuvent être utilisées pour maximiser la de ce problème de maximisation.

Le calcul de ce nouvel estimateur permet de retourner à l'étape du calcul de l'espérance et ainsi de suit. une fois le seuil de tolérance est atteint on possède l'estimation de la vraisemblance maximale de notre ensemble complet de données X .

Remarque 1.2. Comme tout procédé itératif, la mise en oeuvre de cet algorithme nécessite

une valeurs initiale $\theta^{(0)}$. Le choix de cette dernière dépend de la nature des données et de la fonction d'autocovariance $\gamma(k)$ du processus. En effet, trois cas sont possible :

- (a) La fonction $\gamma(k)$ est connue;
- (b) On dispose d'un long segment d'observations;
- (c) On ne dispose pas la long segment d'observations;

Dans le cas de la situation (a), une seul utilisation de l'étape E pour estimer les valeurs manquantes. Dans le cas de la situation (b), on utilise le segment pour estimer les paramètres du modèles (procédure de Box-Jenkins), la convergence de l'algorithme est obtunue en peu d'intération. Dans le dernier cas, nous supposons que les données peuvent être modélisées par un ARMA(p,q) fini d'équation

$$X_t = \sum_{i=1}^p \phi_i X_{t-i} + \epsilon_t + \sum_{j=1}^q \theta_j \epsilon_{t-j}.$$

Ainsi, zéro est la valeur initiale la plus appropriée, i.e. nous supposons que X_1, X_2, \dots, X_N sont collectées suivant un processus bruit blanc. Dans ce cas, l'interpolation des valeurs manquantes est la moyenne des valeurs observées.

Chapitre 2

Estimations des modèles ARCH en présence de données manquantes

2.1 Introduction

Dans ce chapitre, nous allons présenter une méthode d'estimation des paramètres du modèles ARCH en présence de données manquantes. C'est une généralisation de la méthode de Bose et Mukherjee (2003) réalisée par P. Bondon en 2012, au cas, où le processus ARCH est partiellement observé. Elle consiste à appliquer la méthode des moindres carrés en deux étapes. La première donnera lieu à un estimateur préliminaire $\hat{\alpha}_{pr}$, qui sera utilisé pour construire l'estimateur d'intérêt $\hat{\alpha}$. Ce dernier est moins bon que celui du maximum de vraisemblance mais présente l'avantage de posséder une formule explicite. Suivant Parzan (1963), une série de temps avec observations manquantes peut être considérée comme une version modulé au amplitude de la série temporelle originale, *i.e.*

$$X_t^* = a_t X_t \tag{2.1}$$

où X_t^* est supposée défini pour tout les temps les v.a. a_t sont

$$a_t = \begin{cases} 1 & \text{si } X_t \text{ est observée} \\ 0 & \text{si } X_t \text{ est manquante} \end{cases} \tag{2.2}$$

et X_t représente la valeur réelle observée.

Rappelons au passage, que l'estimateur de Bose et Mukherjee (2003) vérifie la consistance et la normalité asymptotique mais nécessite l'existence de moment d'ordre 4 pour la consistance et d'ordre 8 pour la normalité. Nous allons dans ce chapitre rappeler les définitions d'un modèle GARCH/ARCH ainsi l'étude de la stationnarité stricte et la stationnarité faible de ces modèles. ensuite, nous étudions la normalité asymptotique est la consistance

de l'estimateurs du (M.C.O) en deux étapes dans le cas où les données sont complete. L'objectif principal sera présenté à la dernière section. Il consiste à généraliser les résultats de consistance et de normalité asymptotique aux processus ARCH partiellement observés.

2.2 Modèles ARCH-GARCH

Nous donnons une première définition d'un processus GARCH fondé sur les deux premiers moments conditionnelles de $\{X_t\}$.

Définition 2.1 (Modèle GARCH faible). On dit que X_t est un processus GARCH(p,q) semi fort (faible) si les deux condition suivantes sont satisfaites:

- (i) $\mathbb{E}(X_t/X_{t-1}) = 0$ pour tout $t \in \mathbb{Z}$.
- (ii) il existe une constante $\alpha_0, \alpha_i, i = 1, \dots, q$ et $\beta_j, j = 1, \dots, p$ telle que

$$\sigma_t^2 = \text{var}(X_t/X_{t-1}) = \alpha_0 + \sum_{i=1}^q \alpha_i X_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2, \quad t \in \mathbb{Z} \quad (2.3)$$

où p et q sont des entiers.

Remarque 2.1. L'équation 2.3 peut être écrite de manière symbolique sous la forme plus compact

$$\sigma_t^2 = \alpha(B)X_t^2 + \beta(B)\sigma_t^2 \quad t \in \mathbb{Z}, \quad (2.4)$$

où B est l'opérateur retard ($B^i X_t^2 = X_{t-i}^2$ et $B^i \sigma_t^2 = \sigma_{t-i}^2$ pour tout entier i), α et β sont les polynômes de degré q et p :

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

si $\beta(B) = 0$ on a

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i X_{t-i}^2, \quad (2.5)$$

et le processus est appelé ARCH(q). L'innovation du processus X_t^2 est par définition la variable $\eta_t = X_t^2 - \sigma_t^2$. En remplaçant, dans l'équation 2.3 les variables σ_{t-j}^2 par $X_{t-j}^2 - \eta_{t-j}$ on obtient la représentation

$$X_t^2 = \alpha_0 + \sum_{i=1}^r (\alpha_i + \beta_j) X_{t-i}^2 + \eta_t - \sum_{j=1}^p \beta_j \eta_{t-j}, \quad t \in \mathbb{Z} \quad (2.6)$$

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

On retrouve ainsi dans cette équation la structure linéaire des modèles ARMA, permettant par exemple un calcul très simple des prévisions linéaires. Sous des hypothèses supplémentaires (impliquant la stationnarité de $\{X_t^2\}$, on peut affirmer que si $\{X_t\}$ est un GARCH(p,q), $\{X_t^2\}$ est un processus ARMA(r,p). En particulier, le carré d'un processus ARCH(q) admet, s'il est stationnaire, une représentation AR(q). Ces représentations ARMA seront utiles pour l'estimation et l'identification des processus GARCH. Elle seront en revanche du peu d'utilité pour l'étude de la stationnarité du processus $\{X_t\}$ car le bruit η_t dépend, par construction, du passé de $\{X_t\}$.

La définition 1.1 ne fournit pas directement de processus la vérifient. La définition plus restrictive suivante permettra d'obtenir explicitement des processus solutions.

Définition 2.2 (Modèle GARCH fort). $\{X_t\}$ est un processus GARCH(p,q) au sens fort s'il vérifie

$$\begin{cases} X_t = \sigma_t(\alpha)\epsilon_t \\ \sigma_t^2(\alpha) = \alpha_0 + \sum_{i=1}^q \alpha_i X_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2 \end{cases} \quad (2.7)$$

où $\alpha_0 > 0, \alpha_i \geq 0, i = 1, \dots, q$ et $\beta_j \geq 0, j = 1, \dots, p$

et $\{\epsilon_t\}$ est une suite de variable i.i.d, avec $\mathbb{E}(\epsilon_t) = 0$ et $\mathbb{E}(\epsilon_t^2) = 1$. En remplaçant X_{t-1} par $\sigma_{t-1}\epsilon_{t-1}$ dans l'équation 2.3 on obtient:

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

que l'on peut écrire

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^r a(\epsilon_{t-1}) \sigma_{t-i}^2.$$

où $r = \max(p, q)$ et $a(z) = \alpha z^2 + \beta, i = 1, \dots, r$. Cette représentation montre que dans le cas d'un GARCH fort, le processus de volatilité vérifie une équation autorégressive, mais avec coefficients aléatoires.

Cependant il est clair qu'un processus GARCH fort tel que σ_t^2 est mesurable par rapport à la tribu $\{X_u, u < t\}$ est un processus GARCH au sens de la définition 1.1. La réciproque n'est cependant pas vraie.

2.2.1 Stationnarité forte d'un GARCH

Nous examinons d'abord le cas du modèle GARCH(1,1) qui peut se traiter avec des techniques élémentaires.

GARCH(1,1)

Dans le cas où $p = q = 1$ dans le modèle 2.7 on obtient le modèle GARCH(1,1) fort défini

$$\begin{cases} X_t = \sigma_t(\alpha)\epsilon_t \\ \sigma_t^2(\alpha) = \alpha + \alpha_1 X_{t-1}^2 + \beta \sigma_{t-1}^2 \end{cases} \quad (2.8)$$

Théorème 2.1 (Nelson 1990:Stationnarité forte d'un GARCH(1,1)). *Si*

$$\gamma = \mathbf{E}(\log\{\alpha\epsilon_t^2 + \beta\}) < 0$$

la série

$$h_t = \left\{ 1 + \sum_{i=1}^{\infty} a(\epsilon_{t-1}) + \cdots + a(\epsilon_{t-i}) \right\} \alpha_0$$

converge presque sûrement *p.s.*, et le processus $\{X_t\}$ est défini $X_t = \sqrt{h_t}\epsilon_t$ est l'unique solution strictement stationnaire du modèle 2.8. Cette solution est non anticipative et ergodique.

Si $\gamma \geq 0$ et $\alpha_0 > 0$ il n'existe pas des solutions strictement stationnaire.

Remarque 2.2. 1. Les coefficients $\gamma = \mathbf{E}(\log\{\alpha\epsilon_t^2 + \beta\})$ existent toujours dans $]-\infty, +\infty[$. En effet

$$\gamma = \mathbf{E}(\log(\alpha\epsilon_t^2 + \beta)) < \mathbf{E}(\alpha\epsilon_t^2 + \beta) < \alpha + \beta < +\infty.$$

2. La condition $\gamma < 0$ implique que $\beta < 1$.

Par absurde supposons que $\beta \geq 1$ donc :

$$\mathbf{E}(\log(\alpha\epsilon_t^2 + \beta)) = \mathbf{E}(\log[\beta(\frac{\alpha}{\beta}\epsilon_t^2 + 1)]) = \mathbf{E}(\log(\beta) - \log((\frac{\alpha}{\beta}\epsilon_t^2 + 1))) > 0$$

D'où la contradiction.

3. Si $\alpha + \beta < 1$ alors $\gamma < 0$. En effet par application de l'inégalité de Jensen

$$\mathbf{E}(\log(\alpha\epsilon_t^2 + \beta)) \leq \log(\mathbf{E}(\alpha\epsilon_t^2 + \beta)) = \log(\alpha + \beta) < 0.$$

4. Si la condition de la stationnarité stricte du GARCH(1,1) est vérifiée pour α et β alors elle est également satisfaite pour une paire (α_1, β_1) avec $\alpha_1 \leq \alpha$ et $\beta_1 \leq \beta$. En effet

$$(\mathbf{E}(\log(\alpha_1\epsilon_t^2 + \beta_1))) \leq \mathbf{E}(\log(\alpha\epsilon_t^2 + \beta)) < 0.$$

On a également, la stationnarité stricte d'un GARCH(1,1) implique celle du ARCH(1) modèle obtenu en annulant β . En effet

$$\mathbf{E}(\log(\alpha\epsilon_t^2)) \leq \mathbf{E}(\log(\alpha\epsilon_t^2 + \beta)) < 0.$$

Théorème 2.2 (Bougerol et Picard, 1992: Stationnarité forte d'un GARCH(p,q)).

Le modèle GARCH (p, q) a une solution strictement stationnaire ssi

$$\gamma(A) < 0$$

où $\gamma(A)$ est le plus grand exposant de Lyapov de la suite de matrices $\{A_t, t \in \mathbb{Z}\}$ où

$$A_t = \begin{pmatrix} \alpha_1 \epsilon_t^2 & \cdot & \cdot & \cdot & \cdot & \alpha_q \epsilon_t^2 & \beta_1 \epsilon_t^2 & \cdot & \cdot & \cdot & \beta_p \epsilon_t^2 \\ 1 & & & & 0 & 0 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & \cdot & & & 0 & 0 & \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot & \cdot & \cdot & & & & \cdot \\ 0 & & & & 1 & 0 & 0 & \cdot & \cdot & \cdot & 0 \\ \alpha_1 & \cdot & \cdot & \cdot & \cdot & \alpha_q & \beta_1 & \cdot & \cdot & \cdot & \beta_p \\ 0 & \cdot & \cdot & \cdot & \cdot & 0 & 1 & & & & 0 & 0 \\ \cdot & & & & & \cdot & 0 & \cdot & & & 0 & \cdot \\ \cdot & & & & & \cdot & \cdot & \cdot & & & \cdot & \cdot \\ \cdot & & & & & \cdot & \cdot & \cdot & \cdot & & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot & \cdot & 0 & 0 & & & & 1 & 0 \end{pmatrix} \quad (2.9)$$

lorsqu'elle existe la solution stationnaire est unique, anticipative et ergodique.

2.2.2 Stationnarité faible d'un GARCH

Les théorèmes suivant donnent des conditions nécessaires et suffisantes de stationnarité au second ordre.

Théorème 2.3 (stationnarité faible d'un GARCH(1,1)). *Supposons $\alpha_0 \geq 0$*

- *Le modèle GARCH(1,1) a une solution stationnaire au second ordre non anticipative et ergodique ssi*

$$\alpha + \beta < 1$$

De plus, le processus solution $\{X_t\}$ est un bruit blanc et il n'existe pas d'autre solution stationnaire au second ordre et non anticipative.

- *Si $\alpha + \beta \geq 1$ alors il n'existe pas de solutions GARCH(1,1) non anticipative et stationnaire au second ordre.*

Théorème 2.4 (Bollerslev, 1986: stationnarité faible d'un GARCH(p,q)). *S'il existe un processus GARCH(p,q), au sens de la Définition 1.1, stationnaire au second ordre et non anticipative, et si $\alpha > 0$, alors*

$$\sum_{i=1}^q \alpha_i + \sum_{j=1}^p \beta_j < 1 \quad (2.10)$$

Inversement, si 2.10 est vérifiée, l'unique solution strictement stationnaire du modèle 2.7 est un bruit blanc (donc stationnaire au second ordre). Il n'existe pas d'autres solutions stationnaires au second ordre. Autrement dit, Le modèle GARCH(p, q) a une solution stationnaire au second ordre non anticipative ssi 2.10 est vérifiée.

Si

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

il n'existe pas de solution GARCH(p, q) non anticipative et stationnaire au second ordre.

Remarque 2.3. 1. Sous les conditions du théorème, l'unique solution stationnaire du modèle 2.4 est un bruit blanc de variance

$$\mathbf{V}(X_t) = \frac{\alpha_0}{1 - \left(\sum_{i=1}^q \alpha_i + \sum_{j=1}^p \beta_j \right)}.$$

En effet

$$\mathbf{V}(X_t) = \mathbf{E}(\mathbf{X}_t^2) = \mathbf{E}(\mathbf{E}(X_t^2/X_u \quad u < t)) = \mathbf{E}(\sigma_t^2)$$

et comme

$$\begin{aligned} \mathbf{E}(\mathbf{X}^2) &= \mathbf{E}\left(\alpha_0 + \sum_{i=1}^q \alpha_i X_{t-1}^2 + \sum_{j=1}^p \beta_j \sigma_{t-1}^2\right) \\ &= \alpha_0 + \sum_{i=1}^q \alpha_i \mathbf{E}(X_{t-1}^2) + \sum_{j=1}^p \beta_j \mathbf{E}(\sigma_{t-1}^2) \\ &= \alpha_0 + \sum_{i=1}^q \alpha_i \mathbf{E}(X_t^2) + \sum_{j=1}^p \beta_j \mathbf{E}(X_t^2) \\ &= \frac{\omega}{1 - \left(\sum_{i=1}^q \alpha_i + \sum_{j=1}^p \beta_j \right)} \end{aligned}$$

alors

$$\mathbf{V}(X_t) = \frac{\omega}{1 - \left(\sum_{i=1}^q \alpha_i + \sum_{j=1}^p \beta_j \right)}$$

2. Les conditions des Théorèmes 1.2 et 1.4, étant nécessaires et suffisantes, on a forcément

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

puisque la solution stationnaire au second ordre du Théorème 1.4 l'est également strictement .

2.3 Estimation des modèles ARCH par la méthode des moindres carrés

Dans cette partie nous considérons l'estimation par la moindres carrés ordinaires (M.C.O) du modèle ARCH(q):

$$\begin{cases} X_t = \sigma_t(\alpha)\epsilon_t \\ \sigma_t^2(\alpha) = \alpha_0 + \sum_{i=1}^q \alpha_i X_{t-i}^2 \end{cases} \quad (2.11)$$

(ϵ_t) est une suite de variables aléatoire i.i.d, $\mathbb{E}(\epsilon_t) = 0$, $var(\epsilon_t) = 1$.

La méthode consiste à tirer partie de la représentation AR sur le carré du processus observé et à appliquer la méthode des moindres carrés quasi-généralisés (M.C.Q.G). Aucune hypothèse n'est fait sur la loi de ϵ_t . Les estimateurs obtenus sont, au moins pour n grand, moins précis que ceux du quasi-maximum de vraisemblance (Q.M.V) mais plus faciles à obtenir. Ils peuvent également fournir des valeurs initiales pour la procédure d'optimisation utilisée dans l'obtention d'estimateurs du (Q.M.V) plus précis.

La vrai valeur du vecteur des paramètres est noté $\theta_0 = (\alpha_0, \alpha_1, \dots, \alpha_q)$ et nous noterons θ une valeur quelconque.

On déduit de 2.11 la représentation AR(q)

$$X_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i X_{t-i}^2 + \eta_t \quad (2.12)$$

où $\eta_t = X_t^2 - \sigma_t^2 = (\epsilon_t^2 - 1)\sigma_t^2$. La suite $(\eta_t, \mathcal{F}_{t-1})_t$ consiste donc une différence de martingale. On suppose que l'on dispose d'observations X_t, \dots, X_n , réalisation partielle du processus (X_t) , de valeurs initiales X_0, \dots, X_{1-q} . Par exemple ces valeurs initiales peuvent être choisies nulles. Introduisant le vecteur

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

on déduit de 2.12 le système

$$X_t^2 = Z'_{t-1}\theta_0 + \eta_t, \quad t = 1, \dots, n \quad (2.13)$$

soit

$$Y = X\theta_0 + U$$

en définissant la matrice $n * q$ et les vecteurs $n * 1$

$$X = \begin{pmatrix} Z'_{n-1} \\ \vdots \\ Z'_0 \end{pmatrix}, \quad Y = \begin{pmatrix} X_n^2 \\ \vdots \\ X_1^2 \end{pmatrix}, \quad U = \begin{pmatrix} \eta_n \\ \vdots \\ \eta_1 \end{pmatrix}.$$

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

$$\hat{\theta}_n = (X'X)^{-1}X'Y. \quad (2.14)$$

nous serons amenés, pour établir la convergence, à considérer les hypothèses suivantes.

H1 . X_t est solution non anticipative strictement stationnaire du modèle 2.11.

H2 . $\mathbb{E}(X_t^4) < \infty$.

H3 . $P(\epsilon_t^2 = 1) \neq 1$.

Théorème 2.5 (Convergence des estimateurs (M.C.O) pour un ARCH). *Soit $(\hat{\theta}_n)$ une suite d'estimateurs satisfaisant 2.14 sous les hypothèses **H1-H3***

$$\hat{\theta}_n \xrightarrow{p.s} \theta_0 \quad n \rightarrow \infty.$$

Démonstration. La preuve comporte plusieurs étapes.

- i) Nous avons vu (Théorème 2.2) que l'unique solution stationnaire non anticipative (X_t) est ergodique. Le processus (Z_t) est également ergodique car Z_t s'écrit comme fonction mesurable des X_{t-i} . Le théorème ergodique appliqué au processus stationnaire (Z_t) entraîne

$$\frac{1}{n}X'X = \frac{1}{n} \sum_{t=1}^n Z_{t-1}Z'_{t-1} \xrightarrow{p.s} \mathbb{E}_{\theta_0}(Z_{t-1}Z'_{t-1}) \quad \text{quand } n \rightarrow \infty \quad (2.15)$$

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

$$\frac{1}{n}X'Y = \frac{1}{n} \sum_{t=1}^n Z_{t-1}X_t^2 \xrightarrow{p.s} \mathbb{E}_{\theta_0}(Z_{t-1}X_t^2) \quad \text{quand } n \rightarrow \infty \quad (2.16)$$

- ii) Montrons par l'absurde l'inversibilité de la matrice $\mathbb{E}_{\theta_0}Z_{t-1}Z'_{t-1} = \mathbb{E}_{\theta_0}Z_tZ'_t$.
 Supposons qu'il existe c vecteur non nul de \mathbb{R}^{q+1} tel que $c' \mathbb{E}_{\theta_0}Z_tZ'_t = 0$.
 Donc $\mathbb{E}_{\theta_0}[c'Z_t(c'Z_t)'] = 0$, d'où l'on déduit que $c'Z_t$ est p.s constant. Par suite, il existe une combinaison linéaire p.s égale à une constante des variables $X_t^2, \dots, X_{t-q+1}^2$.

On peut supposes sans perte de généralité que, dans cette combinaison, le coefficient de $X_t^2 = \epsilon_t^2 \sigma_t^2$ est 1. Donc ϵ_t s'exprime p.s comme fonction mesurable des variables X_{t-1}, \dots, X_{t-q} .

Or d'après le caractère non anticipatif de la solution, ϵ_t^2 est indépendante de ces variables.

Ceci implique que ϵ_t^2 est p.s égale à une constante. Cette constante ne peut être que 1, mais on aboutit alors à une contradiction avec **H3**.

Donc $\mathbb{E}_{\theta_0} Z_{t-1} Z'_{t-1}$ est inversible.

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

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

iv) Rappelons que le processus η_t est l'innovation forte de (X_t^2) . On a donc, en particulier, les relation d'orthogonalité

$$\mathbb{E}_{\theta_0}(\eta_t) = \mathbb{E}_{\theta_0}(\eta_t X_{t-1}^2) = \dots = \mathbb{E}_{\theta_0}(\eta_t X_{t-q}^2) = 0.$$

C'est-à-dire

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

d'où l'on déduit, en utilisant 2.13

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

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

□

Pour la normalité asymptotique de l'estimateur des (M.C.O), nous devons faire l'hypothèse supplémentaire

$$[\mathbf{H4}] \mathbb{E}_{\theta_0}(X_t^8) < \infty.$$

Introduisons les matrices carrées symétrique de la taille $q + 1$

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

L'inversibilité de la matrice \mathbf{A} à été établie dans la preuve du Théorème (2.5), celle de \mathbf{I} sera montrée dans la preuve du résultat suivant, qui établit la normalité asymptotique de

l'estimateur des (M.C.O).

on note $\mu_4 = \mathbb{E}(\epsilon_t^4)$.

Théorème 2.6. *Sous les hypothèses **H1** – **H4**,*

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

Démonstration. On a, d'après 2.13

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

Donc

$$\sqrt{n}(\hat{\theta}_n - \theta_0) = \left(\frac{1}{n} \sum_{t=1}^n Z_{t-1} Z'_{t-1} \right)^{-1} \left\{ \frac{1}{n} \sum_{t=1}^n Z_{t-1} \eta_t \right\}. \quad (2.17)$$

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

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

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

$$\frac{1}{\sqrt{n}} \sum_{t=1}^n Z_{t-1} \eta_t \xrightarrow{\mathcal{L}} \mathcal{N}(0, (\mu_4 - 1) \mathbf{I}). \quad (2.18)$$

On montre que cette loi limite est non dégénérée, c'est-à-dire que \mathbf{I} est inversible, par le même raisonnement que celui utilisée pour établir l'inversibilité de \mathbf{A} dans la preuve du Théorème (2.5). Par suite, on déduit de 2.14, 2.17 et 2.18, par un raisonnement classique, que $\sqrt{n}(\hat{\theta}_n - \theta_0)$ est asymptotiquement normal, de moyenne de vecteur nul, et de variance la matrice du théorème. \square

2.4 Estimation des modèles ARCH(q) par la méthode des (M.C.O) en présence des données manquantes

Tout d'abord nous allons construire un estimateur préliminaire $\hat{\alpha}_{pr}$ par la (M.C.O) qui vérifie la consistance et la normalité asymptotique. Il sera utilisé par la suite pour la construction de $\hat{\alpha}$.

2.4.1 Construction de l'estimateur préliminaire $\hat{\alpha}_{pr}$

Afin d'introduire l'estimateur préliminaire, posons $Y_t = X_t^2$
Ainsi,

$$Z_t = (1, Y_{t-1}, \dots, Y_{t-q})',$$

et $\eta_t = \epsilon_t^2 - 1$.

la deuxième équation de 2.11 sera équivalente à $\sigma_t^2 = Z_t' \alpha$, et en élevant la première équation au carré on obtient

$$Y_t = Z_t' \alpha + \sigma_t^2(\alpha) \eta_t \quad (2.19)$$

Où $\mathbb{E}(\sigma_t^2 \eta_t / \mathcal{F}_{t-1}^X) = \sigma_t^2 \mathbb{E}(\eta_t / \mathcal{F}_{t-1}^X) = 0$, \mathcal{F}_{t-1}^X est la σ -algèbre générée par $(X_j)_{j \leq t}$. On retrouve ainsi dans l'équation 2.19 la structure linéaire des modèles autorégressif d'ordre q, dont l'innovation est une différence de martingale. Il est important de noter que l'innovation est conditionnellement hétéroscédastique.

Suivant Parzen(1963), et d'après l'équation 2.2, nous sommes amenés, pour établir la convergence, à considérer les hypothèses suivantes.

H1 . Le processus (a_t) est strictement stationnaire.

H2 . Les processus (X_t) et (a_t) sont indépendants.

Lemme 2.1. Soit $A_t = \prod_{i=1}^p a_{t-i}$, et

$$Y_t^* = A_t Y_t = A_t Z_t' \alpha + A_t \sigma_t^2 \eta_t = Z_t^* \alpha + \sigma_t^{*2} \eta_t, \quad p+1 \leq t < n. \quad (2.20)$$

L'équation 2.20 coincide avec 2.19 quand $(p+1)$ données consécutives sont observées.

En ignorant la partie aléatoire de $\sigma_t^{*2}(\alpha)$ et la présence de α dans son expression, on obtient par les M.C.O un estimateur préliminaire que nous notons

$$\hat{\alpha}_{pr} = \left[\sum_{t=p+1}^n Z_t^* Z_t^{*'} \right]^{-1} \sum_{t=p+1}^n Z_t^* Y_t. \quad (2.21)$$

2.4.2 Comportement asymptotique de l'estimateur préliminaire $\hat{\alpha}_{pr}$

La consistance et la normalité asymptotique de $\hat{\alpha}_{pr}$ sont données par le théorème suivant.

Théorème 2.7 (P.Bondon (2012)). *i.* $\hat{\alpha}_{pr} \xrightarrow{p.s} \alpha$ si $\mathbb{E}(X_0^4) < \infty$.

ii. $\sqrt{n}(\hat{\alpha}_{pr} - \alpha) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, (\mathbb{E}[A_0])^{-1} \text{var}(\epsilon_0^2) \mathbb{E}[(\alpha' Z_0)^2 Z_0 Z_0'] \{\mathbb{E}[Z_0 Z_0']\}^{-1}\right)$ si $\mathbb{E}(X_0^8) < \infty$.

Remarque 2.4. En comparant les variances asymptotique de $\hat{\alpha}_{pr}$ et $\hat{\alpha}$ avec celles de l'estimateur de Bose et Mukherjee (2003), nous constatons que les données manquantes augmente les variances avec le facteur multiplicatif $(\mathbb{E}(A_0))^{-1}$. En particulier, la variance asymptotique de $\hat{\alpha}$ est égale à la variance asymptotique du (M.V.Q) avec aucune donnée manquante multiplié par $(\mathbb{E}(A_0))^{-1}$. Lorsque (a_t) est une sequence de Bernoulli independentes avec $q = P\{a_t = 1\}$, $(\mathbb{E}(A_0))^{-1} = q^{-(p+1)}$. Á partir de cette expression nous tirons deux conclusions suivantes:

- i) .lorsque $q \rightarrow 1$, la variance asymptotique de $\hat{\alpha}$ tend vers la variance asymptotique du (M.V.Q) sans données manquantes.
- ii) La variance asymptotique de $\hat{\alpha}$ augment à mesure que q diminue.

Remarque 2.5. Un autre estimateur de α peut être obtenue en appliquant la méthode de Yule-Walker proposé par Dunsuir et Robinson (1981) au modèle autorégressif. En effet il résulte de 2.19 que

$$\alpha_0 = \left(1 - \sum_{i=1}^p \alpha_i\right) \mathbb{E}Y_0, \quad (2.22)$$

avec $(\alpha_1, \dots, \alpha_q)$ est donné par la solution unique à l'équation de Yule-Walker

$$\begin{pmatrix} \gamma_Y(0) & \gamma_Y(1) & \dots & \gamma_Y(p-1) \\ \gamma_Y(1) & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \gamma_Y(1) \\ \gamma_Y(p-1) & \dots & \gamma_Y(1) & \gamma_Y(0) \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_p \end{pmatrix} = \begin{pmatrix} \gamma_Y(1) \\ \gamma_Y(1) \\ \vdots \\ \gamma_Y(p) \end{pmatrix}$$

Où $\gamma_Y(k) = \text{cov}(Y_0, Y_k)$. Les estimateurs $(\hat{\alpha}_{Y_{w,1}}, \dots, \hat{\alpha}_{Y_{w,q}})$ peuvent être obtenus en remplaçant $\gamma_Y(k)$ par un estimation appropriée dans le système ci dessus. D'après Dunsuir et Robinson, nous estimons $\gamma_Y(k)$ par

$$\hat{\gamma}(k) = \frac{\sum_{t=1}^{n-k} a_t a_{t-k} (Y_t - \hat{\mu}_Y)(Y_{t+k} - \hat{\mu}_Y)}{\sum_{t=1}^{n-k} a_t a_{t+k}},$$

où $\hat{\mu}_Y = \frac{\sum_{t=1}^n a_t Y_t}{\sum_{t=1}^n a_t}$. Puis $\hat{\mu}_{Yw,0}$ est obtenu en remplaçant dans la partie droite de 2.22 α_i par $\hat{\alpha}_{Yw,i}$ et $\mathbb{E}Y_0$ par $\hat{\mu}_Y$.

Preuve. D'après les équations 2.20 et 2.21, on a:

$$\hat{\alpha}_{pr} - \alpha = \left[n^{-1} \sum_{t=p+1}^n A_t^2 Z_t Z_t' \right]^{-1} n^{-1} \sum_{t=p+1}^n U_t \quad (2.23)$$

où $U_t = A_t^2 (\alpha' Z_t) Z_t \eta_t$. Comme (X_t) est stationnaire est ergodique et **H1** et **H2** sont vérifiées, alors d'après Hannan (1973), le processus combiné (X_t, a_t) est strictement stationnaire et ergodique, il s'en suit $(A_t^2 Z_t Z_t')$ et (U_t) sont strictement stationnaire et ergodique. Lorsque $\mathbb{E}(X_0^4) < \infty$ et $\mathbb{E}(A_0^2) < \infty$, les quantités $\mathbb{E}(A_t^2 Z_t Z_t')$ et $\mathbb{E}(U_t)$ sont finies, et d'après le théorème ergodique pour les séquences stationnaires (Stout, 1974), il découle

$$\frac{1}{n} \sum_{t=p+1}^n A_t^2 Z_t Z_t' \xrightarrow{p.s} \mathbb{E}[A_0^2] \mathbb{E}[Z_0 Z_0'], \quad (2.24)$$

$$\frac{1}{n} \sum_{t=p+1}^n U_t \xrightarrow{p.s} \mathbb{E}[A_0^2] \mathbb{E}[(\alpha' Z_0) Z_0] \mathbb{E}[\eta_0] = 0, \quad (2.25)$$

quand $n \rightarrow \infty$. Comme $A_0 \neq 0$ p.s alors $\mathbb{E}[A_0^2 Z_0 Z_0']$ est inversible. Donc **i.** se déduit de 2.23, 2.24, 2.41

Pour démontrer la normalité asymptotique, nous allons considérons \mathcal{F}_t la σ -algèbre générée par les variables $(U_j)_{j \leq t}$ et \mathcal{G}_t la σ -algèbre générée par les variables $(X_j)_{j \leq t}$. Nous avons $\mathbb{E}[U_t / \mathcal{G}_{t-1}] = A_t^2 (\alpha' Z_t) Z_t \mathbb{E}[\eta_t / \mathcal{G}_{t-1}]$. Il en résulte de **H2** que les variables $\{\epsilon_t, (X_j)_{j \leq t-1}\}$ et (a_t) sont mutuellement indépendantes. Ainsi, ϵ_t est indépendant de $(X_j)_{j \leq t-1}$ et donc indépendant de \mathcal{G}_{t-1} et $\mathbb{E}[\eta_t / \mathcal{G}_{t-1}] = \mathbb{E}[\eta_t] = 0$, ce qui entraîne que $\mathbb{E}[U_t / \mathcal{G}_{t-1}] = 0$. Comme $\mathcal{F}_t^U \subset \mathcal{G}_t$ alors $\mathbb{E}[U_t / \mathcal{F}_{t-1}^U] = \mathbb{E}[\mathbb{E}(U_t / \mathcal{G}_{t-1}) / \mathcal{F}_{t-1}^U] = 0$. En utilisant le théorème central limite de Billingsley (1961) pour les différences de martingales ergodiques et stationnaires, on obtient

$$\frac{1}{\sqrt{n}} \sum_{t=p+1}^n U_t \xrightarrow{\mathcal{L}} N(0, W). \quad (2.26)$$

Lorsque $\mathbb{E}[X_0^8] < \infty$, les éléments de $\mathbb{E}[(\alpha' Z_0)^2 Z_0 Z_0']$ sont finis. De plus pour tout $m > 0$, si $\mathbb{E}[X_0^m] < \infty$, alors $\mathbb{E}[\epsilon_0^m] < \infty$. Par conséquent, quand $\mathbb{E}[|X_0^8|] < \infty$ et $\mathbb{E}[A_0^4] < \infty$, les éléments de W sont finis et $W = \mathbb{E}(A_0^4) \mathbb{E}[(\alpha' Z_0)^2 Z_0 Z_0'] \text{var}(\epsilon_0^2)$. En utilisant 2.24, 2.26 et 2.23, on obtient

$$\sqrt{n}(\hat{\alpha}_{pr} - \alpha) \xrightarrow{\mathcal{L}} N \left[0, \frac{\mathbb{E}[A_0^2]}{\{\mathbb{E}[A_0^2]^2\}} \text{var}(\epsilon_0^2) (\mathbb{E}[Z_0 Z_0'])^{-1} \mathbb{E}[(\alpha' Z_0)^2 Z_0 Z_0'] \mathbb{E}[Z_0 Z_0']^{-1} \right],$$

qui est équivalent à **ii)** lorsque a_t est un processus à valeurs dans $\{0,1\}$

2.4.3 Construction de l'estimateur d'intérêt $\hat{\alpha}$

A présent, nous allons utiliser $\hat{\alpha}_{pr}$ pour construire un estimateur amélioré $\hat{\alpha}$. En divisant 2.20 par $\sigma_t^2(\alpha)$, on obtient

$$\frac{Y_t^*}{\sigma_t^2(\alpha)} = \frac{Z_t^{*'}}{\sigma_t^2(\alpha)}\alpha + A_t\eta_t.$$

Dans cette expression, l'erreur $A_t\eta_t$ est homoscédastique. En remplaçant α par $\hat{\alpha}_{pr}$ dans

σ_t^2

$$\frac{Y_t^*}{\sigma_t^2(\hat{\alpha}_{pr})} \approx \frac{Z_t^{*'}}{\sigma_t^2(\hat{\alpha}_{pr})}\alpha + A_t\eta_t.$$

Par suite, en procédant à l'estimation de α par les (M.C.O), on obtient

$$\hat{\alpha} = \left[\sum_{t=p+1}^n \frac{Z_t^* Z_t^{*'}}{\sigma_t^4(\hat{\alpha}_{pr})} \right]^{-1} \sum_{t=p+1}^n \frac{Z_t^* Y_t^*}{\sigma_t^4(\hat{\alpha}_{pr})} \quad (2.27)$$

Le théorème qui suit démontre la consistance et la normalité asymptotique de $\hat{\alpha}$.

Théorème 2.8 (Bondon (2012)). *Soit (X_t) un ARCH(q) défini par 1.9 où $\alpha_i > 0$ pour tout $i = 1, \dots, q$ et (a_t) un processus à valeurs dans $\{0,1\}$ vérifiant [H1] et [H2] ou $A_0 \neq 0$ p.s. Alors quand $n \rightarrow \infty$,*

i. $\hat{\alpha} \xrightarrow{p.s.} \alpha$ lorsque $\mathbb{E}(X_0^6) < \infty$,

ii. $n^{1/2}(\hat{\alpha} - \alpha) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, (\mathbb{E}[A_0])^{-1} \text{var}(\epsilon_0^2) \left\{ \mathbb{E}[(\alpha' Z_0)^{-2} Z_0 Z_0']^{-1} \right\} \right)$ si $\mathbb{E}(X_0^8) < \infty$.

La démonstration de ce théorème se base sur les arguments du théorème 1.

Preuve. *D'après 2.20 et 2.21 on a*

$$\hat{\alpha} - \alpha = \left[n^{-1} \sum_{t=p+1}^n \frac{A_t^2 Z_t Z_t'}{\sigma_t^4(\hat{\alpha}_{pr})} \right]^{-1} n^{-1} \sum_{t=p+1}^n \frac{A_t^2 \sigma_t^2(\alpha) Z_t \eta_t}{\sigma_t^4(\hat{\alpha}_{pr})}. \quad (2.28)$$

Pour établir [i.] nous montrons que

$$n^{-1} \sum_{t=p+1}^n \left[\frac{1}{\sigma_t^4(\hat{\alpha}_{pr})} - \frac{1}{\sigma_t^4(\alpha)} \right] A_t^2 Z_t Z_t' \xrightarrow{p.s.} 0, \quad (2.29)$$

$$n^{-1} \sum_{t=p+1}^n \left[\frac{1}{\sigma_t^4(\hat{\alpha}_{pr})} - \frac{1}{\sigma_t^4(\alpha)} \right] A_t^2 \sigma_t^2(\alpha) Z_t \eta_t \xrightarrow{p.s.} 0, \quad (2.30)$$

$$n^{-1} \sum_{t=p+1}^n \frac{A_t^2 Z_t Z_t'}{\sigma_t^4(\alpha)} \xrightarrow{p.s.} \mathbb{E}(A_0^2) \mathbb{E}\{(\alpha' Z_0)^{-2} Z_0 Z_0'\} \quad (2.31)$$

$$n^{-1} \sum_{t=p+1}^n \frac{A_t^2 Z_t \eta_t}{\sigma_t^2(\alpha)} \xrightarrow{p.s} \mathbb{E}(A_0^2) \mathbb{E}\{(\alpha' Z_0)^{-1} Z_0\} \mathbb{E}(\eta_0) = 0 \quad (2.32)$$

Quand $n \rightarrow \infty$, Suivant Bose et Mukherjee (2003), nous utilisons le développement de Taylor pour $u, v > 0$,

$$\frac{1}{u^2} - \frac{1}{v^2} = -2 \frac{(u-v)}{\chi^3} = -2 \frac{(u-v)}{v^3} + 3 \frac{(u-v)^2}{\xi^4} \quad (2.33)$$

où χ et ξ satisfont $0 < \frac{1}{\chi}, \frac{1}{\xi} \leq \frac{1}{v} + \frac{1}{u}$.

Supposons que $\mathbb{E}X_0^4 < \infty$ et $0 < \mathbb{E}A_0^2 < \infty$. d'après le théorème 1 [i.], $\hat{\alpha}_{pr} \xrightarrow{p.s} \alpha$, i.e, il existe un événement $E \in \mathcal{F}$ à la probabilité d'une telle sorte que $\hat{\alpha}_{pr} \rightarrow \alpha$ pour $n \rightarrow \infty$ pour tout les résultats dans E .

Dans ce qui suit. Pour tout $v > 0$, il existe n_0 de telle sorte que $|\hat{\alpha}_{pr,i} - \alpha_i| \leq v, \forall n > n_0$ et $\forall i \in \{0, \dots, q\}$. Prendre $v = \frac{1}{2} \min_{i \in \{0, \dots, q\}} \alpha_i$. Puis $0 < \frac{1}{2} \alpha_i \leq \alpha_i - v \leq \hat{\alpha}_{pr,i} \forall n \geq n_0$ et $\forall i \in \{0, \dots, q\}$. Donc $\forall n \geq n_0$ et $\forall t \in \mathbb{Z}$, $0 < \frac{1}{2} \alpha_0 \leq \frac{1}{2} Z_t' \alpha \leq Z_t' \hat{\alpha}_{pr}$ i.e $0 < \frac{1}{2} \alpha_0 \leq \frac{1}{2} \sigma_t^2(\alpha) \leq \sigma_t^2(\hat{\alpha}_{pr}) \forall n \geq n_0$.

En utilisant 2.33 avec $u = \sigma_t^2(\hat{\alpha}_{pr})$ et $v = \sigma_t^2(\alpha)$, on en déduit que les poits intermédianes correspondants $\chi_{t,n}$ et $\xi_{t,n}$ satisfont $0 < \frac{1}{\chi_{t,n}}, \frac{1}{\xi_{t,n}} \leq \frac{(1+\frac{v}{u})}{v} \leq \frac{3}{\alpha_0} \forall n \geq n_0$ et $t \in \mathbb{Z}$. Soit $\delta_{n,i}$ et $Z_{t,i}$ pour $i = 0, \dots, q$ soit les composantes du $\hat{\alpha}_{pr} - \alpha$ et Z_t , respectivement.

Pour démontrer 2.29, on a pour tout $n \geq n_0$,

$$\begin{aligned} n^{-1} \sum_{t=p+1}^n \left[\frac{1}{\sigma_t^4(\hat{\alpha}_{pr})} - \frac{1}{\sigma_t^4(\alpha)} \right] A_t^2 Z_t Z_t' &= n^{-1} \sum_{t=p+1}^n \left[\frac{1}{u^2} - \frac{1}{v^2} \right] A_t^2 Z_t Z_t' \\ &= n^{-1} \sum_{t=p+1}^n \left[-2 \frac{u-v}{\chi_{t,n}^3} \right] A_t^2 Z_t Z_t' \\ &= n^{-1} \sum_{t=p+1}^n \left[-2 \frac{(\hat{\alpha}_{pr,i} - \alpha_i) Z_{t,i}}{\chi_{t,n}^3} \right] A_t^2 Z_t Z_t' \\ &= \sum_{i=1}^p -2(\hat{\alpha}_{pr,i} - \alpha_i) \sum_{t=p+1}^n \frac{A_t^2 Z_{t,i} Z_t Z_t'}{\chi_{t,n}^3} \\ &= -2 \sum_{i=1}^p \delta_{n,i} n^{-1} \sum_{t=p+1}^n \frac{A_t^2 Z_{t,i} Z_t Z_t'}{\chi_{t,n}^3}, \end{aligned}$$

et on fait entrer (k,l) à l'intérieur de la somme par rapport à "t", alors $\sum_{t=p+1}^n \frac{A_t^2 Z_{t,i} Z_t Z_t'}{\chi_{t,n}^3}$ satisfait

$$0 \leq \frac{A_t^2 Z_{t,i} Z_{t,k} Z_{t,l}}{\chi_{t,n}^3} \leq \left(\frac{3}{\alpha_0} \right)^3 A_t^2 Z_{t,i} Z_{t,k} Z_{t,l}.$$

Le processus $(A_t^2 Z_{t,i} Z_{t,k} Z_{t,l})$ est strictement stationnaire et ergodique (on sait déjà que le processus $\{X_t\}$ est stationnaire et ergodique et **H1** et **H2** sont vérifiées, alors d'après hannan (1973), le processus combiné (X_t, a_t) est strictement stationnaire et ergodique) avec une moyenne finie lorsque $\mathbb{E}X_0^6 < \infty$ et $\mathbb{E}A_0^2 < \infty$.

Ainsi, $n^{-1} \sum_{t=p+1}^n A_t^2 Z_{t,i} Z_{t,k} Z_{t,l}$ converge p.s et $n^{-1} \sum_{t=p+1}^n \frac{A_t^2 Z_{t,i} Z_{t,k} Z_{t,l}}{\chi_{t,n}^3}$ est bornée p.s. puis $\delta_{n,i} \xrightarrow{p.s.} 0$ pour $i = 0, \dots, p$, d'où on déduit 2.29

Pour démontrer 2.30 noter que pour tout $n \geq n_0$ on a

$$n^{-1} \sum_{t=p+1}^n \left[\frac{1}{\sigma_t^4(\hat{\alpha}_{pr})} - \frac{1}{\sigma_t^4(\alpha)} \right] A_t^2 \sigma_t^2(\alpha) Z_t \eta_t = -2 \sum_{i=1}^p \delta_{n,i} n^{-1} \sum_{t=p+1}^n \frac{A_t^2 \sigma_t^2(\alpha) Z_{t,i} Z_t \eta_t}{\chi_{t,n}^3} \quad (2.34)$$

où

$$n^{-1} \left| \sum_{t=p+1}^n \frac{A_t^2 \sigma_t^2(\alpha) Z_{t,i} Z_{t,k} \eta_t}{\chi_{t,n}^3} \right| \leq \left(\frac{3}{\alpha_0} \right)^3 n^{-1} \sum_{t=p+1}^n A_t^2 \sigma_t^2(\alpha) Z_{t,i} Z_{t,k} |\eta_t|, \quad (2.35)$$

et le processus $(A_t^2 \sigma_t^2(\alpha) Z_{t,i} Z_{t,k} |\eta_t|)$ est strictement stationnaire et ergodique avec espérance finie quand $\mathbb{E}X_0^6 < \infty$ et $\mathbb{E}A_0^2 < \infty$. Puis le côté droite de 2.35 converge p.s, et le côté gauche est borné p.s. Ainsi, 2.30 se déduit à partir de 2.34.

et 2.31 et 2.32 sont vérifiant à partir de théorème ergodique pour les séquences stationnaires $(A_t^2 \sigma_t^{-4}(\alpha) Z_t Z_t')$ et $(A_t^2 \sigma_t^{-2}(\alpha) Z_t \eta_t)$ respectivement, quand $\mathbb{E}X_0^4 < \infty$ et $\mathbb{E}A_0^2 < \infty$. D'où **i.** est vérifié.

Pour démontrer **ii.**, nous allons montrons que:

$$n^{-1/2} \sum_{t=p+1}^n \frac{A_t^2 \sigma_t^2(\alpha) Z_t \eta_t}{\sigma_t^4(\hat{\alpha}_{pr})} \xrightarrow{\mathcal{L}} \mathcal{N}[0, \mathbb{E}A_0^4 \text{var}(\epsilon_t^2) \mathbb{E}\{(\alpha' Z_0)^{-2} Z_0 Z_0'\}], \quad (2.36)$$

Où $\mathbb{E}X_0^8 < \infty$ et $0 < \mathbb{E}A_0^4 < \infty$ puis, nous montrons que sous les équations 2.28, 2.29, 2.31, et 2.36 on a

$$n^{1/2}(\hat{\alpha} - \alpha) \xrightarrow{\mathcal{L}} \mathcal{N}\left[0, \frac{\mathbb{E}A_0^4}{\{\mathbb{E}A_0^2\}^2} \text{var}(\epsilon_0^2) \{\mathbb{E}\{(\alpha' Z_0)^{-2} \alpha Z_0 Z_0'\}\}^{-1}\right], \quad (2.37)$$

qui est équivalente à **[ii.]** quand (a_t) est un processus de $\{0,1\}$.

Pour établir 2.36, nous montrons que

$$n^{-1/2} \sum_{t=p+1}^n \left[\frac{1}{\sigma_t^4(\hat{\alpha}_{pr})} - \frac{1}{\sigma_t^4(\alpha)} \right] A_t^2 \sigma_t^2(\alpha) Z_t \eta_t \xrightarrow{p} 0, \quad (2.38)$$

$$n^{-1/2} \sum_{t=p+1}^n \frac{A_t^2 Z_t \eta_t}{\sigma_t^2(\alpha)} \xrightarrow{\mathcal{L}} \mathcal{N}[0, \mathbb{E}A_0^4 \text{var}(\epsilon_t^2) \mathbb{E}\{(\alpha' Z_0)^{-2} Z_0 Z_0'\}]. \quad (2.39)$$

Pour démontrer 2.37 noter que d'après 2.33, $\forall n \geq n_0$, on a

$$\begin{aligned}
 n^{-1/2} \sum_{t=p+1}^n \left[\frac{1}{\sigma_t^4(\hat{\alpha}_{pr})} - \frac{1}{\sigma_t^4(\alpha)} \right] A_t^2 \sigma_t^2(\alpha) Z_t \eta_t &= n^{-1/2} \sum_{t=p+1}^n \sigma_t^2(\alpha) \left[\frac{1}{u^2} - \frac{1}{v^2} \right] A_t^2 Z_t \eta_t \\
 &= n^{-1/2} \sum_{t=p+1}^n \sigma_t^2(\alpha) \left[-2 \frac{(u-v)}{v^3} \right] A_t^2 Z_t \eta_t \\
 &+ n^{-1/2} \sum_{t=p+1}^n \sigma_t^2(\alpha) \left[3 \frac{(u-v)^2}{\xi_{t,n}} \right] A_t^2 Z_t \eta_t \\
 &= -2 \sum_{i=1}^p n^{-1/2} (\hat{\alpha}_{pr,i} - \alpha_i) \sum_{t=p+1}^n \frac{A_t^2 Z_t \eta_t}{\sigma_t^4(\alpha)} \\
 &+ 3 \sum_{i,j=1}^p n^{-1/2} (\hat{\alpha}_{pr,i} - \alpha_i) \sum_{t=p+1}^n \frac{A_t^2 \sigma_t^2(\alpha) Z_{t,i} Z_{t,j} Z_t \eta_t}{\xi_{t,n}^4} \\
 &= -2 \sum_{i=0}^p (n^{1/2} \delta_{n,i}) n^{-1} \sum_{t=p+1}^p \frac{A_t^2 Z_{t,i} Z_t \eta_t}{\sigma_t^4(\alpha)} \\
 &+ 3 \sum_{i,j=0}^p (n^{1/2} \delta_{n,i}) (n^{1/2} \delta_{n,j}) n^{-3/2} \sum_{t=p+1}^n \frac{A_t^2 \sigma_t^2(\alpha) Z_{t,i} Z_{t,j} Z_t \eta_t}{\xi_{t,n}^4} \\
 &= -2T_1 + 3T_2
 \end{aligned}$$

D'après le théorème précédent [ii.], on a

$$n^{1/2} \delta_{n,i} = O_p(1)$$

. Le théorème ergodique pour le séquence stationnaire $\frac{A_t^2 Z_{t,i} Z_t \eta_t}{\sigma_t^4(\alpha)}$ implique que

$$n^{-1} \sum_{t=p+1}^p \frac{A_t^2 Z_{t,i} Z_t \eta_t}{\sigma_t^4(\alpha)} \xrightarrow{p.s} \mathbb{E} A_0^2 \mathbb{E} \{ (\alpha' Z_0)^{-2} Z_{0,i} Z_0 \} \mathbb{E} \eta_0 = 0,$$

et alors $T_1 \xrightarrow{p} 0$. Maintenant, $\forall n \geq n_0$,

$$n^{-1} \left| \sum_{t=p+1}^n \frac{A_t^2 \sigma_t^2(\alpha) Z_{t,i} Z_{t,j} Z_{t,k} \eta_t}{\xi_{t,n}^4} \right| \leq \left(\frac{3}{\alpha_0} \right)^4 n^{-1} \sum_{t=p+1}^n A_t^2 \sigma_t^2(\alpha) Z_{t,i} Z_{t,j} Z_{t,k} | \eta_t |, \quad (2.40)$$

Et le processus $(A_t^2 \sigma_t^2(\alpha) Z_{t,i} Z_{t,j} Z_{t,k} | \eta_t |)$ est strictement stationnaire et ergodique avec une moyenne finie lorsque $\mathbb{E} X_0^8 \leq \infty$ et $\mathbb{E} A_0^2 \leq \infty$. Puis le côté droite de 2.40 converge p.s et le côté gauche est bornée p.s.

Ainsi, $n^{-3/2} \sum_{t=p+1}^n \frac{A_t^2 \sigma_t^2(\alpha) Z_{t,i} Z_{t,j} Z_{t,k} \eta_t}{\xi_{t,n}^4} \xrightarrow{p.s} 0$.

alors $T_2 \xrightarrow{P} 0$. D'où 2.38 est vérifié.

Enfin, pour démontrer 2.39 comme dans la preuve du théorème précédent pour le processus (U_t) , en considérant que le procédé $U_t = (A_t^2 \sigma_t^{-2}(\alpha) Z_t \eta_t)$ est la séquence de différence de martingale, en effet :

Comme X_t est stationnaire et ergodique et **H1** et **H2** sont vérifiées, alors d'après Hannan (1973), le processus combiné (X_t, a_t) est strictement stationnaire et ergodique, il s'en suit (U_t) est strictement stationnaire et ergodique. Lorsque $\mathbb{E}(X_0^4) < \infty$ et $\mathbb{E}(A_0^2) < \infty$, la quantité $\mathbb{E}(U_t)$ est finie, et d'après le théorème ergodique pour les séquences stationnaires (Stout, 1974), il découle

$$\frac{1}{n} \sum_{t=p+1}^n U_t \xrightarrow{p.s} \mathbb{E}[A_0^2] \mathbb{E}[(\alpha' Z_0)^{-2} Z_0] \mathbb{E}[\eta_0] = 0, \quad (2.41)$$

quand $n \rightarrow \infty$.

Soit \mathcal{F}_t la σ -algèbre générée par les variables $(U_j)_{j \leq t}$ et \mathcal{G}_t la σ -algèbre générée par les variables $(X_j)_{j \leq t}$. Nous avons $\mathbb{E}[U_t / \mathcal{G}_{t-1}] = A_t^2 (\alpha' Z_t)^{-2} Z_t \mathbb{E}[\eta_t / \mathcal{G}_{t-1}]$. Il en résulte de **H2** que les variables $\{\epsilon_t, (X_j)_{j \leq t-1}\}$ et (a_t) sont mutuellement indépendantes. Ainsi, ϵ_t est indépendant de $(X_j)_{j \leq t-1}$ et donc indépendant de \mathcal{G}_{t-1} et $\mathbb{E}[\eta_t / \mathcal{G}_{t-1}] = \mathbb{E}[\eta_t] = 0$, ce qui entraîne que $\mathbb{E}[U_t / \mathcal{G}_{t-1}] = 0$. Comme $\mathcal{F}_t^U \subset \mathcal{G}_t$ alors $\mathbb{E}[U_t / \mathcal{F}_{t-1}^U] = \mathbb{E}[\mathbb{E}(U_t / \mathcal{G}_{t-1}) / \mathcal{F}_{t-1}^U] = 0$. En utilisant le théorème central limite de Billingsley (1961) pour les différences de martingales ergodiques et stationnaires, on obtient 2.39.

Lorsque $\mathbb{E}[X_0^8] < \infty$, les éléments de $\mathbb{E}[(\alpha' Z_0)^2 Z_0 Z_0']$ sont finis. De plus pour tout $m > 0$, si $\mathbb{E}[X_0^m] < \infty$, alors $\mathbb{E}[\epsilon_0^m] < \infty$. Par conséquent, quand $\mathbb{E}[|X_0^8|] < \infty$ et $\mathbb{E}[A_0^4] < \infty$, les éléments $\mathbb{E}A_0^4$, $\mathbb{E}\{(\alpha' Z_0)^{-2} Z_0 Z_0'\}$ sont finis, en utilisant 2.38 et 2.39, on obtient **[ii.]**.

Chapitre 3

Illustration numérique

À base du langage de programmation R, nous allons effectuer des simulations intensive dont le but d'illustrer numériquement les résultats du chapitre précédent. Pour rappel, le langage de programmation R est un logiciel de calcul scientifique interactif libre qui possède une large collection d'outils statistiques et graphiques. Plusieurs sites sont consacrés à ce logiciel, en particulier le site <http://www.r-project.org/> offre une description exhaustive sur le langage R et fournit les liens indispensables pour les différents téléchargements, accéder aux différentes bibliothèques. Des versions compilées de R sont disponibles pour Linux, Windows et Mac OS X.

Dans ce chapitre, il sera question de reprendre la procédure numérique faite par P.bondon en 2012. Elle consiste à étudier les propriétés de l'estimateur d'intérêt $\hat{\alpha}$. Pour ce faire, nous considérons une suite de v.a de loi de Bernoulli de paramètre q , avec $q = p\{a_t = 1\}$ et $1 - q$ représente la proportion de données manquantes. Différents tailles de l'échantillon n et de probabilité $(1 - q)$ seront considérées.

Touts les expériences sont basées sur $n.100$ répliques d'un processus ARCH(1) donnée par l'équation

$$\begin{cases} X_t &= \sigma_t \epsilon_t \\ \sigma_t^2 &= 0.3 + 0.5X_{t-1}^2 \end{cases}$$

où (ϵ_t) est une suite de v.a (*i.i.d* de lois normales centrées et réduite. La moyenne empirique et l'erreur standard (e.s) sont les critères retenus pour effectuer une étude comparative de $\hat{\alpha}$ avec celles de $\hat{\alpha}_{YW}$ et d'autres estimateurs basés sur un ensemble de données complète obtenues après le remplissage des observations manquantes par des procédures d'imputations.

en considerant les estimateurs de Bose et Mukherejee (2003), $\hat{\alpha}_b$, $\hat{\alpha}_m$, $\hat{\alpha}_p$ et $\hat{\alpha}_p^n$, lorsque la

donnée manquante est remplacée respectivement par: par:

- 1- le maximum des données observées.
- 2- la moyenne empirique des données observées,
- 3- la première donnée observée qui la précède
- 4- l'interpolation entre la donnée qui la précède et celle qui la succède.

Avec les estimateurs de maximum de vraisemblance $\tilde{\alpha}_b$, $\tilde{\alpha}_m$, $\tilde{\alpha}_p$ et $\tilde{\alpha}_p^n$ lorsque la donnée manquante est remplacée respectivement par:

- 1- le maximum des données observées.
- 2- la moyenne empirique des données manquante .
- 3- la première donnée donnée observée qui la précède.
- 4- l'interpolation entre la donnée qui la précède et celle qui la succède.

Ces simulations sont à base du package **fgarch**, qui donne directement les valeurs simulées d'un ARCH(1), et ce en faisant varier le n . ($n = 500$, $n = 1000$, et $n = 2000$).

A chaque itération, nous avons écraser une proportion $(1 - q)$ données. Les diverses valeurs de $1 - q$ sont = 5%, 10%, 15%, 20%.

A l'étape suivante nous remplaçons les données manquantes par les méthodes d'imputation précédentes. Pour chaque itération, nous calculons la valeur de l'estimateur d'intérêt $\hat{\alpha}$ et celles de $\hat{\alpha}_b$, $\hat{\alpha}_m$, $\hat{\alpha}_p$, $\hat{\alpha}_p^n$, et $\tilde{\alpha}_b, \tilde{\alpha}_m, \tilde{\alpha}_p, \tilde{\alpha}_p^n$. A la dernière étape de chaque itération nous calculons l'erreur et l'écart type des estimateurs.

Programme de simulation

Le programme R qui permet d'exécuter toutes ces tâches citées dans la procédure algorithmique précédentes est :

```
p=0.15 # simulation de données sous forme d'un vecteur z de
dimension n =200. simulator <- fonction(n = 200, omega = 10^(-6),
ar = 0.4, alpha = 0.3){
  spec = garchSpec(model = list(ar = ar, alpha, beta = 0))
  z <- garchSim(spec, n = n)$garch
  z
}

simARCH1 <- fonction(n = 200, ar = 0.3){
```

```
spec = garchSpec(model = list(ar = ar, beta = 0))
z <- garchSim(spec, n = n)$garch
}

#Construction de l'estimateur phiestim <- function(X){
  n <- length(X)
  z <- sum(X[2:n]*X[1:n-1])/sum(X[1:n-1]*X[1:n-1])
  z
}
# theta=(phi, omega,alpha) # phi sera fixé à 0.50

# Désigner la proportion de données manquante et remplacer es
données par des 0 missdata0 <- function(X, q = 0.05){
  n <- length(X)
  w <- runif(n)
  for (i in 1:n){
    if (w[i] < q) X[i] = 0
  }
  X
}

# Données manquantes remplacées par missdata1 <- function(X, q =
0.05, arr) {
  n <- length(X)
  w <- runif(n)
  for (i in 2:(n-1)){
    if (w[i] < q) X[i] <- (X[i-1]+X[i+1])*(arr/(1+(arr^2)))
  }
  X
}

# Calcul de l'erreur quadratique moyenne errq <- function(X, rval
= 0) {
  n <- length(n)
```

```
S <- 0
for (i in 1:n) {
  S <- S + ((X[i] - rval)^2)
}
z <- (1/n)*S
z
}

# Dessin du tableau TABLE <- matrix(vector(mode = "numeric",
length = 10*(5+2)), nrow=5+2) TABLE[1,] = c("try", "q = 0", "",
"", "m.data.0", "", "", "m.data.est.", "", "") TABLE[2,] = c("n",
"moy", "var", "er.q.m", "moy", "var", "er.q.m", "moy", "var",
"er.q.m")

#

# Nombre d'expériences nbrExp <- 5 # 5

for (j in 3:3) {
  n = 125*(2^j) # 250, 500, 1000, 2000, 4000
  TABLE[j+2,1] = n
  rep = 10*n
  ar = 0.5
  omega = 0.2
  alpha = 0.3
  v = vector(mode = "numeric", length = n)
  v_m0 = vector(mode = "numeric", length = n)
  v_m1 = vector(mode = "numeric", length = n)
  est_mc_0 <- vector(mode = "numeric", length = rep)
  est_mc_m0 <- vector(mode = "numeric", length = rep)
  est_mc_m1 <- vector(mode = "numeric", length = rep)
  for (i in 1:rep) {
    # ARCH-GARCH
    # v <- simulator(n=n,omega=omega,ar=ar,alpha=alpha)
```

```
#
v <- simARCH1(n = n, ar = ar)
est_mc_0[i] <- phiestim(v)

v_m0 <- missdata0(v, q = p) # q = ?
est_mc_m0[i] <- phiestim(v_m0)

v_m1 <- missdata1(v, q = p, arr = mean(est_mc_0))
est_mc_m1[i] <- phiestim(v_m1)
}
# remplissage du tableau
TABLE[j+2,2] <- mean(est_mc_0)
TABLE[j+2,3] <- var(est_mc_0)
TABLE[j+2,4] <- errq(est_mc_0, rval = ar)

TABLE[j+2,5] <- mean(est_mc_m0)
TABLE[j+2,6] <- var(est_mc_m0)
TABLE[j+2,7] <- errq(est_mc_m0, rval = ar)

TABLE[j+2,8] <- mean(est_mc_m1)
TABLE[j+2,9] <- var(est_mc_m1)
TABLE[j+2,10] <- errq(est_mc_m1, rval = ar)
}

# simulation de données sous forme d'un vecteur de dimension n
simulator <- function(n = 200, omega = 0.3, ar = 0.5, alpha =
0.6){
  spec = garchSpec(model = list(ar = ar, alpha, beta = 0))
  z <- garchSim(spec, n = n)$garch
  z
} #simARCH1 <- function(n = 200, ar = 0.3){
```

```
# spec = garchSpec(model = list(ar = ar, beta = 0))
# z <- garchSim(spec, n = n)$garch
#}

#Construction de l'estimateur phiestim <- function(X){
  n <- length(X)
  z <- sum(X[2:n]*X[1:n-1])/sum(X[1:n-1]*X[1:n-1])
  z
}
# theta=(phi, omega,alpha) # phi sera fixé à 0.50

# Désigner la proportion de données manquante et remplacer les
données par des 0 missdata0 <- function(X, q = 0.01){
  n <- length(X)
  w <- runif(n)
  for (i in 1:n){
    if (w[i] < q) X[i] = 0
  }
  X
}

# Données manquantes remplacées par une valeur estimée missdata1
<- function(X, q = 0.01, ar) {
  n <- length(X)
  w <- runif(n)
  for (i in 2:(n-1)){
    if (w[i] < q) X[i] <- (X[i-1]+X[i+1])*(ar/(1+(ar^2)))
  }
  X
}

# Calcul de l'erreur quadratique moyenne errq <- function(X, rval
= 0) {
  n <- length(n)
```

```
S <- 0
for (i in 1:n) {
  S <- S + ((X[i] - rval)^2)
}
z <- (1/n)*S
z
}

# faire varier le n ici

nbrExp <- 3 # Nombre d'expériences
  # n = 125*(2^k), k in 1:nbrExp
repMult <- 1000 # multiplicateur de répétitions
  # rep = repMult * n
Q <- c(0.05,0.10,0.15,0.20) # nombre de changements de q

# Préparation du graphique par(mfrow=c(3,length(Q)*nbrExp))

TABLE <- matrix(vector(mode = "numeric", length =
10*((length(Q)+2)*nbrExp)), ncol=10)

for (k in 1:nbrExp) {
  n = 15*(1+k) # 250, 500, 1000, 2000, 4000

  # Dessin du tableau

  TABLE[(((length(Q)+2)*(k-1))+1),] = c("try", "q = 0", "", "", "m.data.0", "", "", "m.d
  TABLE[(((length(Q)+2)*(k-1))+2),] = c("n", "moy", "et", "er.q.m", "moy", "et", "er.q.m
  TABLE[(((length(Q)+2)*(k-1))+2),1] = n

  for (j in 1:length(Q)) {
    TABLE[(((length(Q)+2)*(k-1))+2+j),1] = Q[j]
    rep = repMult*n
    ar = 0.5
```

```

omega = 0.3
alpha = 0.6
v = vector(mode = "numeric", length = n)
v_m0 = vector(mode = "numeric", length = n)
v_m1 = vector(mode = "numeric", length = n)
est_mc_0 <- vector(mode = "numeric", length = rep)
est_mc_m0 <- vector(mode = "numeric", length = rep)
est_mc_m1 <- vector(mode = "numeric", length = rep)
for (i in 1:rep) {
  # ARCH-GARCH
  # v <- simulator(n=n,omega=omega,ar=ar,alpha=alpha)
  #
  v <- simulator(n = n, omega = omega, ar = ar, alpha = alpha)
  est_mc_0[i] <- phiestim(v)

  v_m0 <- missdata0(v, q = Q[j]) # q = ?
  est_mc_m0[i] <- phiestim(v_m0)

  v_m1 <- missdata1(v, q = Q[j], ar = mean(est_mc_m1))
  est_mc_m1[i] <- phiestim(v_m1)
}
# remplissage du tableau
TABLE[(((length(Q)+2)*(k-1))+2+j),2] <- mean(est_mc_0)
TABLE[(((length(Q)+2)*(k-1))+2+j),3] <- var(est_mc_0)
TABLE[(((length(Q)+2)*(k-1))+2+j),4] <- errq(est_mc_0, rval = ar)
# hist(est_mc_0, breaks=256)

TABLE[(((length(Q)+2)*(k-1))+2+j),5] <- mean(est_mc_m0)
TABLE[(((length(Q)+2)*(k-1))+2+j),6] <- sqrt(var(est_mc_m0))
TABLE[(((length(Q)+2)*(k-1))+2+j),7] <- errq(est_mc_m0, rval = ar)
# hist(est_mc_m0, breaks=256)

TABLE[(((length(Q)+2)*(k-1))+2+j),8] <- mean(est_mc_m1)
TABLE[(((length(Q)+2)*(k-1))+2+j),9] <- sqrt(var(est_mc_m1))

```

```

TABLE[(((length(Q)+2)*(k-1))+2+j),10] <- errq(est_mc_m1, rval = ar)
# hist(est_mc_m1, breaks=256)
} }
# Affichage du tableau TABLE

```

Après exécution, les résultats obtenus sont résumés par les tableaux ci dessous

Tableau 1. pour $n = 500$

$1 - q$	0%	5%	10%	15%	20%
$\hat{\alpha}$	0.305 (3.161e - 2)	0.305 (3.895e - 2)	0.305 (4.422e - 2)	0.306 (8.414e - 2)	0.307 (5.794e - 2)
$\hat{\alpha}_{YW}$	0.483 (9.618e - 2)	0.483 (4.728e - 1)	0.480 (3.792e - 1)	0.477 (5.255e - 1)	0.475 (5.984e - 1)
$\hat{\alpha}_b$	0.356 (6.816e - 2)	0.357 (7.115e - 2)	0.358 (7.706e - 2)	0.359 (8.437e - 2)	0.359 (9.480e - 2)
$\tilde{\alpha}_b$	0.394 (1.121e - 1)	0.392 (1.202e - 1)	0.390 (1.290e - 1)	0.388 (1.378e - 1)	0.386 (1.478e - 1)
$\hat{\alpha}_m$	0.305 (3.161e - 2)	0.312 (3.400e - 2)	0.321(3.659e - 2)	0.330 (3.997e - 2)	0.340 (4.272e - 2)
$\tilde{\alpha}_m$	0.483 (9.618e - 2)	0.470 (1.375e - 1)	0.455 (1.089e - 1)	0.439 (1.645e - 1)	0.423 (1.145e - 1)
$\hat{\alpha}_p$	0.302 (3.011e - 2)	0.310 (3.223e - 2)	0.318 (3.455e - 2)	0.327 (3.716e - 2)	0.336 (3.999e - 2)
$\tilde{\alpha}_p$	0.491 (8.944e - 1)	0.478 (9.317e - 2)	0.464 (9.712e - 2)	0.449 (1.017e - 1)	0.434 (1.064e - 1)
$\hat{\alpha}_p^n$	0.305 (3.161e - 2)	0.318 (4.133e - 2)	0.325 (4.760e - 2)	0.328 (5.199e - 2)	0.325 (5.297e - 2)
$\tilde{\alpha}_p^n$	0.483 (9.618e - 2)	0.423 (1.054e - 1)	0.372 (7.926e - 1)	0.329 (1.184e - 1)	0.289 (4.015e - 1)
	0.302 (3.011e - 2)	0.317 (3.976e - 2)	0.324 (4.532e - 2)	0.326 (4.897e - 2)	0.323 (5.026e - 2)
	0.491 (8.944e - 2)	0.427 (9.019e - 2)	0.377 (8.939e - 2)	0.335 (8.854e - 2)	0.299 (8.772e - 2)
	0.305 (3.161e - 2)	0.296 (3.191e - 2)	0.288 (3.324e - 2)	0.279 (3.273e - 2)	0.270 (3.326e - 2)
	0.483(9.618e - 2)	0.498 (1.018e - 1)	0.512 (1.315e - 1)	0.527 (9.721e - 2)	0.543(9.686e - 2)
	0.302 (3.011e - 2)	0.294 (3.061e - 2)	0.0.286 (3.113e - 2)	0.277 (3.170e - 2)	0.268 (3.230e - 2)
	0.491 (8.944e - 2)	0.502 (8.851e - 2)	0.519 (8.766e - 2)	0.534 (8.715e - 2)	0.550 (8.676e - 2)
	0.305 (3.161e - 2)	0.294 (3.180e - 2)	0.282 (3.226e - 2)	0.270 (3.222e - 2)	0.258 (3.216e - 2)
	0.483 (9.618e - 2)	0.492 (1.199e - 1)	0.501 (9.841e - 2)	0.512 (9.741e - 2)	0.524 (1.396e - 1)
	0.302 (3.011e - 2)	0.291 (3.058e - 2)	0.0.279 (3.110e - 2)	0.267 (3.113e - 2)	0.254 (3.123e - 2)
	0.491 (8.944e - 2)	0.500 (9.054e - 2)	0.512 (9.179e - 2)	0.525 (9.309e - 2)	0.540 (9.447e - 2)

Tableau 2.pour $n = 1000$

$1 - q$	0%	5%	10%	15%	20%
$\hat{\alpha}$	0.302 (2.192e - 2)	0.302 (2.925e - 2)	0.303 (3.670e - 2)	0.303 (4.445e - 2)	0.307 (5.144e - 2)
	0.492 (6.630e - 2)	0.492 (4.897e - 1)	0.491 (7.368e - 1)	0.489 (4.640e - 1)	0.487 (7.792e - 1)
$\hat{\alpha}_{YW}$	0.346 (5.433e - 2)	0.347 (5.976e - 2)	0.348 (6.489e - 2)	0.349 (7.286e - 2)	0.349 (7.747e - 2)
	0.415 (9.763e - 2)	0.413 (1.053e - 1)	0.412 (1.136e - 1)	0.410 (1.223e - 1)	0.408 (1.316e - 1)
$\hat{\alpha}_b$	0.302 (2.192e - 2)	0.310 (2.382e - 2)	0.318 (2.546e - 2)	0.327 (2.728e - 2)	0.336 (2.991e - 2)
	0.492 (6.630e - 2)	0.478 (7.156e - 2)	0.464 (7.489e - 2)	0.449 (7.880e - 2)	0.432 (1.003e - 1)
$\tilde{\alpha}_b$	0.301 (2.114e - 2)	0.309 (2.264e - 2)	0.317 (2.429e - 2)	0.326 (2.616e - 2)	0.335 (2.816e - 2)
	0.495 (6.283e - 2)	0.482 (6.567e - 2)	0.469 (6.852e - 2)	0.454 (7.167e - 2)	0.439 (7.503e - 2)
$\hat{\alpha}_m$	0.302 (2.192e - 2)	0.317 (2.999e - 2)	0.324 (3.384e - 2)	0.327 (3.929e - 2)	0.325 (3.833e - 2)
	0.492 (6.630e - 2)	0.430 (7.087e - 2)	0.377 (8.374e - 2)	0.332 (7.733e - 2)	0.294 (7.618e - 2)
$\tilde{\alpha}_m$	0.301 (2.114e - 2)	0.317 (2.906e - 2)	0.324 (3.233e - 2)	0.326 (3.818e - 2)	0.323 (3.637e - 2)
	0.495 (6.283e - 2)	0.430 (6.391e - 2)	0.380 (6.339e - 2)	0.338 (6.279e - 2)	0.302 (6.194e - 2)
$\hat{\alpha}_p$	0.302 (2.192e - 2)	0.505 (7.344e - 2)	0.519 (6.590e - 2)	0.534 (6.893e - 2)	0.550 (7.774e - 2)
	0.492 (6.630e - 2)	0.293 (2.151e - 2)	0.285 (2.191e - 2)	0.276 (2.239e - 2)	0.267 (2.284e - 2)
$\tilde{\alpha}_p$	0.301 (2.144e - 2)	0.293 (2.151e - 2)	0.285 (2.191e - 2)	0.276 (2.239e - 2)	0.267 (2.284e - 2)
	0.495 (6.283e - 2)	0.509 (6.228e - 2)	0.523 (6.194e - 2)	0.537 (6.153e - 2)	0.553 (6.118e - 2)
$\hat{\alpha}_p^n$	0.302 (2.192e - 2)	0.291 (2.225e - 2)	0.280 (2.250e - 2)	0.268 (2.241e - 2)	0.256 (2.274e - 2)
	0.492 (6.630e - 2)	0.500 (9.561e - 2)	0.509 (6.714e - 2)	0.520 (7.911e - 2)	0.533 (7.335e - 2)
$\tilde{\alpha}_p^n$	0.301 (2.114e - 2)	0.290 (2.150e - 2)	0.278 (2.174e - 2)	0.266 (2.193e - 2)	0.253 (2.204e - 2)
	0.495 (6.283e - 2)	0.505 (6.381e - 2)	0.516 (6.469e - 2)	0.529 (6.562e - 2)	0.544 (6.651e - 2)

Lemme 3.1. *Supposons que $\{X_t\}$ un processus purment non déterminable de représentation $MA(\infty)$ donné par*

$$X_t = \sum_{i=0}^{\infty} c_i \epsilon_{t-i}$$

où $\epsilon_t = X_t - P_{t-1}X_t$, $c_0 = 1$, et $\sum_{i=0}^{\infty} c_i^2 < \infty$. La représentation autorégressive $AR(\infty)$ de $\{X_t\}$

$$X_t = \sum_{i=1}^{\infty} a_i X_{t-i}$$

Alors les paramètres de la représentation autorégressive et la représentation moyenne mobile sont liés par les relations.

$$a_0 = 1$$

$$a_i = \sum_{j=0}^{\infty} a_j c_{i-j}$$

Interprétation

À partir des tableaux ci dessus, nous constatons que $\hat{\alpha}, \hat{\alpha}_{YW}, \hat{\alpha}_b, \tilde{\alpha}_b, \hat{\alpha}_m$ et $\tilde{\alpha}_m$.

sur-estiment α_0 et sous-estiment α_1 . Aussi, $\hat{\alpha}_p, \tilde{\alpha}_p, \hat{\alpha}_p^n$ et $\tilde{\alpha}_p^n$ sous-estiment α_0 et sur-estiment α_1 .

Pour tous les estimateurs l'écart type de α_1 est plus grand que celui de α_0 .

Sans aucune surprise, $\hat{\alpha}_{YW}$ n'est pas un bon estimateur de α même dans le cas d'absence de données manquantes.

Pour toutes les méthodes d'imputations, il est claire que le biais de l'estimateur augmente en fonction de la proportion de données manquante ($1 - q$), mais l'écart type reste sensiblement le même.

Toutefois, il est à signaler que le biais de $\hat{\alpha}$ varie longtement lorsque la proportion de données manquantes augment. En d'autres termes, ces estimateur résiste "plus" au manque de données.

Pour l'estimateur du maximum de vraisemblance, en absence de données manquantes le bien fondé de la théorie est confirmé.

La supériorité des estimateurs de P.Bondon (2012) est confirmée par ces trois tableaux. En effet, pour toutes les méthodes d'estimations par imputation, la moyenne empirique augmente fortement quand le $1 - q$ augmente, et lorsque n augmente, la qualité de l'estimateur $\hat{\alpha}$ est nettement meilleure, ceci est constaté indépendamment de de la proportion de données manquantes.

Ceci qui confirme qu'il est meilleurs par rapport aux autres.

Conclusion générale

Dans de nombreuses situations pratiques les données ne sont pas observées. En effet, dans les applications, on est très souvent en présence d'observations pour lesquelles on ne dispose de l'ensemble des valeurs des variables descriptives, et ceci se produit pour des nombreuses raisons. Nous citons : valeurs non enregistrées, valeurs aberrantes, erreur de saisies, données recueillis difficilement, etc. Ce mémoire est une rétrospective de traitement de ce problème de données manquantes. Après une présentation de la technique d'interpolation de données manquantes, nous avons étudié un problème très particulier, qui consiste à estimer les paramètres du modèle ARCH et à étudier leurs propriétés. Aussi, nous nous sommes basés sur des travaux de simulation pour quantifier la qualité de cette procédure, initiée par P.Bondon 2012. Ce travail est loin d'être achevé et ouvre des perspectives futures. Nous citons le cas de généralisation de cette procédure à d'autres modèles, en particulier, Les GARCH et les ARMA GARCH.

Annexe

Nous allons fournir les outils, probabilistes, employés dans la théorie asymptotique

Définition 3.1 (Martingale). Soit $(X_t)_{t \in \mathbb{N}}$ une suite de variables aléatoires réelles (v.a.r) sur un espace probabilisé (Ω, \mathcal{F}, P) , et $(\mathcal{F}_t)_{t \in \mathbb{N}}$ est une suite de tribus.

La suite $\{(X_t, \mathcal{F}_t), t = 1, 2, \dots\}$ est une martingale si et seulement si:

1. $\mathcal{F}_t \subset \mathcal{F}_{t+1}$.
2. X_t est \mathcal{F}_t -mesurable.
3. $E(|X_t|) < \infty$.
4. $E(X_{t+1}/\mathcal{F}_t) = X_t$.

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

Définition 3.2 (Différence de martingale). Soient $\{X_t\}_{t \in \mathbb{N}}$ une suite de variables aléatoires réelles (v.a.r), et $(\mathcal{F}_t)_{t \in \mathbb{N}}$ est une suite de tribus. La suite $\{(X_t, \mathcal{F}_t), t = 1, 2, \dots\}$ est une différence de martingale (ou une suite d'accroissements de martingale) si et seulement si :

1. $\mathcal{F}_t \subset \mathcal{F}_{t+1}$.
2. X_t et \mathcal{F}_t -mesurable.
3. $E(|X_t|) < \infty$.
4. $E(X_{t+1}/\mathcal{F}_t) = 0$.

Théorème 3.1 (Théorème de Cramer-Wold). Pour une suite (Z_n) de vecteurs aléatoire d , $Z_n \xrightarrow{L} Z$ si et seulement si pour tout $\lambda \in \mathbb{R}^d$, on a

$$\lambda' Z_n \xrightarrow{\mathcal{L}} \lambda' Z$$

Théorème central limite (T.C.L) pour différence de martingale stationnaire
[Théorème de Billingsley(1961)]

Théorème 3.2. Si (ν_t, \mathcal{F}_t) est une différence (ν_t est \mathcal{F}_t -mesurable et $E(\nu_t/\mathcal{F}_t) = 0$, sta-

tionnaire ergodique, de carré intégrable, telle que $V(\nu_t) = \sigma_t^2 \neq 0$, alors

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

[(T.C.L) de Lindeberg]

Théorème 3.3. On suppose que pour chaque $n > 0$, $(\eta_{nk}^2, \mathcal{F}_{nk})_{k \in \mathbb{N}}$ est une différence de martingale de carré intégrable. Soit $\sigma_{nk}^2 = E(\eta_{nk}^2 / \mathcal{F}_{n(k-1)})$. Si

$$\sum_{k=1}^n \sigma_{nk}^2 \xrightarrow{p} \sigma_0^2 \quad \text{quand } n \rightarrow \infty$$

Où σ_0^2 est une constante strictement positive, et

$$\sum_{k=1}^n E(\sigma_{nk}^2 \mathbb{I}_{\{|\eta_{nk}| > \epsilon\}}) \rightarrow 0 \quad \text{quand } n \rightarrow \infty$$

Pour chaque réel positif ϵ , alors

$$\sum_{k=1}^n \eta_{nk} \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma_0^2)$$

Définition 3.3 (Ergodicité). On dit qu'une suite stationnaire est ergodique si elle satisfait la loi forte des grands nombres. Certaines transformations de suites ergodiques restent ergodiques.

Théorème 3.4. Si $(Z_t)_{t \in \mathbb{Z}}$, est une suite fortement stationnaire et ergodique. et si $f : \mathbb{R}^\infty \rightarrow \mathbb{R}$ une fonction mesurable, et soit $Y_t = f(\dots, Z_{t-1}, Z_t, Z_{t+1}, \dots)$, alors $(Y_t)_{t \in \mathbb{Z}}$ reste une suite fortement stationnaire et ergodique.

Théorème 3.5. Si $(Z_t)_{t \in \mathbb{Z}}$ est strictement stationnaire et ergodique, si f est mesurable et si $E\{|f(\dots, Z_{t-1}, Z_t, Z_{t+1}, \dots)|\} < \infty$, alors: $\frac{1}{n} \sum_{t=1}^n f(\dots, Z_{t-1}, Z_t, Z_{t+1}, \dots) \rightarrow E\{f(\dots, Z_{t-1}, Z_t, Z_{t+1}, \dots)\}$ p.s.

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

Définition 3.4 (Convergence en probabilité vers zero). On dit que $\{X_n\}$ converge vers zero en probabilité, on écrit $X_n = o_p(1)$ ou $X_n \rightarrow 0$, si pour tout $\epsilon > 0$

$$P(|X_n| > \epsilon) \rightarrow 0, \quad \text{quand } n \rightarrow \infty.$$

Définition 3.5 (Borgnitude en probabilité). On dit que (X_n) est borné en probabilité, on écrit $X_n = O_p(1)$, si pour tout $\epsilon > 0$, il existe $\delta(\epsilon) \in (0, \infty)$;

$$P(|X_n| > \delta(\epsilon)) < \epsilon, \quad \forall n$$

Définition 3.6 (Convergence en probabilité vers une variable aléatoire). 1. X_n converge en probabilité vers une variable aléatoire X , on écrit $X_n \longrightarrow X$, si et seulement si $X_n - X = o_p(1)$.

2. $X_n = o_p(a_n)$ si et seulement si $a_n^{-1}X_n = o_p(1)$.

3. $X_n = O_p(a_n)$ si et seulement si $a_n^{-1}X_n = O_p(1)$.

Propriété 3.1. Si X_n et Y_n , $n = 1, 2, \dots$, des variables aléatoires dans un même espace probabilité et $a_n > 0$, $b_n > 0$, $n = 1, 2, \dots$, alors.

1. Si $X_n = o_p(a_n)$ et $Y_n = o_p(b_n)$, on aura

- $X_n Y_n = o_p(a_n b_n)$.
- $X_n + Y_n = o_p(\max(a_n, b_n))$.
- $|X_n|^r = o_p(a_n^r)$ pour $r > 0$.

2. Si $X_n = O_p(a_n)$ et $Y_n = O_p(b_n)$ on aura:

$$X_n Y_n = O_p(a_n b_n).$$

Remarque 3.1. La première propriété reste vrai même si on remplace o_p par O_p .

Définition 3.7 (convergence en probabilité d'un vecteur aléatoire). 1. X_n converge en probabilité vers le vecteur aléatoire X , et on écrit, $X_n \xrightarrow{Pr} X$ si et seulement si

$$X_n - X = o_p(1).$$

2. $X_n = o_p(a_n)$ si et seulement si $X_{nj} = o_p(1)$, $j = 1, \dots, k$.

3. $X_n = O_p(a_n)$ si et seulement si $X_{nj} = O_p(1)$, $j = 1, \dots, k$.

La convergence en probabilité de $\{X_n\}$ vers X peut aussi être caractérisée en fonction de la distance euclidienne $|X_n - X| = [\sum_{j=1}^k (X_{nj} - X_j)^2]^{1/2}$.

Développement de Taylor en probabilité

Si g est continue en a et $X_n = a + o_p(1)$, alors la dernière proposition nous permet d'écrire $g(X_n) = g(a) + o_p(1)$.

Si nous renforçons les hypothèses sur g en mentionnant l'existence de produits dérivés, il est alors possible de calculer de manière probabiliste des développements de Taylor de

fonctions non aléatoires au voisinage d'un point donné a .

Proposition 3.1. Soit $\{X_n\}$ une suite de variables aléatoires telles que $X_n = a + O_p(r_n)$, où $a \in \mathbb{R}$ et $0 < r_n \rightarrow \infty$. Si g est une fonction n fois dérivée au point a alors

$$g(X_n) = \sum_{j=0}^n \frac{g^{(j)}(a)}{j!} (X_n - a)^j + o_p(r_n^s),$$

où $g^{(j)}$ est la j^{eme} dérivée de g , et $g^{(0)} = g$.

Convergence en loi, et presque sur

Définition 3.8 (convergence en loi). On dit que $\{X_n\}$ converge en vers la variable aléatoire X si la fonction de répartition $F_n(t)$ de X_n converge vers $F(t)$, celle de X en tout point t où F est continue (c'est à dire en tout points t tels que $P(X = t) = 0$). On note la convergence en loi :

$$X_n \xrightarrow{\mathcal{L}} X$$

Remarque 3.2. Si X est une variable à densité (de fonction de répartition continue), il n'y a plus d'hypothèse restrictive à faire sur les points de continuité.

Proposition 3.2. Si $X_n \xrightarrow{\mathcal{L}} X$ et $Y_n \xrightarrow{\mathcal{L}} 0$ alors $X_n + Y_n \xrightarrow{\mathcal{L}} X$.

Proposition 3.3. $X_n \xrightarrow{\mathcal{L}} X$ si et seulement si pour toute fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ continue bornée

$$\lim_{n \rightarrow \infty} \mathbf{E}[f(X_n)] = \mathbf{E}[f(X)].$$

Définition 3.9 (convergence presque sure). Une suite $\{X_n\}$ de variables aléatoires converge presque sûrement vers X si

$$P(\omega \in \Omega / \lim_{n \rightarrow \infty} X_n(\omega) = x(\omega)) = 1.$$

Et On note $X_n \xrightarrow{P.S} X$.

Proposition 3.4. 1. Convergence presque sure \Rightarrow convergence en probabilité \Rightarrow convergence en loi

2. Soit C une constante alors : $X_n \xrightarrow{\mathcal{L}} C \Rightarrow X_n \xrightarrow{P} C$.

Définition 3.10 (Le processus bruit blanc :). On dit que $\{X_t\}_{t \in \mathbb{Z}}$ est un bruit blanc faible de moyenne nulle et de variance σ^2 noté $X_t \sim BB(0, \sigma^2)$ lorsque $Cov(X_t, X_i) = 0$, $\forall t \neq i$

On dit que $\{X_t\}$ est un bruit blanc fort de moyenne nulle est de variance σ^2 lorsque: $X_t \perp X_i$, $\forall t \neq i$.

Bibliographie

- [1] Brockwell, P.J. and Davis R.A. (1991). Time Series: Theory and Methods (2nd eds), New York: Springer Verlag.
- [2] Bollerslev, T. (1986), Generalized autoregressive conditional heteroscedasticity, Journal of Econometrics. 31 (3), 307-327.
- [3] bondon, P and Bahamonde, N. (2012) Least squares estimation of ARCH models with missing observations.
- [4] Bose, A. and Mukherjee, K. (2003) Estimating the ARCH parameters by solving linear equations. Journal of Time Series Analysis 24, 127–136.
- [5] Engle, R.F. 1982, Autoregressive conditional heteroscedasticity with estimates of the variance of Unit Kingdom inflation, Econometrica, 50, 987-1007.
- [6] Jensen, S.T. and Rahbek, A. (2004a) Asymptotic normality of the QMLE estimator of ARCH in the nonstationary case. Econometrica 72, 641–646.
- [7] Parzen, E. (1983). Proc. of time series analysis of irregularly observed data, Lecture Notes in Statistics, 25, Springer, New York.
- [8] Pourahmadi, M. (1989). Estimation and interpolation of missing values of stationary time series, Journal of Time Series Analysis, 10 (2), 149-169.