الجمهورية الجزائرية الديمقراطية الشعبية
وزارة التعليم العالي و البحث العلمي

Université Abou Bekr Belkaid — جامعة أبي بكر بلقايد
Tlemcen Algérie — تلمسان الجزائر

MÉMOIRE

Pour l'obtention du grade de

Master

Faculté des Sciences Tlemcen

Département de Mathématiques

Probabilité et Statistique

Présenté par

CHAOUCHE Meryem

Sur un modèle autorégressif à coefficients aléatoires d'ordre un RCA(1)

Soutenu le 21 Septembre 2015
Devant la Commission d'Examen

JURY

M. T Mourid	Professeur à l'université Abou Bekr Belkaid	Président
M. A Labbas	Maitre de conférence à l'université Abou Bekr Belkaid	Examinateur
M. A. ALLAM	Maitre de conférence à l'université Abou Bekr Belkaid	Encadreur

Remerciement

A notre maître et président du jury Professeur Mourid

Monsieur le Professeur Mourid.

Vous nous faites l'honneur et le plaisir de présider ce jury.
Qu'il me soit permis de vous exprimer toute ma reconnaissance et mon profond respect car vous avez été pour moi un mentor remarquable qui m'a appris comment construire mon intuition mathématique.

A mon encadreur et Professeur Allam

Monsieur le Professeur Allam.

Je tiens à vous remercier très sincèrement pour votre sympathie, votre disponibilité, et vos conseils intéressants et combien profitables pour l'avancement de ce travail. Tout au long de ma thèse j'ai eu la chance d'apprendre et de profiter de votre expérience et de vos compétences.

Je remercie chaleureusement Monsieur Abbes d'avoir bien voulu me faire l'honneur de faire partie du jury.

Table des matières

Introduction

Les travaux de recherches relatifs à la modélisation des séries chronologiques ont été dévoués surtout aux modèles linéaires, en particulier les modèles autoregressifs et les moyennes mobiles qui ont été étendus jusqu'aux variables exogènes. Par la suite, l'analyse des modèles non linéaires a connu un intérêt croissant dans la littérature. Granger et Anderson (1978) ont étudié les modèles bilinéaires. Les modèles autoregressifs à coefficients aléatoires forment une autre classe de modèles non linéaires qui a été reconsidérée par de nombreux auteurs.

Andel (1976) a mis le point sur le fait que la modélisation statistique dans les domaines comme l'économie, la biologie, l'hydraulique ..., les coefficients du modèle sont les résultats de plusieurs processus compliqués et actions de caractère aléatoire, d'où l'intérêt de considérer des modèles à coefficients aléatoires qui constituent une généralisation des modèles à coefficients constants.

La classe des processus autogressifs à coefficients aléatoires a été introduite par Andel (1976) qui a établi des conditions pour la stationnarité faible pour cette classe dans le cas univarié. Nicholls et Quinn (1980-1982) ont généralisé ces résultats au cas multivarié et ont traité l'estimation de cette classe.

Ce mémoire se base essentiellement sur la thèse de :

WANG, DAZHE (Frequentist and Bayesian analysis of Random Coefficient Autoregressive Models), et vise à développer quelque résultats sur les modèles RCA(1), notamment sur les problèmes d'estimation et des previsions.

Il est composé de trois chapitres. Dans le premier, on présente quelques généralités sur les processus autorégressifs à coefficients aléatoires RCA(1), plus particulièrement l'existence de la solution ergodique et strictement stationnaire.

Quelques exemples de simulation d'un RCA(1) sont aussi présentés.

Dans le deuxième chapitre, on traite le problème de l'estimation des paramètres d'un modèle $RCA(1)$ par la méthode des moindres carrées ensuite par la méthode du maximum de vraisemblance. Nous démontrons les deux principaux résultats. Le premier sur la consistance forte et le deuxième sur la loi asymptotique du vecteur des estimateurs obtenu par la méthode du maximum de vraisemblance.

Le troisième chapitre est consacré à l'étude de l'effet de l'existence de la racine unitaire sur les prévisions. Nous présentons aussi un test sur la racine unitaire basé sur la statistique de Wald.

Chapitre 1

Le modèle autorégressif à coefficients aléatoires RCA(1)

1.1 Généralités sur les processus

Définition 1.1. Un processus $(X_t, t \in \mathbb{Z})$ est dit strictement stationnaire ou fortement stationnaire si pour tous $t_1, t_2, ..., t_n \in \mathbb{Z}$ avec $t_1 < t_2 < ... < t_n$ et $h \in \mathbb{Z}$, le vecteur $(X_{t_1+h}, ..., X_{t_n+h})$ a la même loi de probabilité que $(X_{t_1}, ..., X_{t_n})$.

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

(i) $\forall t \in \mathbb{Z}, \qquad E(X_t^2) < \infty$

(ii) $\forall t \in \mathbb{Z}, \qquad E(X_t) = m$ est indépendant de t

(iii) $\forall (t, h) \in \mathbb{Z}^2, \qquad Cov(X_t, X_{t+h}) = E((X_t - m)(X_{t+h} - m)) = \gamma(h)$.

Définition 1.3. Soit $(\Omega, \mathcal{F}, P)$ un espace de probabilité. On dit que $T : \Omega \longrightarrow \Omega$ est une transformation qui préserve la mesure P si :

$$\forall F \in \mathcal{F}, \qquad P(T^{-1}(F)) = P(F).$$

Définition 1.4. L'opérateur shift S est défini sur $\mathbb{R}^\infty$ par $S(..., x_{i-1}, x_i, x_{i+1}, ...) = (..., x_i, x_{i+1}, x_{i+2}, ...)$.

Remarque : Si $X = (X_t, t \in \mathbb{Z})$ est un processus strictement stationnaire alors le shift S préserve PX^{-1}. la loi de X. En effet :

Soit $F = \prod\limits_{i \in Z} F_i \in \mathcal{B}_{\mathbb{R}_\infty}$ la tribu borélienne de $\mathbb{R}^\infty$ où $F_i \in \mathcal{B}_{\mathbb{R}}$. On a

$$\begin{aligned} S^{-1}(F) &= \{(..., x_i, x_{i+1}, ..., x_0, x_1, ...) \in \mathbb{R}^\infty \mid S(..., x_i, x_{i+1}, ..., x_0, x_1, ...) \in F\} \\ &= \{(..., x_i, x_{i+1}, ..., x_0, x_1, ...) \in \mathbb{R}^\infty \mid (..., x_{i-1}, x_i, ...) \in F\} \\ &= \prod_{i \in Z} F_{i-1} \end{aligned}$$

Par la stationnarité, on a

$$\begin{aligned} P(X \in S^{-1}(F)) &= P(\bigcap_{i \in Z}(X_i \in F_{i-1})) \\ &= P(\bigcap_{i \in Z}(X_{i-1} \in F_{i-1})) \\ &= P(X \in F) \end{aligned}$$

Définition 1.5. - Un ensemble mesurable $F \in \mathcal{F}$ est dit invariant par rapport à T si $T^{-1}(F) = F$.

- Une transformation qui préserve la mesure est dite ergodique si pour tout ensemble invariant F, $P(F) = 0$ ou $P(F) = 1$.

Le théorème suivant assure la convergence presque sure de la moyenne d'une suite strictement stationnaire et ergodique.

Théorème 1.1. [4] *(Théorème ergodique)*

Soit $\xi = (\xi_1, \xi_2, ...)$ *une suite aléatoire strictement stationnaire et ergodique avec* $E(\mid \xi_1 \mid) < \infty$. *Alors*

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} \xi_k = E(\xi_1) \qquad p.s$$

1.2 Le modèle autorégressif à coefficients aléatoires

1.2.1 Présentation du modèle RCA(1)

Considérons deux suites de variables aléatoires $(\epsilon_t, t \in \mathbb{Z})$ et $(\beta_t, t \in \mathbb{Z})$ avec les conditions suivantes :

- A1 : (ϵ_t) est une suite de variables aléatoires indépendantes de moyenne nulle et de variance σ^2.
- A2 : (β_t) est une suite de variables aléatoires indépendantes de moyenne μ_β et de variance σ_β^2.
- A3 : (β_t) et (ϵ_t) sont indépendantes.

Définition 1.6. Un processus $(X_t, t \in \mathbb{Z})$ est dit autorégressif à coefficients aléatoires $RCA(1)$, s'il vérifie l'équation suivante :

$$X_t = \alpha + \beta_t X_{t-1} + \epsilon_t \tag{1.1}$$

Remarque : Lorsque $\sigma_\beta^2 = 0$ le modèle $RCA(1)$ devient une série d'un modèle autorégressif $AR(1)$.

1.2.2 L'existence d'une solution strictement stationnaire pour le modèle RCA(1)

Nous commençons par énoncer le résultat de Billingsley (1995), qui sera utilisé dans la démonstration du théorème (2.4).

Lemme 1.1. [*1*] *Soit* $(X_t, t \in \mathbb{Z})$ *un processus stochastique. considérons un processus stochastique* $(Y_t, t \in \mathbb{Z})$ *avec* $Y_t = \Phi(..., X_{t-1}, X_t, X_{t+1}, ...)$, *ou* ϕ *est mesurable. Si* $(X_t, t \in \mathbb{Z})$ *est strictement stationnaire et ergodique, en particulier, si les* X_t *sont indépendants et identiquement distribués, alors* Y *est strictement stationnaire et ergodique.*

Quinn (1982) a étudié l'existence d'une solution strictement stationnaire et ergodique pour les modèles bilinéaires, définis par l'équation

$$X_n = aX_{n-1} + be_nX_{n-1} + e_n$$

où (e_n) une suite de variables aléatoires strictement stationnaire et ergodique.
Il a montré que les conditions $E(ln \mid a+be_n \mid) < 0$ et $(E(ln \mid a+be_n \mid) \leq 0)$ sont nécessaires et suffisantes pour l'existence de cette solution. Les mêmes arguments appliqués é un modèle $RCA(1)$ donne

Théorème 1.2. [5] *Considérons l'équation suivante :*

$$Y_n = \beta_n Y_{n-1} + \epsilon_n \tag{1.2}$$

où $\{(\beta_n, \epsilon_n)\}$ *est une suite de vecteurs aléatoires strictement stationnaire et ergodique avec*

$$E \mid ln \mid \epsilon_n \mid\mid < \infty \quad et \quad E \mid ln \mid \beta_n \mid\mid < \infty$$

Si $E(ln \mid \beta_n \mid) < 0$, *alors il existe une solution* $\{Y_n\}$ *strictement stationnaire et ergodique pour l'équation (1.2), et telle que pour tout* $n \in \mathbb{Z}$, Y_n *est mesurable par rapport à* $\mathcal{F}_n = \sigma(\{(\beta_t, \epsilon_t), t \leq n\})$ *la tribu engendrée par* $\{(\beta_t, \epsilon_t), t \leq n\}$ *.*

Démonstration. Supposons que

$$E(ln \mid \beta_n \mid) = \alpha < 0.$$

On définit la suite $(T_j)_{j\in\mathbb{N}}$ par :

$$T_0 = \epsilon_n \quad \text{et} \quad T_j = \prod_{i=0}^{j-1} \beta_{n-i}\epsilon_j, \quad \text{pour} \quad j \geq 1.$$

Soit $S_r = \sum\limits_{j=0}^{r} T_j$ on a :

$$\begin{aligned} ln \mid T_j \mid &= ln \prod_{i=0}^{j-1} \mid \beta_{n-i} \mid\mid \epsilon_j \mid \\ &= ln \mid \epsilon_j \mid + \sum_{i=0}^{j-1} ln \mid \beta_{n-i} \mid \end{aligned}$$

D'où

$$\frac{1}{j} ln \mid T_j \mid = \frac{1}{j} ln \mid \epsilon_{n-j} \mid + \frac{1}{j} \sum_{i=0}^{j-1} ln \mid \beta_{n-i} \mid .$$

Par le théorème ergodique, on obtient

$$\frac{1}{j} ln \mid T_j \mid \xrightarrow{p.s} E(ln \mid \beta_n \mid),$$

soit

$$\mid T_j \mid^{\frac{1}{j}} \xrightarrow{p.s} exp[E(ln \mid \beta_n \mid)] = exp(\alpha) < 1.$$

Maintenant on prend toutes les réalisations (T_j^*) pour lesquelles la convergence ci-dessus est vérifiée.

Étant donné que $exp(\alpha) < 1$ alors il existe $0 < \delta < 1$ tel que :

$$\mid T_j^* \mid < \delta^j < 1, \qquad \forall j > R_\delta.$$

D'où $\sum\limits_{j=0}^{r} \mid T_j^* \mid$ converge et par conséquent $S_r = \sum\limits_{j=0}^{r} T_j$ converge presque surement quand $r \longrightarrow \infty$.

Maintenant on pose :

$$\begin{aligned} Y_n^* &= \lim_{r \to \infty} S_r \\ &= \lim_{r \to \infty} (\epsilon_n + \sum_{j=1}^{r} T_j) \\ &= \epsilon_n + \sum_{j=1}^{\infty} \prod_{i=0}^{j-1} \beta_{n-i} \epsilon_j \end{aligned}$$

Il est évident que Y_n^* est une solution $\mathcal{F}_n$ mesurable pour l'équation (1.2). De plus par hypothèse $\{(\epsilon_n, \beta_n)\}$ est une suite de vecteurs strictement stationnaire et ergodique et par le lemme(1.1) on conclut que (Y_n^*) est strictement stationnaire et ergodique. $\square$

Théorème 1.3. [5] *Soit $(X_t, t \in \mathbb{Z})$ un $RCA(1)$ défini par la relation 1.2 et supposons que les hypothèses $A1 - A3$ sont satisfaites. Si les (β_t, ϵ_t) ont des distributions normales, alors une condition suffisante pour l'existence d'une solution strictement stationnaire et ergodique est :*

$$ln\,(\sigma_\beta^2) < \zeta + ln(2) - 2\int_0^1 \frac{1 - exp(-\lambda(1 - w^2))}{1 - w^2}\,dw, \tag{1.3}$$

où $\zeta \approx 0.57721$, ζ est la constante d'Euler, et $\lambda = \frac{\mu_\beta^2}{2\sigma_\beta^2}$.

Démonstration. Soit $e_t = \alpha + \epsilon_t$, il est clair que $((e_t, \beta_t), t \in \mathbb{Z})$ est une suite strictement stationnaire et ergodique.

Il est facile de voir que

$$E(\mid ln \mid e_t \mid\mid) < \infty \quad \text{et} \quad E(\mid ln \mid \beta_t \mid\mid) < \infty.$$

Afin de faire appel au théorème (1.1), nous devons montrer que

$$E(ln \mid \beta_t \mid) = E(ln \mid \mu_\beta + \sigma_\beta v_t \mid) < 0$$

où $v_t \sim \mathcal{N}(0,1)$.

Nous considérons deux cas :

1) Si $\mu_\beta = 0$, alors

$$\begin{aligned} E(ln \mid \beta_t \mid) &= E(ln \mid \sigma_\beta v_t \mid) \\ &= ln \mid \sigma_\beta \mid + E(ln \mid v_t \mid) \\ &= \frac{1}{2} ln(\sigma_\beta^2) + E(ln \mid v_t \mid) \\ &= \frac{1}{2} ln(\sigma_\beta^2) - \frac{1}{2}(\zeta + ln2) \\ &< 0. \end{aligned}$$

2) Si $\mu_\beta \neq 0$, alors :

$$\begin{aligned}E[ln \mid \beta_t \mid] &= E[ln \mid \mu_\beta + \sigma_\beta v_t \mid] \\ &= E[ln(\mid \mu_\beta \mid\mid 1 + \frac{\sigma_\beta}{\mu_\beta} v_t \mid)] \\ &= ln \mid \mu_\beta \mid + E[ln \mid 1 + \frac{\sigma_\beta}{\mu_\beta} v_t \mid] \\ &= \frac{1}{2} ln(\mu_\beta^2) + E[ln \mid 1 + Z \mid]\end{aligned}$$

avec $Z \sim \mathcal{N}(0, \frac{\sigma_\beta^2}{\mu_\beta^2})$

En utilisant l'évaluation de $E(ln \mid 1 + Z \mid)$, dans [3], on obtient

$$E[ln \mid \beta_t \mid] < 0 \Longleftrightarrow ln\, \sigma_\beta^2 < \zeta + ln2 - 2 \int_0^1 \frac{1 - exp(-\lambda(1 - w^2)}{1 - w^2}.$$

$\square$

1.2.3 L'existence d'une solution faiblement stationnaire

Nicholls et Quinn (1982) [3] ont obtenu les conditions nécessaires et suffisantes pour l'existence d'une solution faiblement stationnaire pour les modèles RCA d'ordre p. En particulier, le théorème suivant donne les conditions pour le modèle RCA(1).

Théorème 1.4. *Soit $\mathcal{F}_t$ la tribu engendrée par $((\beta_s, \epsilon_s), s \leq t)$. Alors il existe une unique solution $\mathcal{F}_t$ mesurable, faiblement stationnaire pour (1.1) si et seulement si $\mu_\beta^2 + \sigma_\beta^2 < 1$.*

1.3 Simulation du modèle RCA(1)

Nous présentons plusieurs séries du modèle RCA (1) (avec le logiciel R v3.2.0) simulé de longueur n = 500.

Pour générer la série RCA, nous avons fixé $\alpha = 0$ et $\sigma^2 = 1$ et utilisé des valeurs différentes pour μ_β et σ_β^2.

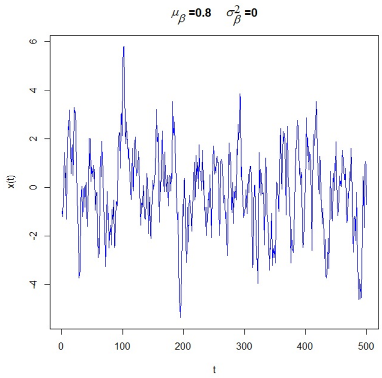

FIGURE 1.1 – X(t) avec $\mu_\beta = 0.8$ et $\sigma_\beta^2 = 0$

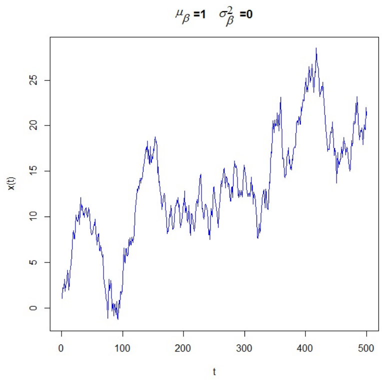

FIGURE 1.2 – X(t) avec $\mu_\beta = 1$ et $\sigma_\beta^2 = 0$

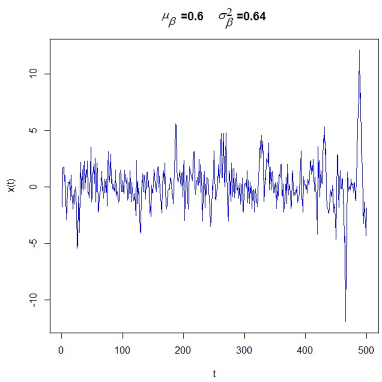

FIGURE 1.3 – X(t) avec $\mu_\beta = 0.6$ et $\sigma_\beta^2 = 0.64$

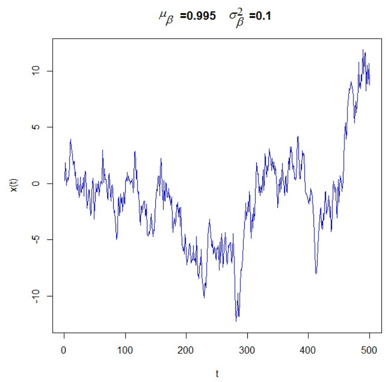

FIGURE 1.4 – X(t) avec $\mu_\beta = 0.995$ et $\sigma^2_\beta = 0.1$

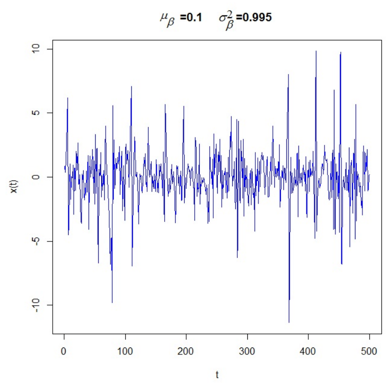

FIGURE 1.5 – X(t) avec $\mu_\beta = 0.1$ et $\sigma^2_\beta = 0.995$

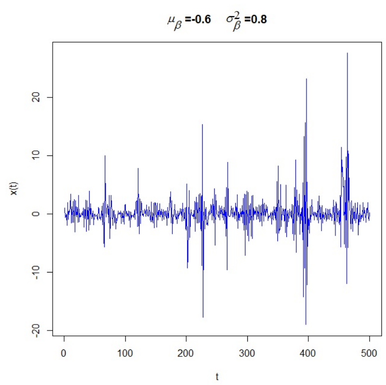

FIGURE 1.6 – X(t) avec $\mu_\beta = -0.6$ et $\sigma^2_\beta = 0.8$

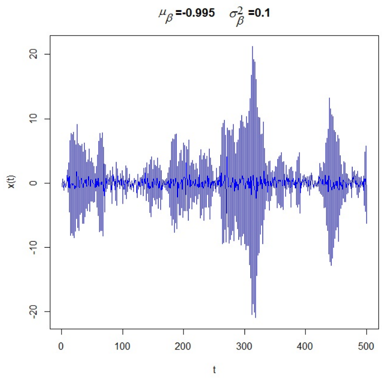

FIGURE 1.7 – X(t) avec $\mu_\beta = -0.995$ et $\sigma^2_\beta = 0.1$

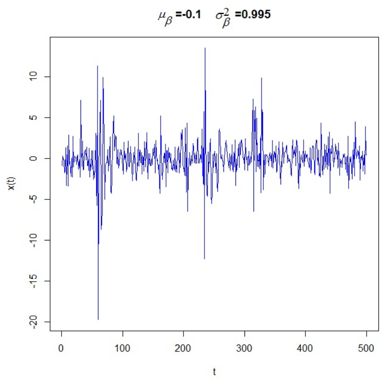

FIGURE 1.8 – X(t) avec $\mu_\beta = -0.1$ et $\sigma^2_\beta = 0.995$

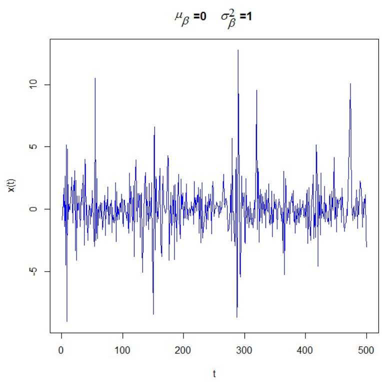

FIGURE 1.9 – X(t) avec $\mu_\beta = 0$ et $\sigma^2_\beta = 1$

Chapitre 2

Estimation du processus RCA(1)

Nicholls et Quinn suggèrent plusieurs méthodes pour estimer les paramètres du modèle RCA. Ils excluent le terme α et le modèle RCA(1) devient :

$$X_t = \beta_t X_{t-1} + \epsilon_t \tag{2.1}$$

où (β_t) est une suite de variable iid de moyenne μ_β et de variance σ_β^2, et (ϵ_t) une suite de variable iid de moyenne 0 et de variance σ^2, et (β_t) est indépendante de (ϵ_t) .

2.1 Estimation des paramètres du modèle RCA(1) par la méthode des moindres carrés

La première méthode d'estimation qu'ils proposent est basée sur le critère des moindres carrés, son procédé s'effectue en deux étapes.
Soit $\mathcal{F}_t = \sigma(X_s, s \leq t)$ la tribu engendrée par le processus jusqu'à l'instant t on a :

$$\begin{aligned} E(X_t \mid \mathcal{F}_{t-1}) &= E(\beta_t X_{t-1} + \epsilon_t \mid \mathcal{F}_{t-1}) \\ &= E(\beta_t X_{t-1} \mid \mathcal{F}_{t-1}) + E(\epsilon_t \mid \mathcal{F}_{t-1}) \end{aligned}$$

comme β_t et ϵ_t sont indépendantes de $\mathcal{F}_{t-1}$ et X_{t-1} est $\mathcal{F}_{t-1}$ mesurable alors

$$E(X_t \mid \mathcal{F}_{t-1}) = \mu_\beta X_{t-1}$$

Soit $u_t = X_t - \mu_\beta X_{t-1}$. Il en découle :

$$\begin{aligned} E(u_t \mid \mathcal{F}_{t-1}) &= E(X_t - \mu_\beta X_{t-1} \mid \mathcal{F}_{t-1}) \\ &= E(X_t \mid \mathcal{F}_{t-1}) - \mu_\beta X_{t-1} \\ &= 0, \end{aligned}$$

et

$$\begin{aligned} E(u_t^2 \mid F_{t-1}) &= E(X_t^2 \mid F_{t-1}) + \mu_\beta^2 X_{t-1}^2 - 2X_{t-1}\mu_\beta E(X_t \mid F_{t-1}) \\ &= X_{t-1}^2(\sigma_\beta^2 + \mu_\beta^2) + \sigma^2 + \mu_\beta^2 X_{t-1}^2 - 2\mu_\beta^2 X_{t-1}^2 \\ &= \sigma^2 + \sigma_\beta^2 X_{t-1}^2 \end{aligned}$$

Considérons un échantillon $(X_0, ..., X_n)$.

La première étape consiste à estimer μ_β en minimisant $\sum\limits_{t=1}^{n} u_t^2$ par rapport à μ_β. Puisque

$$\frac{\partial \sum\limits_{t=1}^{n} u_t^2}{\partial \mu_\beta} = -2 \sum_{t=1}^{n} X_{t-1}(X_t - \mu_\beta X_{t-1}),$$

et

$$\frac{\partial^2 \sum\limits_{t=1}^{n} u_t^2}{\partial \mu_\beta^2} = 2 \sum_{t=1}^{n} X_{t-1}^2 > 0,$$

alors l'estimateur $\hat{\mu}_\beta$ est donné par :

$$\hat{\mu}_\beta = (\sum_{t=1}^{n} X_{t-1}^2)^{-1} \sum_{t=1}^{n} X_{t-1} X_t.$$

La deuxième étape de l'estimation consiste à former le résidu

$$\hat{u}_t = X_t - \hat{\mu}_\beta X_{t-1}.$$

Pour estimer σ^2 et σ_β^2, on minimise

$$l(\sigma^2, \sigma_\beta^2) = \sum_{t=1}^{n} (\hat{u}_t^2 - \sigma^2 - \sigma_\beta^2 X_{t-1}^2)^2$$

par rapport à ces deux paramètres.

On a :

$$\frac{\partial}{\partial \sigma_\beta^2} \sum_{t=1}^{n} \left(\hat{u}_t^2 - \sigma^2 - \sigma_\beta^2 X_{t-1}^2\right)^2 = -2 \sum_{t=1}^{n} X_{t-1}^2 (\hat{u}_t^2 - \sigma^2 - \sigma_\beta^2 X_{t-1}^2) = 0$$

et

$$\frac{\partial^2}{\partial (\sigma_\beta^2)^2} \sum_{t=1}^{n} \left(\hat{u}_t^2 - \sigma^2 - \sigma_\beta^2 X_{t-1}^2\right)^2 = 2 \sum_{t=1}^{n} X_{t-1}^4 > 0,$$

donc

$$\hat{\sigma}_\beta^2 = \frac{\sum\limits_{t=1}^{n} X_{t-1}^2 \hat{u}_t^2 - \sigma^2 \sum\limits_{t=1}^{n} X_{t-1}^2}{\sum\limits_{t=1}^{n} X_{t-1}^4}.$$

Ensuite

$$\frac{\partial}{\partial \sigma^2} \sum_{t=1}^{n} \left(\hat{u}_t^2 - \sigma^2 - \sigma_\beta^2 X_{t-1}^2\right)^2 = -2 \sum_{t=1}^{n} (\hat{u}_t^2 - \sigma^2 - \sigma_\beta^2 X_{t-1}^2)$$

L'estimateur de $\hat{\sigma}^2$ est donné donc par :

$$\hat{\sigma}^2 = n^{-1} \sum_{t=1}^{n} \hat{u}_t^2 - n^{-1} \hat{\sigma}_\beta^2 \sum_{t=1}^{n} X_{t-1}^2,$$

soit

$$\hat{\sigma}^2 = n^{-1} \sum_{t=1}^{n} \hat{u}_t^2 - \hat{\sigma}_\beta^2 \overline{z},$$

avec $\overline{z} = n^{-1} \sum\limits_{t=1}^{n} X_{t-1}^2$.

2.2 Estimation du modèle RCA(1) par la méthode du maximum de vraisemblance

Dans cette section nous considérons Les hypothèses suivantes :

- A4 : $E(\epsilon_t^4) < \infty$ et $E(\beta_t^4) < \infty$.
- A5 : $\sigma^2 \geq \delta_1 > 0$ et $\sigma_\beta^2 \geq \delta_2 > 0$.
- A6 : $(\mu_\beta, \sigma_\beta^2) \in D$ où D est un ensemble fermé contenu dans la région de la stationnarité stricte et l'ergodicité donnée par l'inégalité (1.3).

L'hypothèse A4 sera utilisée pour prouver le théorème central limite pour l'estimateur du maximum de vraisemblance, et les hypothèses A5 et A6 sont imposées afin que l'ensemble des paramètres du modèle RCA(1)
$\Theta = \{(\alpha, \mu_\beta, \sigma_\beta^2, \sigma^2 : \alpha \in \mathcal{R}, (\mu_\beta, \sigma_\beta^2) \in D, \sigma^2 \geq \delta_1, \sigma_\beta^2 \geq \delta_2\}$ soit compact.

2.2.1 Calcul de l'estimateur du maximum de vraisemblance

Considérons l'échantillon $(X_0, X_1, ..., X_n)$ à partir d'un modèle RCA(1) qui satisfait (1.1), $\{X_t\}$ est strictement stationnaire, ergodique et $\mathcal{F}_t$ mesurable et satisfait l'équation (1.1) sous les conditions A1 et A6, nous fixons un X_0, et nous dérivons la fonction de vraisemblance en supposant la normalité de β_t et ϵ_t.
Il est facile de voir que :

$$E(X_t|X_{t-1}) = \alpha + \mu_\beta X_{t-1}$$

et

$$Var(X_t|X_{t-1}) = \sigma^2 + \sigma_\beta^2 X_{t-1}^2$$

Ici la fonction de vraisemblance conditionnelle par rapport à X_0 a la forme suivante :

$$\begin{aligned} f_n(X_1, ..., X_n|X_0) &= \prod_{t=1}^{n} f(X_t|X_{t-1}) \\ &= (2\pi)^{-\frac{n}{2}} \prod_{t=1}^{n} \{(\sigma^2 + \sigma_\beta^2 X_{t-1})^{\frac{-1}{2}} exp[-\frac{(X_t - \alpha - \mu_\beta X_{t-1})^2}{2(\sigma^2 + \sigma_\beta^2 X_{t-1}^2)}]\} \\ &= L_n(\alpha, \mu_\beta, \sigma_\beta^2, \sigma^2) \end{aligned}$$

On pose :

$$\begin{aligned} \tilde{l}_n(\alpha, \mu_\beta, \sigma_\beta^2, \sigma^2) &= -\frac{2}{n} ln[L_n(\alpha, \mu_\beta, \sigma_\beta^2, \sigma^2)] - ln(2\pi) \\ &= \frac{1}{n}\sum_{t=1}^{n} ln(\sigma^2 + \sigma_\beta^2 X_{t-1}^2) + \frac{1}{n}\sum_{t=1}^{n} \frac{(X_t - \alpha - \mu_\beta X_{t-1})^2}{\sigma^2 + \sigma_\beta^2 X_{t-1}^2} \end{aligned} \quad (2.2)$$

Il est plus pratique de minimiser $\tilde{l}_n(\alpha, \mu_\beta, \sigma_\beta^2, \sigma^2)$ que de maximiser $L_n(\alpha, \mu_\beta, \sigma_\beta^2, \sigma^2)$. Pour cela on pose $\tau = \frac{\sigma_\beta^2}{\sigma^2}$,et on donne la fonction $\tilde{l}_n$ qui dépend de τ.

$$\begin{aligned} \bar{l}_n(\alpha, \mu_\beta, \tau, \sigma^2) &= \tilde{l}_n(\alpha, \mu_\beta, \sigma_\beta^2, \sigma^2) \\ &= ln\sigma^2 + n^{-1}\sum_{t=1}^{n} ln(1+\tau X_{t-1}^2) + \sigma^{-2}n^{-1}\sum_{t=1}^{n}\frac{(X_t - \alpha - \mu_\beta X_{t-1})^2}{1+\tau X_{t-1}^2} \end{aligned} \tag{2.3}$$

En dérivant par rapport aux paramètres α et μ_β, on obtient

$$\frac{\partial}{\partial \alpha}\bar{l}_n(\alpha, \mu_\beta, \tau, \sigma^2) = -2\sigma^{-2}n^{-1}\sum_{t=1}^{n}\frac{(X_t - \alpha - \mu_\beta X_{t-1})}{1+\tau X_{t-1}^2}$$

$$\frac{\partial}{\partial \mu_\beta}\bar{l}_n(\alpha, \mu_\beta, \tau, \sigma^2) = -2\sigma^{-2}n^{-1}\sum_{t=1}^{n}\frac{(X_t - \alpha - \mu_\beta X_{t-1})X_{t-1}}{1+\tau X_{t-1}^2}$$

Maintenant on a :

$$\frac{\partial}{\partial \alpha}\bar{l}_n(\alpha, \mu_\beta, \tau, \sigma^2) = 0$$
$$\frac{\partial}{\partial \mu_\beta}\bar{l}_n(\alpha, \mu_\beta, \tau, \sigma^2) = 0$$

On trouve

$$\mu_\beta = \mu_{\beta,n}(\tau) = \frac{c_1c_2 - c_3c_4}{c_1c_5 - c_4^2}$$

$$\alpha = \alpha_n(\tau) = \frac{c_3 - \mu_{\beta,n}(\tau)c_4}{c_1}$$

Et

$$\sigma^2 = \sigma_n^2(\tau) = n^{-1}\sum_{t=1}^{n}\frac{[X_t - \alpha_n(\tau) - \mu_{\beta,n}(\tau)X_{t-1}]^2}{1+\tau X_{t-1}^2}$$

Où :

$$c_1 = \sum_{t=1}^{n}\frac{1}{1+\tau X_{t-1}^2}, c_2 = \sum_{t=1}^{n}\frac{X_tX_{t-1}}{1+\tau X_{t-1}^2}$$
$$c_3 = \sum_{t=1}^{n}\frac{X_t}{1+\tau X_{t-1}^2}, c_4 = \sum_{t=1}^{n}\frac{X_{t-1}}{1+\tau X_{t-1}^2}$$
$$c_5 = \sum_{t=1}^{n}\frac{X_{t-1}^2}{1+\tau X_{t-1}^2}$$

L'estimateur du maximum de vraisemblance de $\hat{\alpha}_n, \hat{\mu}_{\beta,n}, \hat{\sigma}^2_n$ et $\hat{\sigma}^2_{\beta,n}$, peut être obtenu en calculant $\hat{\tau}_n$.

Nous pouvons considérer le minimum de la fonction suivante :

$$l_n(\alpha, \mu_\beta, \tau) = \inf_{\sigma^2} \bar{l}_n(\alpha, \mu_\beta, \tau, \sigma^2) - 1$$

où

$$\bar{l}_n(\alpha, \mu_\beta, \tau, \sigma^2) = ln\, \sigma^2 + n^{-1} \sum_{t=1}^{n} ln(1 + \tau X^2_{t-1}) + \sigma^{-2} n^{-1} \sum_{t=1}^{n} \frac{(X_t - \alpha - \mu_\beta X_{t-1})^2}{1 + \tau X^2_{t-1}}$$

Calculons $\inf_{\sigma^2} \bar{l}_n(\alpha, \mu_\beta, \tau, \sigma^2)$

$$\frac{\partial \bar{l}_n(\alpha, \mu_\beta, \tau, \sigma^2)}{\partial \sigma^2} = \sigma^{-2} - \sigma^{-4} n^{-1} \sum_{t=1}^{n} \frac{(X_t - \alpha - \mu_\beta X_{t-1})^2}{1 + \tau X^2_{t-1}}$$

D'où :

$$\sigma^2 = n^{-1} \sum_{t=1}^{n} \frac{(X_t - \alpha - \mu_\beta X_{t-1})^2}{1 + \tau X^2_{t-1}}$$

Par suite

$$\inf_{\sigma^2} \bar{l}_n(\alpha, \mu_\beta, \tau, \sigma^2) = ln(n^{-1} \sum_{t=1}^{n} \frac{(X_t - \alpha - \mu_\beta X_{t-1})^2}{1 + \tau X^2_{t-1}}) + n^{-1} \sum_{t=1}^{n} ln(1 + \tau X^2_{t-1}) + 1$$

D'où :

$$l_n(\alpha, \mu_\beta, \tau) = ln(n^{-1} \sum_{t=1}^{n} \frac{(X_t - \alpha - \mu_\beta X_{t-1})^2}{1 + \tau X^2_{t-1}}) + n^{-1} \sum_{t=1}^{n} ln(1 + \tau X^2_{t-1})$$

$\hat{\tau}_n$ est obtenu en minimisant la fonction $l_n(\hat{\alpha}, \hat{\mu}_\beta, \tau_n)$ par rapport à τ. Comme $\tau = \frac{\sigma^2_\beta}{\sigma^2}$, alors $\hat{\sigma}^2_{\beta,n} = \hat{\tau}_n \hat{\sigma}^2_n$

2.2.2 La consistance forte de l'estimateur du maximum de vraisemblance

Il convient de noter que le minimum de $l_n(\alpha, \mu_\beta, \tau)$ est le même pour $l_n^*(\alpha, \mu_\beta, \tau)$ avec :

$$\begin{aligned}
l_n^*(\alpha, \mu_\beta, \tau) &= l_n(\alpha, \mu_\beta, \tau) - l_n(\alpha_0, \mu_{\beta,0}, \tau_0) \\
&= n^{-1}\sum_{t=1}^{n} ln(1+\tau X_{t-1}^2) + ln[n^{-1}\sum_{t=1}^{n}\frac{(X_t-\alpha-\mu_\beta X_{t-1})^2}{1+\tau X_{t-1}^2}] \\
&- n^{-1}\sum_{t=1}^{n} ln(1+\tau_0 X_{t-1}^2) - ln[n^{-1}\sum_{t=1}^{n}\frac{(X_t-\alpha_0-\mu_{\beta,0} X_{t-1})^2}{1+\tau_0 X_{t-1}^2}] \\
&= n^{-1}\sum_{t=1}^{n} ln\frac{1+\tau X_{t-1}^2}{1+\tau_0 X_{t-1}^2} + ln\{n^{-1}\sum_{t=1}^{n}\frac{(X_t-\alpha-\mu_\beta X_{t-1})^2}{1+\tau X_{t-1}^2}\} \\
&- ln\{n^{-1}\sum_{t=1}^{n}\frac{(X_t-\alpha_0-\mu_{\beta,0} X_{t-1})^2}{1+\tau_0 X_{t-1}^2}\}
\end{aligned} \tag{2.4}$$

Où $\alpha_0, \mu_{\beta,0}, \tau_0$ désignent certaines valeurs fixes de α, μ_β et τ

Lemme 2.1. *Sous les hypothèses A1–A3, la variable aléatoire Y_t définie par $Y_t = X_t^2$ est presque surement non constante pour tout t, où X_t est une solution strictement stationnaire et ergodique de l'équation (1.1).*

Lemme 2.2. *Considérons la fonction de la forme suivante :*

$$g(x) = \frac{ax^2+bx+c}{dx^2+e}$$

où $d > 0$ et $e > 0$. Alors il existe une constante $M > 0$ telle que $| g(x) | \leq M$ pour tout x.

Démonstration. On prend

$$M = \left|\frac{a}{d}\right| + \left|\frac{b}{2\sqrt{de}}\right| + \left|\frac{c}{e}\right| + 1$$

$\square$

Théorème 2.1. *Soit $\{X_t\}$ une suite $\mathcal{F}_t$ mesurable, strictement stationnaire et ergodique qui satisfait l'équation (1.1) avec $\theta = (\alpha, \mu_\beta, \sigma_\beta^2, \sigma^2)'$ sous les conditions A1, A3, A5 et A6. On définit $\pi = (\alpha, \mu_\beta, \tau)'$ où $\tau = \frac{\sigma_\beta^2}{\sigma^2}$. Soit Θ^* l'image de Θ par l'application continue $T(\theta) = \pi$, où Θ est l'espace des paramètres du modèle RCA(1).*
Soit $\pi_0 = (\alpha_0, \mu_{\beta,0}, \tau_0)'$, où $\tau_0 = \frac{\sigma_{\beta,0}^2}{\sigma_0^2}$. Alors la $\lim\limits_{n\to\infty} l_n^*(\alpha, \mu_\beta, \tau)$ existe presque surement pour tout $(\alpha, \mu_\beta, \tau) \in \Theta^*$ et sa limite $l^*(\alpha, \mu_\beta, \tau)$ est minimisée uniquement sur Θ^* en π_0 à condition que $\pi_0 \in int(\Theta^*)$.

Démonstration. Nicholls et Quinn [2] montrent que sous les conditions A1, A3 , A5 et A6, δ_1 et δ_2 peuvent être choisies de telle sorte que $(\alpha_0, \mu_{\beta,0}, \sigma_{\beta,0}^2, \sigma_0^2)' \in int(\Theta)$ et que l'ensemble Θ soit compact .
Comme T est une application continue et Θ est compact, alors Θ^* est compact de plus $\pi_0 \in int(\Theta^*)$ car $\pi_0 = T(\theta_0)$.
Maintenant par le lemme (1.1) les suites $(ln\frac{1+\tau X_{t-1}^2}{1+\tau_0 X_{t-1}^2})$, $(ln\{n^{-1}\sum\limits_{t=1}^{n}\frac{(X_t-\alpha-\mu_\beta X_{t-1})^2}{1+\tau X_{t-1}^2}\})$ et $(ln\{n^{-1}\sum\limits_{t=1}^{n}\frac{(X_t-\alpha_0-\mu_{\beta,0} X_{t-1})^2}{1+\tau_0 X_{t-1}^2}\})$ sont ergodiques. Ainsi, le théorème ergodique (1.1) assure la convergence des trois termes de
$l_n^*(\alpha, \mu_\beta, \tau) = n^{-1}\sum\limits_{t=1}^{n} ln\frac{1+\tau X_{t-1}^2}{1+\tau_0 X_{t-1}^2} + ln\{n^{-1}\sum\limits_{t=1}^{n}\frac{(X_t-\alpha-\mu_\beta X_{t-1})^2}{1+\tau X_{t-1}^2}\} - ln\{n^{-1}\sum\limits_{t=1}^{n}\frac{(X_t-\alpha_0-\mu_{\beta,0} X_{t-1})^2}{1+\tau_0 X_{t-1}^2}\}$.
Ces trois termes convergent respectivement vers $E[ln(\frac{1+\tau X_{t-1}^2}{1+\tau_0 X_{t-1}^2})]$, $ln[E(\frac{(X_t-\alpha-\mu_\beta X_{t-1})^2}{1+\tau X_{t-1}^2})]$ et $ln[E(\frac{(X_t-\alpha_0-\mu_{\beta,0} X_{t-1})^2}{1+\tau_0 X_{t-1}^2})]$, à condition que ces espérances sont finies.
Pour le premier terme, il est facile de voir que

$$max(1, \frac{\tau}{\tau_0}) \geq \frac{1+\tau X_{t-1}^2}{1+\tau_0 X_{t-1}^2} \geq min(1, \frac{\tau}{\tau_0})$$

Par suite $E \mid ln(\frac{1+\tau X_{t-1}^2}{1+\tau_0 X_{t-1}^2}) \mid < \infty$.

Pour le second terme, par le lemme (2.2) on a :

$$\begin{aligned}
E(\frac{(X_t-\alpha-\mu_\beta X_{t-1})^2}{1+\tau X_{t-1}^2}) &= \mid E\frac{(X_t-\alpha-\mu_\beta X_{t-1})^2}{1+\tau X_{t-1}^2} \mid \\
&\leq E \mid \frac{(X_t-\alpha-\mu_\beta X_{t-1})^2}{1+\tau X_{t-1}^2} \mid \\
&< \infty.
\end{aligned}$$

Pour le 3^{eme} terme, on a :

$$\begin{aligned}
0 < E\left[\frac{(X_t-\alpha_0-\mu_{\beta,0}X_{t-1})^2}{1+\tau_0 X_{t-1}^2}\right] &= E\left[E\left(\frac{(X_t-\alpha_0-\mu_{\beta,0}X_{t-1})^2}{1+\tau_0 X_{t-1}^2}|F_{t-1}\right)\right] \\
&= \sigma_0^2 \\
&< \infty
\end{aligned}$$

De plus : $E[\frac{(X_t-\alpha-\mu_\beta X_{t-1})^2}{1+\tau X_{t-1}^2}] = \sigma^2 > 0$ car autrement nous aurions $X_t = \alpha + \mu_\beta X_{t-1} = 0$ presque surement, ce qui est exclu par les hypothèses A5 et A6.

Il s'en suit que $l_n^*(\alpha,\mu_\beta,\tau)$ converge presque surement vers

$$l^*(\alpha,\mu_\beta,\tau) = E[ln(\frac{1+\tau X_{t-1}^2}{1+\tau_0 X_{t-1}^2})] + ln[E(\frac{(X_t-\alpha-\mu_\beta X_{t-1})^2}{1+\tau X_{t-1}^2})] - ln(\sigma_0^2)$$

D'autre part

$$\begin{aligned}
E[(\frac{(X_t-\alpha-\mu_\beta X_{t-1})^2}{1+\tau X_{t-1}^2})] &= E\left\{\frac{[(X_t-\alpha_0-\mu_{\beta,0}X_{t-1})+(\alpha_0-\alpha)+(\mu_{\beta,0}-\mu_\beta)X_{t-1}]^2}{1+\tau X_{t-1}^2}\right\} \\
&= E\left[\frac{(X_t-\alpha_0-\mu_{\beta,0}X_{t-1})^2}{1+\tau X_{t-1}^2}\right] + E\left\{\frac{[(\alpha_0-\alpha)+(\mu_{\beta,0}-\mu_\beta)X_{t-1}]^2}{1+\tau X_{t-1}^2}\right\} \\
&\geq E\left[\frac{(X_t-\alpha_0-\mu_{\beta,0}X_{t-1})^2}{1+\tau X_{t-1}^2}\right] \\
&\geq E\left[\frac{E\left((X_t-\alpha_0-\mu_{\beta,0}X_{t-1})^2 \mid \mathcal{F}_{t-1}\right)}{1+\tau X_{t-1}^2}\right] \\
&\geq E\left[\frac{\sigma_0^2+\sigma_{\beta,0}^2 X_{t-1}^2}{1+\tau X_{t-1}^2}\right] \\
&\geq \sigma_0^2 E[\frac{1+\tau_0 X_{t-1}^2}{1+\tau X_{t-1}^2}]
\end{aligned}$$

Pour avoir l'égalité, il suffit que : $(\alpha_0 - \alpha) + (\mu_{\beta,0} - \mu_\beta)X_{t-1} = 0$ presque surement, c'est à dire $\alpha = \alpha_0$ et $\mu_\beta = \mu_{\beta,0}$, puisque X_{t-1} est une variable aléatoire non constante par le lemme (2.1).

Par conséquent,

$$\begin{aligned}
\inf_{(\alpha,\mu_\beta)} l^*(\alpha,\mu_\beta,\tau) &= l^*(\alpha_0,\mu_{\beta,0},\tau) \\
&= E\left[ln(\frac{1+\tau X_{t-1}^2}{1+\tau_0 X_{t-1}^2})\right] + ln\left[E(\frac{(X_t-\alpha_0-\mu_{\beta,0}X_{t-1})^2}{1+\tau X_{t-1}^2})\right] - ln(\sigma_0^2) \\
&= E\left[ln(\frac{1+\tau X_{t-1}^2}{1+\tau_0 X_{t-1}^2})\right] + ln\left[E(\sigma^2\frac{1+\tau_0 X_{t-1}^2}{1+\tau X_{t-1}^2})\right] - ln(\sigma_0^2) \\
&= E\left[\ln\left(\frac{1+\tau X_{t-1}^2}{1+\tau_0 X_{t-1}^2}\right)\right] + \ln\left[E\left\{\frac{1+\tau_0 X_{t-1}^2}{1+\tau X_{t-1}^2}\right\}\right]
\end{aligned}$$

En suite $\inf\limits_{(\alpha,\mu_\beta,\tau)} l^*(\alpha,\mu_\beta,\tau) = l^*(\alpha_0,\mu_{\beta,0},\tau)$.

Maintenant, pour toute variable aléatoire non négative Y d'espérance 1, nous avons par l'inégalité de Jensen $E[ln(Y)] \leq ln[E(Y)] = 0$, l'égalité est vraie lorsque $Y = 1$ presque surement.

Soit $W = c^{-1}\frac{1+\tau_0 X_{t-1}^2}{1+\tau X_{t-1}^2}$, où $c = E[\frac{1+\tau_0 X_{t-1}^2}{1+\tau X_{t-1}^2}]$.Il est claire que $E(W) = 1$ et

$$\begin{aligned}
\inf_{(\alpha,\mu_\beta)} l^*(\alpha,\mu_\beta,\tau) &= E\left[\ln\left(\frac{1+\tau X_{t-1}^2}{1+\tau_0 X_{t-1}^2}\right)\right] + \ln\left[E\left\{\frac{1+\tau_0 X_{t-1}^2}{1+\tau X_{t-1}^2}\right\}\right] \\
&= E[ln(c^{-1}W^{-1})] + ln[E(Wc)] \\
&= E[-ln(cW)] + ln(c) \\
&= -E[ln(W)] \\
&\geq 0
\end{aligned}$$

L'égalité n'est vérifiée que lorsque $W = c^{-1}\frac{1+\tau_0 X_{t-1}^2}{1+\tau X_{t-1}^2} = 1$ presque surement.

D'où $(\tau_0 - c\tau)X_{t-1}^2 + (1-c) = 0$ presque surement.

Par le lemme(2.1), cela se produit uniquement lorsque $\tau = \tau_0$. Par conséquent $l^*(\alpha,\mu_\beta,\tau)$ est minimisé uniquement lorsque $\alpha = \alpha_0$, $\mu_\beta = \mu_{\beta,0}$ et $\tau = \tau_0$. $\square$

Corollaire 2.1. $\lim_{n\to\infty} \tilde{l}_n(\alpha, \mu_\beta, \sigma_\beta^2, \sigma^2)$*existe presque surement et elle est minimisée uniquement quand* $\alpha = \alpha_0$, $\mu_\beta = \mu_{\beta,0}$, $\sigma_\beta^2 = \sigma_{\beta,0}^2$ *et* $\sigma^2 = \sigma_0^2$.

Démonstration. Rappelons que :

$$\begin{aligned} l_n^*(\alpha, \mu_\beta, \tau) &= l_n(\alpha, \mu_\beta, \tau) - l_n(\alpha_0, \mu_{\beta,0}, \tau_0) \\ &= \inf_{\sigma^2} \bar{l}_n(\alpha, \mu_\beta, \tau, \sigma^2) - 1 - l_n(\alpha_0, \mu_{\beta,0}, \tau_0) \\ &= \inf_{\sigma^2} \tilde{l}_n(\alpha, \mu_\beta, \sigma_\beta^2, \sigma^2) - 1 - l_n(\alpha_0, \mu_{\beta,0}, \tau_0) \end{aligned}$$

Par le théorème (2.1) et la définition de $l_n(\alpha, \mu_\beta, \tau)$, on voit que $\lim_{n\to\infty} \tilde{l}_n(\alpha, \mu_\beta, \sigma_\beta^2, \sigma^2)$ existe presque surement et elle est uniquement minimisée quand $\alpha = \alpha_0$, $\mu_\beta = \mu_{\beta,0}$,
$\sigma^2 = \sigma^{2*} = E[\frac{(X_t - \alpha_0 - \mu_{\beta,0} X_{t-1})^2}{1+\tau_0 X_{t-1}^2}]$, et $\sigma_\beta^2 = \tau_0 \sigma^{2*}$
Mais

$$\begin{aligned} \sigma^{2*} &= E\left\{E\left[(X_t - \alpha_0 - \mu_{\beta,0} X_{t-1})^2 | \mathcal{F}_{t-1}\right] / (1 + \tau_0 X_{t-1}^2)\right\} \\ &= E\left[(\sigma_0^2 + \sigma_{\beta,0}^2 X_{t-1}^2)/(1 + \tau_0 X_{t-1}^2)\right] \\ &= \sigma_0^2, \end{aligned}$$

alors $\lim_{n\to\infty} \tilde{l}_n(\alpha, \mu_\beta, \sigma_\beta^2, \sigma^2)$ est minimisée uniquement quand $\alpha = \alpha_0$,
$\mu_\beta = \mu_{\beta,0}, \sigma_\beta^2 = \sigma_{\beta,0}^2$ et $\sigma^2 = \sigma_0^2$. □

Le théorème (2.1) est nécessaire dans la démonstration du théorème suivant, qui affirme la convergence presque sure des estimateurs du maximum de vraisemblance.

Théorème 2.2. *Soit* $l_n(\alpha, \mu_\beta, \tau)$ *minimisé sur* Θ^* *en* $\alpha = \hat{\alpha}_n$, $\mu_\beta = \hat{\mu}_{\beta,n}$ *et* $\tau = \hat{\tau}_n$ *où* Θ^* *est défini dans le théorème (2.1). Soit* $\hat{\pi}_n = (\hat{\alpha}_n, \hat{\mu}_{\beta,n}, \hat{\tau}_n)$. *Alors* $\hat{\pi}_n$ *converge presque surement vers* $\pi_0 = (\alpha_0, \mu_{\beta,0}, \tau_0)$ *à condition que* $\pi_0 \in int(\Theta^*)$. *De plus* $\hat{\sigma}_{\beta,n}^2$ *et* $\hat{\sigma}_n^2$ *convergent presque surement vers* σ_β^2 *et* σ^2.

2.2.3 Le théorème central limite

Théorème 2.3. *[1]*

Soit $\{\xi_t\}$ *une suite de variables aléatoires avec la propriété que* ξ_t *peut être exprimée comme une fonction qui ne dépend pas de* t*, mesurable par rapport à* $\mathcal{F}_t$ *générée par une suite* $\{\alpha_t, \alpha_{t-1}, ...\}$ *qui est strictement stationnaire et ergodique. De plus on suppose que* $E(\xi_t \mid \mathcal{F}_{t-1}) = 0$ *et* $E(\xi_t^2) = c^2 < \infty$*, alors* $(c^2 n)^{-\frac{1}{2}} \sum\limits_{t=1}^{n} \xi_t$ *converge en distribution vers une variable normale standard.*

Lemme 2.3. *Soit* (f_n) *une suite de fonctions dérivable sur un compact* Θ*.*
Si $\lim\limits_{n\to\infty} \sup\limits_{\theta\in\Theta} \| \frac{\partial f_n(\theta)}{\partial\theta} \| < \infty$ *alors* (f_n) *est équicontinue sur* Θ*.*

Démonstration. Nous devons montrer que pour tout $\varepsilon > 0$, il existe un entier N et un $\delta > 0$, qui dépend de ε tel que $|f_n(\theta_1) - f_n(\theta_2)| < \varepsilon$ pour $n > N$ quand $\|\theta_1 - \theta_2\| < \delta$.
Comme $f_n(\theta)$ est continue et dérivable sur Θ, par le théorème de la moyenne, on a pour $\theta_1, \theta_2 \in \Theta$

$$f_n(\theta_1) - f_n(\theta_2) = (\theta_1 - \theta_2)' \frac{\partial f_n(\theta^*)}{\partial\theta}$$

où

$$\theta^* = \lambda\theta_1 + (1-\lambda)\theta_2 \qquad \text{pour} \qquad \lambda \in (0,1)$$

On a :

$$\left| (\theta_1 - \theta_2)' \frac{\partial f_n(\theta^*)}{\partial\theta} \right|^2 \leq \|\theta_1 - \theta_2\|^2 \left\| \frac{\partial f_n(\theta^*)}{\partial\theta} \right\|^2$$

Il suivra que $f_n(\theta)$ est équicontinue si $\lim\limits_{n\to\infty} \sup\limits_{\theta\in\Theta} \| \frac{\partial f_n(\theta)}{\partial\theta} \| < \infty$. □

Le lemme indiqué ci dessus peut être utilisé pour montrer l'équicontinuité de la dérivée seconde de $\tilde{l}_n(\theta)$ sur un voisinage compact de θ_0.
Suivant ce lemme, il suffit de montrer que chaque élément de la 3^{eme} dérivée de $\tilde{l}_n(\theta)$ est

borné uniformément sur un voisinage de θ_0.

Par exemple, pour

$$\frac{\partial^2 \widetilde{l_n}(\theta)}{\partial \mu_\beta^2} = 2n^{-1} \sum_{t=1}^{n} \lambda_t^{-1} X_{t-1}^2$$

où $\lambda_t = \sigma^2 + \sigma_\beta^2 X_{t-1}^2$, on a

$$\begin{aligned} \mid \frac{\partial^3 \widetilde{l_n}(\theta)}{\partial \mu_\beta^2 \partial \sigma^2} \mid &= 2n^{-1} \mid \sum_{t=1}^{n} \lambda_t^{-2} X_{t-1}^2 \mid \\ &\leq 2n^{-1} \sum_{t=1}^{n} \mid \lambda_t^{-1} X_{t-1}^2 \mid\mid \lambda_t^{-1} \mid \\ &\leq 2(\frac{1}{\sigma_\beta^2} + 1)(\frac{1}{\sigma^2} + 1) \\ &< \infty \end{aligned}$$

On procède de la même manière pour les autres éléments de la dérivée de $\left\{ \frac{\partial^2 \widetilde{l_n}(\theta)}{\partial \theta \partial \theta'} \right\}$.

Lemme 2.4. *Soit $\{\theta_n\}$ une suite de variables aléatoires telle que θ_n converge presque surement vers une constante θ_0. Soit $\{f_n(.)\}$ une suite de fonction telle que f_n converge uniformément vers une fonction continue f sur un voisinage de θ_0 $(N(\theta_0))$. Alors $f_n(\theta_n)$ converge presque surement vers $f(\theta_0)$.*

Démonstration. Soit ω tel que $\theta_n(\omega)$ converge vers $\theta_0(\omega)$. On a alors pour un n suffisamment grand, $\theta_n(\omega) \in N(\theta_0)$.

$$\begin{aligned} |f_n(\theta_n) - f(\theta_0)| &= |f_n(\theta_n) - f(\theta_n) + f(\theta_n) - f(\theta_0)| \\ &\leq |f_n(\theta_n) - f(\theta_n)| + |f(\theta_n) - f(\theta_0)| \end{aligned}$$

Puisque $f_n(.)$ converge uniformément vers $f(.)$ alors

$$|f_n(\theta_n(\omega)) - f(\theta_n(\omega))| < \frac{\varepsilon}{2}$$

f est continue sur $N(\theta_0)$ donc il existe $N \in \mathbb{N}$ tel que si $n \geq N$, alors

$$|f_n(\theta_n) - f(\theta_0)| < \varepsilon$$

D'où $f_n(\theta_n)$ converge presque surement vers $f(\theta_0)$. $\square$

Lemme 2.5. *La suite* $\left\{\frac{\partial^2 \widetilde{l}_n(\theta)}{\partial\theta\partial\theta'}\right\}$ *converge uniformément et presque surement sur un compact au voisinage de* θ_0 *vers* $\left\{\frac{\partial^2 \widetilde{l}(\theta)}{\partial\theta\partial\theta'}\right\}$, *où* $\widetilde{l}(\theta) = \lim\limits_{n\to\infty} \widetilde{l}_n(\theta)$.

Démonstration. On définit $\lambda_t = \sigma^2 + \sigma_\beta^2 X_{t-1}^2$ et $u_t = X_t - \alpha - \mu_\beta X_{t-1}$. On a

$$\begin{aligned} E(u_t \mid \mathcal{F}_{t-1}) &= E[(X_t - \alpha_0 - \mu_{\beta,0}X_{t-1}) + (\alpha_0 - \alpha) + (\mu_{\beta,0} - \mu_\beta)X_{t-1} \mid \mathcal{F}_{t-1}] \\ &= \alpha_0 - \alpha + (\mu_{\beta,0} - \mu_\beta)X_{t-1} \end{aligned}$$

et

$$\begin{aligned} E(u_t^2 \mid \mathcal{F}_{t-1}) &= E\{[(X_t - \alpha_0 - \mu_{\beta,0}X_{t-1}) + (\alpha_0 - \alpha) + (\mu_{\beta,0} - \mu_\beta)X_{t-1}]^2 \mid \mathcal{F}_{t-1}\} \\ &= \sigma_0^2 + \sigma_{\beta,0}^2 X_{t-1}^2 + [(\alpha_0 - \alpha) + (\mu_{\beta,0} - \mu_\beta)X_{t-1}]^2 \end{aligned}$$

Maintenant les dérivées secondes de $\tilde{l}_n(\theta)$ sont données par :

$$\frac{\partial^2 \tilde{l}_n(\theta)}{\partial \alpha^2} = 2n^{-1}\sum_{t=1}^{n} \lambda_t^{-1}$$

$$\frac{\partial^2 \tilde{l}_n(\theta)}{\partial \alpha \partial \mu_\beta} = 2n^{-1}\sum_{t=1}^{n} \lambda_t^{-1} X_{t-1}$$

$$\frac{\partial^2 \tilde{l}_n(\theta)}{\partial \alpha \sigma_\beta^2} = 2n^{-1}\sum_{t=1}^{n} \lambda_t^{-2} u_t X_{t-1}^2$$

$$\frac{\partial^2 \tilde{l}_n(\theta)}{\partial \alpha \partial \sigma^2} = 2n^{-1}\sum_{t=1}^{n} \lambda_t^{-2} u_t$$

$$\frac{\partial^2 \tilde{l}_n(\theta)}{\partial \mu_\beta^2} = 2n^{-1}\sum_{t=1}^{n} \lambda_t^{-1} X_{t-1}^2$$

$$\frac{\partial^2 \tilde{l}_n(\theta)}{\partial \mu_\beta \partial \sigma_\beta^2} = 2n^{-1}\sum_{t=1}^{n} \lambda_t^{-2} u_t X_{t-1}^3$$

$$\frac{\partial^2 \tilde{l}_n(\theta)}{\partial \mu_\beta \partial \sigma^2} = 2n^{-1}\sum_{t=1}^{n} \lambda_t^{-2} u_t X_{t-1}$$

$$\frac{\partial^2 \tilde{l}_n(\theta)}{\partial (\sigma_\beta^2)^2} = 2n^{-1}\sum_{t=1}^{n} \lambda_t^{-3} u_t^2 X_{t-1}^4 - n^{-1}\sum_{t=1}^{n} \lambda_t^{-2} X_{t-1}^4$$

$$\frac{\partial^2 \tilde{l}_n(\theta)}{\partial \sigma_\beta^2 \partial \sigma^2} = 2n^{-1}\sum_{t=1}^{n} \lambda_t^{-3} u_t^2 X_{t-1}^2 - n^{-1}\sum_{t=1}^{n} \lambda_t^{-2} X_{t-1}^2$$

$$\frac{\partial^2 \tilde{l}_n(\theta)}{\partial (\sigma^2)^2} = 2n^{-1}\sum_{t=1}^{n} \lambda_t^{-3} u_t^2 - n^{-1}\sum_{t=1}^{n} \lambda_t^2$$

Par le théorème ergodique, la matrice $(\frac{\partial^2 \tilde{l}_n(\theta)}{\partial\theta\partial\theta'})$ converge presque surement vers une matrice

M dont les éléments $(M_{ij})_{ij}$ sont donnés par :

$$\begin{aligned}
M_{11} &= \frac{\partial^2 \tilde{l}(\theta)}{\partial \alpha^2} = 2E(\lambda_t^{-1})\\
M_{12} = M_{21} &= \frac{\partial^2 \tilde{l}(\theta)}{\partial \alpha \partial \mu_\beta} = 2E(\lambda_t^{-1} X_{t-1})\\
M_{13} = M_{31} &= \frac{\partial^2 \tilde{l}(\theta)}{\partial \alpha \partial \sigma_\beta^2} = 2E(\lambda_t^{-2} u_t X_{t-1}^2)\\
M_{14} = M_{41} &= \frac{\partial^2 \tilde{l}(\theta)}{\partial \alpha \partial \sigma^2} = 2E(\lambda_t^{-2} u_t)\\
M_{22} &= \frac{\partial^2 \tilde{l}(\theta)}{\partial \mu_\beta^2} = 2E(\lambda_t^{-1} X_{t-1}^2)\\
M_{23} = M_{32} &= \frac{\partial^2 \tilde{l}(\theta)}{\partial \mu_\beta \partial \sigma_\beta^2} = 2E(\lambda_t^{-2} u_t X_{t-1}^3)\\
M_{24} = M_{42} &= \frac{\partial^2 \tilde{l}(\theta)}{\partial \mu_\beta \partial \sigma^2} = 2E(\lambda_t^{-2} u_t X_{t-1})\\
M_{33} &= \frac{\partial^2 \tilde{l}(\theta)}{\partial (\sigma_\beta^2)^2} = 2E(\lambda_t^{-3} u_t^2 X_{t-1}^4) - E(\lambda_t^{-2} X_{t-1}^4)\\
M_{34} = M_{43} &= \frac{\partial^2 \tilde{l}(\theta)}{\partial \sigma_\beta^2} = 2E(\lambda_t^{-3} u_t^2 X_{t-1}^2) - E(\lambda_t^{-2} X_{t-1}^2)
\end{aligned}$$

et

$$M_{44} = \frac{\partial^2 \tilde{l}(\theta)}{\partial (\sigma^2)^2} = 2E(\lambda_t^{-3} u_t^2) - E(\lambda_t^{-2})$$

Par le lemme(2.3),$(\frac{\partial^2 \tilde{l_n}(\theta)}{\partial \theta \partial \theta'})$ est équicontinue sur un voisinage compact de θ_0, d'où $(\frac{\partial^2 \tilde{l_n}(\theta)}{\partial \theta \partial \theta'})$ converge uniformément et presque surement sur un voisinage de θ_0 vers $(\frac{\partial^2 \tilde{l}(\theta)}{\partial \theta \partial \theta'})$. □

Lemme 2.6. *Pour toute suite $\{\theta_n\}$ qui converge presque surement vers θ_0, on a $\{\frac{\partial^2 \tilde{l}_n(\theta_n)}{\partial \theta \partial \theta'}\}$ converge presque surement vers $\frac{\partial^2 \tilde{l}(\theta_0)}{\partial \theta \partial \theta'} = V$, où V est une matrice symétrique, donnée par :*

$$\begin{pmatrix} V_{11} & V_{12} & V_{13} & V_{14}\\ & V_{22} & V_{23} & V_{24}\\ & & V_{33} & V_{34}\\ & & & V_{44}\end{pmatrix} = \begin{pmatrix} 2E(\lambda_{t,0}^{-1}) & 2E(\lambda_{t,0}^{-1} X_{t-1}) & 0 & 0\\ & 2E(\lambda_{t,0}^{-1} X_{t-1}^2) & 0 & 0\\ & & E(\lambda_{t,0}^{-2} X_{t-1}^4) & E(\lambda_{t,0}^{-2} X_{t-1}^2)\\ & & & E(\lambda_{t,0}^{-2})\end{pmatrix} \quad (2.5)$$

où

$$\lambda_{t,0} = \sigma_0^2 + \sigma_{\beta,0}^2 X_{t-1}^2$$

Démonstration. D'après le lemme (2.5) $\{\frac{\partial^2 \tilde{l}_n(\theta_n)}{\partial\theta\partial\theta'}\}$ converge uniformément et presque surement vers $\frac{\partial^2 \tilde{l}(\theta_0)}{\partial\theta\partial\theta'} = V$, et comme $\theta_n \longrightarrow \theta_0$ ps, alors $\{\frac{\partial^2 \tilde{l}_n(\theta_n)}{\partial\theta\partial\theta'}\}$ converge presque surement vers $\frac{\partial^2 \tilde{l}(\theta_0)}{\partial\theta\partial\theta'} = V$,et chaque élément V_{ij} de V est donné par :

$$V_{11} = \frac{\partial^2 \tilde{l}(\theta_0)}{\partial\alpha^2} = 2E(\lambda_{t,0}^{-1})$$

$$V_{12} = V_{21} = \frac{\partial^2 \tilde{l}(\theta_0)}{\partial\alpha\partial\mu_\beta} = 2E(\lambda_{t,0}^{-1} X_{t-1})$$

$$V_{13} = V_{31} = \frac{\partial^2 \tilde{l}(\theta_0)}{\partial\alpha\partial\sigma_\beta^2} = 2E(\lambda_{t,0}^{-2} u_{t,0} X_{t-1}^2) = 2E(\lambda_{t,0}^{-2} X_{t-1}^2 E(u_{t,0} \mid \mathcal{F}_{t-1})) = 0$$

$$V_{14} = V_{41} = \frac{\partial^2 \tilde{l}(\theta_0)}{\partial\alpha\partial\sigma^2} = 2E(\lambda_{t,0}^{-2} u_{t,0}) = 2E(\lambda_{t,0}^{-2} E(u_{t,0} \mid \mathcal{F}_{t-1})) = 0$$

$$V_{22} = \frac{\partial^2 \tilde{l}(\theta_0)}{\partial\mu_\beta^2} = 2E(\lambda_{t,0}^{-1} X_{t-1}^2)$$

$$V_{23} = V_{32} = \frac{\partial^2 \tilde{l}(\theta_0)}{\partial\mu_\beta\partial\sigma_\beta^2} = 2E(\lambda_{t,0}^{-2} u_{t,0} X_{t-1}^3) = 2E(\lambda_{t,0}^{-2} X_{t-1}^3 E(u_{t,0} \mid \mathcal{F}_{t-1})) = 0$$

$$V_{24} = V_{42} = \frac{\partial^2 \tilde{l}(\theta_0)}{\partial\mu_\beta\partial\sigma^2} = 2E(\lambda_{t,0}^{-2} u_{t,0} X_{t-1}) = 2E(\lambda_{t,0}^{-2} X_{t-1} E(u_{t,0} \mid \mathcal{F}_{t-1})) = 0$$

$$V_{33} = \frac{\partial^2 \tilde{l}(\theta_0)}{\partial(\sigma_\beta^2)^2} = 2E(\lambda_{t,0}^{-3} u_{t,0}^2 X_{t-1}^4) - E(\lambda_{t,0}^{-2} X_{t-1}^4) = E(\lambda_{t,0}^{-2} X_{t-1}^4)$$

$$V_{34} = V_{43} = \frac{\partial^2 \tilde{l}(\theta_0)}{\partial\sigma_\beta^2\partial\sigma^2} = 2E(\lambda_{t,0}^{-3} u_{t,0}^2 X_{t-1}^2) - E(\lambda_{t,0}^{-2} X_{t-1}^2) = E(\lambda_{t,0}^{-2} X_{t-1}^2)$$

$V_{44} = \frac{\partial^2 \tilde{l}(\theta_0)}{\partial(\sigma^2)^2} = 2E(\lambda_{t,0}^{-3} u_{t,0}^2) - E(\lambda_{t-1}^{-2}) = E(\lambda_{t,0}^{-2})$

nous utilisons le fait que : $E(u_{t,0} \mid \mathcal{F}_{t-1}) = 0$ et $E(u_{t,0}^2 \mid \mathcal{F}_{t-1}) = \lambda_{t,0}$
où : $u_{t,0} = X_t - \alpha_0 - \mu_{\beta,0} X_{t-1}$. □

Théorème 2.4. *Soit $\{X_t\}$ une suite strictement stationnaire, ergodique, et une solution $\mathcal{F}_t$ mesurable pour l'équation (1.1) sous les conditions $A1 - A6$.*
Alors $\sqrt{n}(\hat{\theta}_n - \theta_0)$ converge vers une loi normale d'ésperance nulle et de matrice de covariance $V^{-1}WV^{-1}$ où V est donné dans (2.5) et W est une matrice symétrique, ces éléments sont donnés par :

$W_{11} = 4E(\lambda_{t,0}^{-1})$, $\quad W_{12} = 4E(\lambda_{t,0}^{-1} X_{t-1})$

$W_{13} = 2E(u_{t,0} \lambda_{t,0}^{-3} X_{t-1}^2 \eta_{t,0})$, $\quad W_{14} = 2E(u_{t,0} \lambda_{t,0}^{-3} \eta_{t,0})$

$W_{22} = 4E(\lambda_{t,0}^{-1} X_{t-1}^2)$, $\quad W_{23} = 2E(u_{t,0} \lambda_{t,0}^{-3} X_{t-1}^3 \eta_{t,0})$

$W_{24} = 2E(u_{t,0} \lambda_{t,0}^{-3} X_{t-1} \eta_{t,0})$, $\quad W_{33} = E(\lambda_{t,0}^{-4} X_{t-1}^4 \eta_{t,0}^2)$

$W_{34} = E(\lambda_{t,0}^{-4} X_{t-1}^2 \eta_{t,0}^2)$, $\quad W_{44} = E(\lambda_{t,0}^{-2} \eta_{t,0}^2)$

où :

$$\lambda_{t,0} = \sigma_0^2 + \sigma_{\beta,0}^2 X_{t-1}^2 \quad u_{t,0} = X_t - \alpha_0 - \mu_{\beta,0} X_{t-1} \quad et \quad \eta_{t,0} = u_{t,0}^2 - \lambda_{t,0}$$

De plus, si $\{\beta_t\}$ et $\{\epsilon_t\}$ sont conjointement normales, alors la matrice de covariance se réduit à $2V^{-1}$.

Démonstration. Nous montrons d'abord que la matrice de covariance V est définie positive. Puisque V est diagonale par bloc, il suffit de montrer que chaque bloc est défini positif.
Considérons la sous matrice :

$$\begin{aligned}
\begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix} &= \begin{pmatrix} 2E(\lambda_{t,0}^{-1}) & 2E(\lambda_{t,0}^{-1}X_{t-1}) \\ 2E(\lambda_{t,0}^{-1}X_{t-1}) & 2E(\lambda_{t,0}^{-1}X_{t-1}^2) \end{pmatrix} \\
&= E[2\lambda_{t,0}^{-1} \begin{pmatrix} 1 \\ X_{t-1} \end{pmatrix} (1, X_{t-1})] \\
&= E(zz^{'})
\end{aligned}$$

où

$$z = (\sqrt{2}\lambda_{t,0}^{-\frac{1}{2}}, \sqrt{2}\lambda_{t,0}^{-\frac{1}{2}} X_{t-1})^{'}.$$

Pour tout vecteur $a = (a_1, a_2)^{'}$, on a :

$$E(a^{'}zz^{'}a) = E[2\lambda_{t,0}^{-1}(a_1 + a_2 X_{t-1})^2] \geq 0$$

et par le lemme(2.1) l'inégalité n'est vraie que lorsque $a_1 = a_2 = 0$.

Par la même méthode on montre aussi que la sous matrice

$$\begin{pmatrix} V_{33} & V_{34} \\ V_{43} & V_{44} \end{pmatrix}$$

est définie positive. D'où V est définie positive.

Par ailleurs, nous avons montré dans le lemme (2.5) que la dérivée seconde de $\{\frac{\partial^2 \tilde{l}_n(\theta)}{\partial\theta\partial\theta'}\}$ converge uniformément et presque surement sur un voisinage compact de θ_0 vers $\{\frac{\partial^2 \tilde{l}(\theta)}{\partial\theta\partial\theta'}\}$.

Le développement de Taylor de $\frac{\partial \tilde{l}_n(\hat{\theta_n})}{\partial\theta}$ autour de θ_0 est donné par :

$$\frac{\partial \tilde{l}_n(\hat{\theta}_n)}{\partial\theta} = \frac{\partial \tilde{l}_n(\theta_0)}{\partial\theta} + \frac{\partial^2 \tilde{l}_n(\tilde{\theta}_n)}{\partial\theta\partial\theta'}(\hat{\theta}_n - \theta_0)$$

Puisque $\hat{\theta_n}$ converge presque surement vers θ_0, alors $\tilde{\theta_n}$ converge presque surement vers θ_0, et par le lemme (2.6), on a $\frac{\partial^2 \tilde{l}_n(\tilde{\theta}_n)}{\partial\theta\partial\theta'}$ converge presque surement vers $\frac{\partial^2 \tilde{l}(\theta_0)}{\partial\theta\partial\theta'} = V$.

Comme $\hat{\theta}_n$ minimise $\tilde{l}_n(\theta)$, on a $\frac{\partial \tilde{l}_n(\hat{\theta}_n)}{\partial\theta} = 0$.

Si nous supposons que $\sqrt{n}\frac{\partial \tilde{l}_n(\theta_0)}{\partial \theta}$ a une distribution limite de moyenne nulle et de matrice de covariance W, alors $\sqrt{n}(\hat{\theta}_n - \theta_0)$ aura la même distribution asymptotique que $-V^{-1}\sqrt{n}\frac{\partial \tilde{l}_n(\theta_0)}{\partial \theta}$. Il en découle que $\sqrt{n}(\hat{\theta}_n - \theta_0)$ converge vers une normale de moyenne nulle et de matrice de covariance $V^{-1}WV^{-1}$.

Maintenant on montre que $\sqrt{n}\frac{\partial \tilde{l}_n(\theta_0)}{\partial \theta}$ est asymptotiquement normale de moyenne nulle et de matrice de covariance W

Rappelons que :

$$\tilde{l}_n(\theta_0) = n^{-1}\sum_{t=1}^{n} ln(\sigma_0^2 + \sigma_{\beta,0}^2 X_{t-1}^2) + n^{-1}\sum_{t=1}^{n} \frac{(X_t - \alpha_0 - \mu_{\beta,0}X_{t-1})^2}{\sigma_0^2 + \sigma_{\beta,0}^2 X_{t-1}^2}$$

Donc :

$$\begin{aligned}
Y_1^n &= \frac{\partial \tilde{l}_n(\theta_0)}{\partial \alpha} = -2n^{-1}\sum_{t=1}^{n} \lambda_{t,0}^{-1} u_{t,0} \\
Y_2^n &= \frac{\partial \tilde{l}_n(\theta_0)}{\partial \mu_\beta} = -2n^{-1}\sum_{t=1}^{n} \lambda_{t,0}^{-1} u_{t,0} X_{t-1} \\
Y_3^n &= \frac{\partial \tilde{l}_n(\theta_0)}{\partial \sigma_\beta^2} = n^{-1}\sum_{t=1}^{n} \lambda_{t,0}^{-1} X_{t-1}^2 - n^{-1}\sum_{t=1}^{n} \lambda_{t,0}^{-2} u_{t,0}^2 \\
Y_4^n &= \frac{\partial \tilde{l}_n(\theta_0)}{\partial \sigma^2} = n^{-1}\sum_{t=1}^{n} \lambda_{t,0}^{-1} - n^{-1}\sum_{t=1}^{n} \lambda_{t,0}^{-2} u_{t,0}^2
\end{aligned}$$

Soit

$$\xi_t(a) = \lambda_{t,0}^{-2}[2u_{t,0}\lambda_{t,0}(a_1 + a_2 X_{t-1}) + \eta_{t,0}(a_4 + a_3 X_{t-1}^2)$$

où $a = (a_1, a_2, a_3, a_4)'$ Il est facile de voir que

$$n^{-1}\sum_{t=1}^{n} \xi_t(a) = -a' \frac{\partial \tilde{l}_n(\theta_0)}{\partial \theta}$$

Comme : $E(u_{t,0} \mid \mathcal{F}_{t-1}) = 0$ et $E(u_{t,0}^2 \mid \mathcal{F}_{t-1}) = \lambda_{t,0}$, alors

$E(\eta_{t,0} \mid \mathcal{F}_{t-1}) = E(u_{t,0}^2 - \lambda_{t,0} \mid \mathcal{F}_{t-1}) = 0$ par suite : $E(\xi_t(a) \mid \mathcal{F}_{t-1}) = 0$.

D'autre part

$$\xi_t^2(a) = 4u_{t,0}^2\lambda_{t,0}^{-2}(a_1 + a_2X_{t-1})^2 + \lambda_{t,0}^{-4}\eta_{t,0}^2(a_4 + a_3X_{t-1}^2)^2$$
$$+ 4u_{t,0}\lambda_{t,0}^{-3}(a_1 + a_2X_{t-1})(a_4 + a_3X_{t-1}^2)\eta_{t,0} \tag{2.6}$$

Par définition $\xi_t(a)$ est strictement stationnaire et ergodique et on verra toute de suite que $E(\xi_t^2(a))$ est finie.

Par le théorème(2.3), $n^{\frac{-1}{2}}\sum_{t=1}^{n}\xi_t(a)$ converge vers une normale de moyenne nulle et de variance $E(\xi_t^2(a))$.

Cette variance peut être exprimée sous forme de $a^{'}Wa$ où W est symétrique et ne dépend pas de a. D'où $\sqrt{n}\frac{\partial\tilde{l}_n(\theta_0)}{\partial\theta}$ converge vers une loi normale de moyenne nulle et de matrice de covariance W.

Il reste à démontrer que :$E(\xi_t^2(a))$ est fini pour tout a.

Pour le premier terme, nous appliquons le lemme (2.2)

$$\begin{aligned} E \mid 4\lambda_{t,0}^{-2}u_{t,0}^2(a_1 + a_2X_{t-1})^2 \mid &= E \mid 4\lambda_{t,0}^{-2}(a_1 + a_2X_{t-1})^2E(u_{t,0}^2 \mid \mathcal{F}_{t-1}) \mid \\ &= 4E \mid \lambda_{t,0}^{-1}(a_1 + a_2X_{t-1})^2 \mid \\ &< \infty \end{aligned}$$

Montrons que le deuxième terme est fini :

Comme $(u_{t,0} \mid \mathcal{F}_{t-1}) \sim \mathcal{N}(0, \lambda_{t,0})$, alors $E(u_{t,0}^4 \mid \mathcal{F}_{t-1}) = 3\lambda_{t,0}^2$

$$\begin{aligned} E(\eta_{t,0}^2 \mid \mathcal{F}_{t-1}) &= E[(u_{t,0}^2 - \lambda_{t,0})^2] \\ &= E(u_{t,0}^4 + \lambda_{t,0}^2 - 2\lambda_{t,0}u_{t,0}^2 \mid \mathcal{F}_{t-1}) \\ &= E(u_{t,0}^4 \mid \mathcal{F}_{t-1}) + \lambda_{t,0}^2 - 2\lambda_{t,0}E(u_{t,0}^2 \mid \mathcal{F}_{t-1}) \\ &= E(u_{t,0}^4 \mid \mathcal{F}_{t-1}) - \lambda_{t,0}^2 \\ &= 2\lambda_{t,0}^2 \end{aligned}$$

D'où :

$$\begin{aligned}
E \mid \lambda_{t,0}^{-4}(a_4 + a_3X_{t-1}^2)^2\eta_{t,0}^2 \mid &= E \mid \lambda_{t,0}^{-4}(a_4 + a_3X_{t-1}^2)^2 E(\eta_{t,0}^2 \mid \mathcal{F}_{t-1}) \mid \\
&= 2 \mid \lambda_{t,0}^{-2}(a_4 + a_3X_{t-1}^2)^2 \mid \\
&= 2 \mid \frac{(a_4 + a_3X_{t-1}^2)^2}{(\sigma_0^2 + \sigma_{\beta,0}^2X_{t-1}^2)^2} \mid \\
&= 2[\frac{a_4 + a_3X_{t-1}^2}{\sigma_0^2 + \sigma_{\beta,0}^2X_{t-1}^2}]^2 \\
&\leq 2(\frac{\mid a_3 \mid}{\sigma_{\beta,0}^2} + \frac{\mid a_4 \mid}{\sigma_0^2} + 1)^2 \\
&< \infty
\end{aligned}$$

Par la même méthode on montre que le 3^{eme} terme est fini.

D'où : $\quad E(\xi_t^2(a)) < \infty$

Ainsi par le théorème (2.3), $\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\xi_t(a)$ converge en distribution vers une variable normale de moyenne nulle et de variance $E(\xi_t^2(a))$. Par conséquent $\sqrt{n}\frac{\partial \tilde{l}_n(\theta_0)}{\partial\theta}$ converge vers une normale de moyenne nulle et de matrice de covariance $W = (W_{ij})$. On a

$$\begin{aligned}
E(\xi_t^2(a)) &= 4E(u_{t,0}^2\lambda_{t,0}^{-2}(a_1 + a_2X_{t-1})^2) + E(\lambda_{t,0}^{-4}\eta_{t,0}^2(a_4) + a_3X_{t-1}^2)^2) \\
&\quad + 4E(u_{t,0}\lambda_{t,0}^{-3}(a_1 + a_2X_{t-1})(a_4 + a_3X_{t-1}^2)\eta_{t,0})
\end{aligned}$$

D'autre part

$$E(a_1Y_1^n + a_2Y_2^n + a_3Y_3^n + a_4Y_4^n)^2 = \sum_{i,j=1}^{4} a_ia_jE(Y_i^nY_j^n)$$

les W_{ij} peuvent être obtenus en identifiant les deux relations de telle sorte que W_{ij} soit le coefficient de a_ia_j dans $E(\xi_t^2(a))$.

D'où :

$$
\begin{aligned}
&W_{11} = 4E(u_{t,0}^2\lambda_{t,0}^{-2}) = E(\lambda_{t,0}^{-1})\\
&W_{12} = W_{21} = 4E(u_{t,0}^2\lambda_{t,0}^{-2}X_{t-1}) = 4E(\lambda_{t,0}^{-1}X_{t-1})\\
&W_{13} = W_{31} = 2E(u_{t,0}\lambda_{t,0}^{-3}X_{t-1}^2\eta_{t,0})\\
&W_{14} = W_{41} = 2E(u_{t,0}\lambda_{t,0}^{-3}\eta_{t,0})\\
&W_{22} = 4E(u_{t,0}^2\lambda_{t,0}^{-2}X_{t-1}^2) = 4E(\lambda_{t,0}^{-1}X_{t-1}^2)\\
&W_{23} = W_{32} = 2E(u_{t,0}\lambda_{t,0}^{-3}X_{t-1}^3\eta_{t,0})\\
&W_{24} = W_{42} = 2E(u_{t,0}\lambda_{t,0}^{-3}X_{t-1}\eta_{t,0})\\
&W_{33} = E(\lambda_{t,0}^{-4}X_{t-1}^4\eta_{t,0}^2)\\
&W_{34} = W_{43} = E(\lambda_{t,0}^{-4}X_{t-1}^2\eta_{t,0}^2)\\
&W_{44} = E(\lambda_{t,0}^{-4}\eta_{t,0}^2)
\end{aligned}
\tag{2.7}
$$

Si β_t et ϵ_t sont conjointement normales, alors $u_{t,0} \mid \mathcal{F}_{t-1} \sim \mathcal{N}(0, \lambda_{t,0})$, d'où

$$
\begin{aligned}
E(u_{t,0}\eta_{t,0} \mid \mathcal{F}_{t-1}) &= E(u_{t,0}^3 - u_{t,0}\lambda_{t,0} \mid \mathcal{F}_{t-1}) = 0\\
E(u_{t,0}^4 \mid \mathcal{F}_{t-1}) &= 3\lambda_{t,0}^2
\end{aligned}
$$

et :

$$
E(\eta_{t,0}^2 \mid \mathcal{F}_{t-1}) = E(u_{t,0}^4 - 2u_{t,0}^2\lambda_{t,0} + \lambda_{t,0}^2 \mid \mathcal{F}_{t-1}) = 2\lambda_{t,0}^2
$$

Par conséquent l'expression de W dans (2.7) devient :

$$W_{11} = 4E(\lambda_{t,0}^{-1})$$

$$W_{12} = W_{21} = 4E(\lambda_{t,0}^{-1} X_{t-1})$$

$$W_{13} = W_{14} = W_{23} = W_{24} = 0$$

$$W_{22} = 4E(\lambda_{t,0}^{-1} X_{t-1}^{2})$$

$$W_{33} = E(\lambda_{t,0}^{-4} X_{t-1}^{4} \eta_{t,0}^{2}) = 2E(\lambda_{t,0}^{-2} X_{t-1}^{4})$$

$$W_{34} = E(\lambda_{t,0}^{-4} X_{t-1}^{2} \eta_{t,0}^{2}) = 2E(\lambda_{t,0}^{-2} X_{t-1}^{2})$$

$$W_{44} = 2E(\lambda_{t,0}^{-2})$$

Nous remarquons que $W = 2V$ sous la condition que ϵ_t et β_t soient conjointement normales.

D'où

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \sim \mathcal{N}(0, 2V^{-1})$$

$\square$

Chapitre 3

Effet de la racine unitaire sur les prévisions et simulation

3.1 Effet de la racine unitaire sur les prévisions

Dans cette section, nous allons considérer l'effet de la racine unitaire pour le modèle RCA(1) en examinant la variance conditionnelle des erreurs de prédiction.

Supposons qu'a l'instant t on a t+1 observations $X_0, X_1, ...X_t$ à partir d'une série de modèles RCA(1) définis dans le premier chapitre .

Supposons que les paramètres de $\theta = (\alpha, \mu_\beta, \sigma_\beta^2, \sigma^2)'$ sont connus.

Le meilleur prédicteur est donnée par :

$$\begin{aligned} \hat{X}_{t+1} &= E(X_{t+1} \mid X_t) \\ &= E(\alpha + \beta_t X_t + \epsilon_{t+1} \mid X_t) \\ &= \alpha + \mu_\beta X_t \end{aligned}$$

La variance conditionnelle est donnée par :

$$\begin{aligned} v_{t+1|t} &= Var(X_{t+1} \mid X_t) \\ &= Var(\alpha + \beta_t X_t + \epsilon_{t+1} \mid X_t) \\ &= \sigma_\beta^2 X_t^2 + \sigma^2 \end{aligned}$$

Proposition 3.1. *Pour $j \geq 2$ la moyenne de l'erreur quadratique a la forme suivante :*
$\hat{X}_{t+j} = \alpha(\frac{1-\mu_\beta^j}{1-\mu_\beta}) + \mu_\beta^j X_t \qquad pour \qquad \mu_\beta < 1.$

Et la variance conditionnelle de l'erreur de prédiction est donnée par :

$v_{t+j|t} = \eta^{j-1} v_{t+1|t} + \sigma^2 \sum_{k=1}^{j-1} \eta^{k-1} + \sigma_\beta^2 \sum_{k=1}^{j-1} \eta^{k-1} \hat{X}_{t+j-k}^2$

Démonstration. 1)Pour $j \geq 2$ on a :

$$\begin{aligned} \hat{X}_{t+j} &= E(X_{t+j} \mid X_t) \\ &= E[E(X_{t+j} \mid X_{t+j-1}) \mid X_t] \\ &= E(\alpha + X_{t+j-1}\mu_\beta \mid X_t) \\ &= \alpha + \mu_\beta \hat{X}_{t+j-1} \end{aligned}$$

D'où :

$$\begin{aligned} \hat{X}_{t+j} &= \alpha + \mu_\beta \hat{X}_{t+j-1} \\ &= \alpha + \alpha\mu_\beta + ... + \alpha\mu_\beta^{j-1} + \mu_\beta^j \hat{X}_t \\ &= \alpha(1 + \mu_\beta + ... + \mu_\beta^{j-1}) + \mu_\beta^j \hat{X}_t \\ &= \alpha[\frac{1-\mu_\beta^j}{1-\mu_\beta}] + \mu_\beta^j X_t \qquad pour \mu_\beta < 1 \end{aligned}$$

2)Pour la j^{eme} étape la variance conditionnelle de l'erreur de prédiction est donnée par :

$$\begin{aligned} v_{t+j|t} &= Var(X_{t+j} \mid X_t) \\ &= E(X_{t+j}^2 \mid X_t) - E^2(X_{t+j} \mid X_t) \\ &= \alpha^2 + (\sigma_\beta^2 + \mu_\beta^2) E(X_{t+j-1}^2 \mid X_t) + \sigma_0^2 + 2\alpha\mu_\beta E(X_{t+j-1} \mid X_t) - (\alpha + \mu_\beta E(X_{t+j-1} \mid X_t))^2 \\ &= (\sigma_\beta^2 + \mu_\beta^2) E(X_{t+j-1}^2 \mid X_t) + \sigma_0^2 - \mu_\beta^2 E^2(X_{t+j-1} \mid X_t) \\ &= \mu_\beta^2 Var(X_{t+j-1} \mid X_t) + \sigma_0^2 + \sigma_\beta^2 E^2(X_{t+j-1} \mid X_t) \\ &= \eta v_{t+j-1|t} + \sigma^2 \hat{X}_{t+j-1}^2 + \sigma^2 \end{aligned}$$

où $\eta = \mu_\beta^2 + \sigma_\beta^2$ est le paramètre de stationnarité. L'évaluation de $v_{t+j|t}$ donne

$$\begin{aligned} v_{t+j|t} &= \eta v_{t+j-1|t} + \sigma_\beta^2 \hat{X}_{t+j-1}^2 + \sigma^2 \\ &= \eta[\eta v_{t+j-2|t} + \sigma_\beta^2 \hat{X}_{t+j-2}^2 + \sigma^2] + \sigma_\beta^2 \hat{X}_{t+j-1} + \sigma^2 \\ &. \\ &. \\ &. \\ &= \eta^{j-1} v_{t+1|t} + \sigma^2 \sum_{k=1}^{j-1} \eta^{k-1} + \sigma_\beta^2 \sum_{k=1}^{j-1} \eta^{k-1} \hat{X}_{t+j-k}^2 \end{aligned}$$

□

Proposition 3.2. *Si le processus RCA(1) a une racine unitaire, c'est à dire $\eta = \mu_\beta^2 + \sigma_\beta^2 = 1$, alors $v_{t+j|t} \longrightarrow \infty$ quand $j \to \infty$, et si le processus est faiblement stationnaire, alors $v_{t+j|t}$ converge.*

Démonstration. Dans le cas où $\eta = 1$, l'équation de $v_{t+j|t}$ devient :

$$v_{t+j|t} = v_{t+1|t} + \sigma^2(j-1) + \sigma_\beta^2 \sum_{k=1}^{j-1} \hat{X}_{t+j-k}^2$$

Il est évident que lorsque j tend vers l'infini, $v_{t+j|t} \longrightarrow \infty$.

Pour le cas où $\eta < 1$, le premier terme tend vers 0, le deuxième terme converge vers $\frac{\sigma^2}{1-\eta}$, et le dernier terme peut s'écrire sous la forme suivante :

$$\begin{aligned}
\sigma_\beta^2 \sum_{k=1}^{j-1} \eta^{k-1} \hat{X}_{t+j-k}^2 &= \sigma_\beta^2 \sum_{k=1}^{j-1} \eta^{k-1} [\alpha \frac{1-\mu_\beta^{j-k}}{1-\mu_\beta} + X_t \mu_\beta^{j-k}]^2 \\
&= \sigma_\beta^2 \sum_{k=1}^{j-1} [\frac{\alpha}{1-\mu_\beta} + (X_t - \frac{\alpha}{1-\mu_\beta}) \mu_\beta^{j-k}]^2 \eta^{k-1} \\
&= \sigma_\beta^2 \sum_{k=1}^{j-1} (a + b\mu_\beta^{j-k})^2 \eta^{k-1} \\
&\leq \sigma_\beta^2 a^2 \sum_{k=1}^{j-1} \eta^{k-1} + 2\sigma_\beta^2 \mid ab \mid \sum_{k=1}^{j-1} \mid \mu_\beta \mid^{j-k} + \sigma_\beta^2 b^2 \sum_{k=1}^{j-1} \mu_\beta^{2(j-k)}
\end{aligned}$$

où $a = \frac{\alpha}{1-\mu_\beta}$ et $b = X_t - \frac{\alpha}{1-\mu_\beta}$

Le premier terme de l'inégalité converge vers $\sigma_\beta^2 a^2 \frac{1}{1-\eta}$, le deuxième et le dernier terme convergent vers 0.

Chaque terme de l'inégalité converge vers un nombre fini quand $j \to \infty$ alors $v_{t+j|t}$ converge quand $\eta < 1$. $\square$

Dans la pratique nous devons estimer le paramètre $\theta = (\alpha, \mu_\beta, \sigma_\beta^2, \sigma^2)'$ au lieu de connaitre sa vraie valeur.

3.2 Test de la racine unitaire pour le processus RCA(1) basé sur la méthode du maximum de vraisemblance

Considérons le modèle RCA(1), on s'intéresse à tester la racine unitaire contre l'hypothèse alternative du modèle qui est faiblement stationnaire.

$$H_0 : \mu_\beta^2 + \sigma_\beta^2 = 1 \text{contre } H_1 : \mu_\beta^2 + \sigma_\beta^2 < 1 \tag{3.1}$$

où $\sigma_\beta^2 \neq 0$. Il a été montré dans la section 2.2.3 que l'estimateur du maximum de vraisemblance est asymptotiquement normale.
Nous pouvons utiliser le critère du test de Wald pour tester l'hypothèse de la racine unitaire.
La distribution limite de la statistique de Wald peut être obtenue en utilisant la méthode delta qui consiste à obtenir une approximation de la distribution asymptotique de la transformée d'une variable aléatoire asymptotiquement normale.

Lemme 3.1. *(méthode delta) Soient $(\hat{\theta}_n)$ une suite de variable aléatoire dans R^p telle que $\sqrt{n}(\hat{\theta}_n - \theta)$ converge en distribution vers $\mathcal{N}(0, \sum)$, et g une fonction de R^p vers R^q qui possède une dérivée continue au voisinage de θ.*
Alors

$$\sqrt{n}(g(\hat{\theta}_n) - g(\theta)) \longrightarrow \mathcal{N}(0, \dot{g}(\theta) \sum \dot{g}^T(\theta))$$

où : $\quad \dot{g}(\theta) = \frac{\partial g(\theta)}{\partial \theta^T}$.

Soit $\theta = (\alpha, \mu_\beta, \sigma_\beta^2, \sigma^2)'$ le paramètre du processus RCA(1), $\hat{\theta} = (\hat{\alpha}, \hat{\mu}_\beta, \hat{\sigma}_\beta^2, \hat{\sigma}^2)'$ l'estimateur du maximum de vraisemblance pour θ basé sur un échantillon fini $X_0, X_1, ..., X_n$, et $\theta_0 = (\alpha_0, \mu_{\beta,0}, \sigma_{\beta,0}^2, \sigma_0^2)$ la vraie valeur de θ.
Il a été montré que sous les conditions A1-A6 le processus RCA(1) est strictement stationnaire, ergodique, et $\sqrt{n}(\hat{\theta} - \theta_0) \sim \mathcal{N}(0, V^{-1}WV^{-1})$. Dans cette section nous allons supposer la normalité conjointe de $\{\beta_t\}$ et $\{\epsilon_t\}$. Dans ce cas, la matrice de covariance asymptotique des estimateurs du maximum de vraisemblance se réduit à $2V^{-1}$, où V est donné dans l'équation (2.6).
On définit $\eta = g(\theta) = \mu_\beta^2 + \sigma_\beta^2$, alors $\frac{\partial g(\theta)}{\partial \theta^T} = (0, 2\mu_\beta, 1, 0)$.

Soit $\hat{\eta} = g(\hat{\theta}) = \hat{\mu}_\beta^2 + \hat{\sigma}_\beta^2$ l'estimateur du maximum de vraisemblance pour le paramètre stationnaire η du modèle RCA(1). En appliquant la méthode delta nous avons :

$$\begin{aligned}\hat{\eta} &= g(\hat{\theta}) \\ &\sim \mathcal{N}(g(\theta_0), 2n^{-1}g^{\cdot}(\theta_0)V^{-1}(\theta_0)g^{\cdot^T}(\theta_0))\end{aligned} \tag{3.2}$$

Rappelons que la statistique de Wald peut être utilisée pour tester la vraie valeur du paramètre, elle est basée sur la propriété de normalité asymptotique de l'estimation du maximum de vraisemblance et se calcule comme suit :

$$T_w = \frac{g(\hat{\theta}) - 1}{\sqrt{2n^{-1}g^{\cdot}(\hat{\theta})\hat{V}^{-1}(\hat{\theta})g^{\cdot}(\hat{\theta}^T)}} \tag{3.3}$$

où $\hat{V}(\hat{\theta})$ est l'estimateur du moment de $V(\theta_0)$. Chaque élément de $\hat{V}(\hat{\theta})$ est obtenu en remplaçant l'espérance par son moment correspondant.
Par exemple $\hat{V}_{11}(\hat{\theta})$ a la forme ,

$$\hat{V}_{11}(\hat{\theta}) = 2n^{-1}\sum_{t=1}^{n}\lambda_{t,n}^{-1} = 2n^{-1}\sum_{t=1}^{n}(\hat{\sigma}^2 + \hat{\sigma}_\beta^2 X_{t-1}^2)^{-1}$$

Par le théorème ergodique, on a $\hat{V}(\hat{\theta})$ converge presque surement vers $V(\theta_0)$, donc par le théorème de Slutsky, sous l'hypothèse H_0, $T_w \longrightarrow Z$ en loi, quand $n \longrightarrow \infty$, où Z est une loi normale.

3.3 Simulation

Notre étude de simulation pour évaluer la performance de la procédure du maximum de vraisemblance pour le modèle RCA (1) est divisée en trois parties [5].
La première partie se résume à l'étude de l'estimation des paramètres et évaluation de leurs performances basées sur la convergence en probabilité.

La deuxième partie se concentre sur l'évaluation des propriétés asymptotiques des estimateurs du maximum de vraisemblance (décrites en section 2.2).

La troisième partie est destinée à l'étude de la fiabilité du test de racine unitaire sur la base de l'estimateur du maximum de vraisemblance et montre que ce test simple fonctionne raisonnablement bien.

Nous générons les données $\{X_t : t = 1, ..., n\}$ pour le modèle RCA(1) en commençant par choisir la valeur de X_0

Pour générer des échantillons à partir du modèle RCA(1) avec une racine unitaire, nous choisissons $X_0 = 0$, et $X_1 \sim \mathcal{N}(\alpha, \sigma^2)$ en supposant que β_t et ϵ_t sont conjointement normales.

Pour générer un échantillon aléatoire à partir d'un processus faiblement stationnaire, nous avons besoin de commencer par générer X_0 à partir d'une distribution appropriée.

Théorème 3.1. *Soit $\{X_t : t = 0, 1, 2, ...\}$ une suite infinie, si pour X_1 on a $E(X_1) = \frac{\alpha}{1-\mu_\beta}$ et $Var(X_1) = \frac{\alpha^2\sigma_\beta^2}{[(1-\mu_\beta)^2(1-\eta)]} + \frac{\sigma^2}{1-\eta}$, où $\eta = \mu_\beta^2 + \sigma_\beta^2 < 1$, et X_1 est indépendant de la suite (β_t) et (ϵ_t), alors le processus généré par l'équation (1.1) est faiblement stationnaire.*

Démonstration. Il est facile de voir que :

$$X_t = w_{t,t-2}X_1 + \alpha \sum_{j=-1}^{t-3} w_{t,j} + \sum_{j=-1}^{t-3} w_{t,j}\epsilon_{t-j-1}$$

où : $\quad w_{t,j} = \beta_t * \beta_{t-1} * ... * \beta_{t-j} \quad$ pour $j = 0, 1, 2, ...$ et $w_{t,-1} = 1$.

Nous avons besoin de vérifier les deux premières conditions :

Comme $E(X_1) = \frac{\alpha}{1-\mu_\beta}$ et $E(w_{t,t-2}) = \mu_\beta^{t-1}$

alors

$$
\begin{aligned}
E(X_t) &= E(w_{t,t-2}X_1) + \alpha \sum_{j=-1}^{t-3} E(w_{t,j}) + \sum_{j=-1}^{t-3} E(w_{t,j}\epsilon_{t-j-1}) \\
&= E(X_1)E(w_{t,t-2}) + \alpha \sum_{j=-1}^{t-3} E(w_{t,j}) \\
&= \frac{\alpha}{1-\mu_\beta}\mu_\beta^{t-1} + \alpha \sum_{j=-1}^{t-3} \mu_\beta^{j-1} \\
&= \frac{\alpha}{1-\mu_\beta}\mu_\beta^{t-1} + \alpha \frac{1-\mu_\beta^{t-1}}{1-\mu_\beta} \\
&= \frac{\alpha}{1-\mu_\beta}
\end{aligned}
$$

Vérifions la deuxième condition :

On a : $X_2 = \alpha + \beta_2 X_1 + \epsilon_2$

comme :$E(\beta_2^2) = \eta$ et :

$$
\begin{aligned}
E(X_1^2) &= Var(X_1) + E^2(X_1) \\
&= \alpha^2[\frac{\sigma_\beta^2}{(1-\mu_\beta)^2(1-\eta)} + \frac{1}{(1-\mu_\beta)^2}] + \frac{\sigma^2}{1-\eta} \\
&= \alpha^2[\frac{1+\mu_\beta}{(1-\mu_\beta)(1-\eta)}] + \frac{\sigma^2}{1-\eta}
\end{aligned}
$$

Par conséquent :

$$
\begin{aligned}
Var(X_2) &= Var(\alpha + \beta_2 X_1 + \epsilon_2) \\
&= Var(\beta_2 X_1) + Var(\epsilon_2) \\
&= E(\beta_2^2)E(X_1^2) - E^2(\beta_2)E^2(X_1) + Var(\epsilon_2) \\
&= \eta[\alpha^2 \frac{1+\mu_\beta}{(1-\mu_\beta)(1-\eta)} + \frac{\sigma^2}{1-\eta}] - \mu_\beta^2 \frac{\alpha^2}{(1-\mu_\beta)^2} + \sigma^2 \\
&= \frac{\alpha^2\sigma_\beta^2}{(1-\mu_\beta)^2(1-\eta)} + \frac{\sigma^2}{1-\eta}
\end{aligned}
$$

Par induction, pour t=1,2,... on a :

$$
Var(X_t) = \frac{\alpha^2\sigma_\beta^2}{(1-\mu_\beta)^2(1-\eta)} + \frac{\sigma^2}{1-\eta}
$$

Montrons par récurrence que cette formule est vraie pour tout $t > 1$.

Supposons qu'elle est vraie à l'ordre t et montrons pour t+1.

On a :

$$X_{t+1} = \alpha + \beta_{t+1} X_t + \epsilon_{t+1}$$

D'où :

$$\begin{aligned}
Var(X_{t+1}) &= Var(\beta_{t+1} X_t) + Var(\epsilon_t) \\
&= E(\beta_{t+1}^2)E(X_t^2) - E^2(\beta_{t+1})E^2(X_t) + Var(\epsilon_t) \\
&= \eta[\frac{\alpha^2\sigma_\beta^2}{(1-\mu_\beta)^2(1-\eta)} + \frac{\sigma^2}{(1-\eta)} + \frac{\alpha^2}{1-\mu_\beta)^2}] - \mu_\beta^2 \frac{\alpha^2}{(1-\mu_\beta)^2} + \sigma^2 \\
&= \frac{\alpha^2\sigma_\beta^2}{(1-\mu_\beta)^2(1-\eta)} + \frac{\sigma^2}{1-\eta}
\end{aligned}$$

Maintenant,

$$\begin{aligned}
E(X_t X_{t+h}) &= E(E(X_t X_{t+h} \mid X_t)) \\
&= E[X_t E(X_{t+h} \mid X_t)] \\
&= E(X_t E(\alpha \sum_{j=-1}^{h-2} w_{t+h,j} + \sum_{j=-1}^{h-2} w_{t+h,j}\epsilon_{t+h-j-1} + w_{t+h,h-1}X_t \mid X_t)] \\
&= \alpha \frac{1-\mu_\beta^h}{1-\mu_\beta} E(X_t) + \mu_\beta^h E(X_t^2) \\
&= \alpha^2 \frac{1-\mu_\beta^h}{(1-\mu_\beta)^2} + \mu_\beta^h[\frac{\alpha^2\sigma_\beta^2}{(1-\mu_\beta)^2(1-\eta)} + \frac{\sigma^2}{1-\eta} + \frac{\alpha^2}{(1-\mu_\beta)^2}]
\end{aligned}$$

D'où :

$$\begin{aligned}
Cov(X_t, X_{t+h}) &= E(X_t X_{t+h}) - E^2(X_t) \\
&= \mu_\beta^h[\frac{\alpha^2\sigma_\beta^2 + \sigma^2(1-\mu_\beta)^2}{(1-\mu_\beta)^2(1-\eta)}]
\end{aligned}$$

Puisque la covariance dépend de h, le processus est faiblement stationnaire. □

Dans notre simulation, nous avons supposé que les β_t et ϵ_t sont conjointement normales. Dans ce cas, nous générons une série de modèles RCA(1) en commençant par $X_0 = 0$ où X_0 suit une distribution normale de moyenne et de variance appropriées données dans le théorème(3.1). Nous avons fixé $\alpha = 0$, $\sigma^2 = 1$ et nous prenons des valeurs différentes pour μ_β et σ^2_β.

Dans l'étude de l'estimation des paramètres basée sur la procédure du maximum de vraisemblance, nous générons les séries de modèles RCA(1) avec différentes valeurs des paramètres :

1. $\alpha = 0, \sigma^2 = 1, \mu_\beta = 0.5, \sigma^2_\beta = 0.25$
2. $\alpha = 0, \sigma^2 = 1, \mu_\beta = 0.6, \sigma^2_\beta = 0.64$
3. $\alpha = 0, \sigma^2 = 1, \mu_\beta = 0.995, \sigma^2_\beta = 0.01$
4. $\alpha = 0, \sigma^2 = 1, \mu_\beta = 0.1, \sigma^2_\beta = 0.99$
5. $\alpha = 0, \sigma^2 = 1, \mu_\beta = -0.6, \sigma^2_\beta = 0.64$
6. $\alpha = 0, \sigma^2 = 1, \mu_\beta = -0.995, \sigma^2_\beta = 0.01$
7. $\alpha = 0, \sigma^2 = 1, \mu_\beta = -0.1, \sigma^2_\beta = 0.99$
8. $\alpha = 0, \sigma^2 = 1, \mu_\beta = 1, \sigma^2_\beta = 0$

Il est à noter que le premier cas est faiblement stationnaire, les cas 2, 4, 5 et 7 ont une racine unitaire, les cas 3 et 6 sont strictement stationnaires, et le cas 8 est une marche aléatoire.

Pour l'estimation des paramètres du modèle RCA(1), nous avons utilisé le logiciel (OpenBUGS V3.2.3 Rev1012). On prend la taille de l'échantillon n = 100 et 500, et répété le modèle 500 fois. La deuxième et la troisième colonne du tableau (3.1)et (3.2) représentent la moyenne et l'écart type de l'estimateur du maximum de vraisemblance.

TABLE 3.1 – Estimations des paramètres du modèle RCA(1) par la méthode du maximum de vraissemblance avec n=100

True	mean	sd
$\alpha = 0$	-0.0185	0.1337
$\mu_\beta = 0.5$	0.5012	0.2776
$\sigma_\beta^2 = 0.25$	0.2524	0.1619
$\sigma^2 = 1$	0.5848	0.08044
$\alpha = 0$	0.2209	0.2533
$\mu_\beta = 0.6$	0.5011	0.2777
$\sigma_\beta^2 = 0.64$	0.2529	0.1633
$\sigma^2 = 1$	0.1651	0.0227
$\alpha = 0$	-0.4522	0.3025
$\mu_\beta = 0.995$	0.5111	0.2762
$\sigma_\beta^2 = 0.01$	0.2508	0.163
$\sigma^2 = 1$	0.2069	0.0285
$\alpha = 0$	-0.09424	0.1677
$\mu_\beta = 0.1$	0.5001	0.2769
$\sigma_\beta^2 = 0.99$	0.2523	0.1622
$\sigma^2 = 1$	0.3813	0.05295
$\alpha = 0$	-0.1801	0.1695
$\mu_\beta = -0.6$	0.4891	0.2784
$\sigma_\beta^2 = 0.64$	0.24	0.1568
$\sigma^2 = 1$	0.3977	0.05513
$\alpha = 0$	-0.09025	0.127
$\mu_\beta = -0.995$	0.4558	0.2665
$\sigma_\beta^2 = 0.01$	0.2326	0.1489
$\sigma^2 = 1$	0.6639	0.09216
$\alpha = 0$	-0.2895	0.2287
$\mu_\beta = -0.1$	0.4989	0.2776
$\sigma_\beta^2 = 0.99$	0.2504	0.1611
$\sigma^2 = 1$	0.2089	0.02902
$\alpha = 0$	-0.02104	0.1073
$\mu_\beta = 1$	0.5164	0.2739
$\sigma_\beta^2 = 0$	0.2534	0.1626
$\sigma^2 = 1$	0.9219	0.127

TABLE 3.2 – Estimations des paramètres du modèle RCA(1) par la méthode du maximum de vraissemblance avec n=500

True	mean	sd
$\alpha = 0$	-0.02754	0.05728
$\mu_\beta = 0.5$	0.5003	0.282
$\sigma_\beta^2 = 0.25$	0.2495	0.1634
$\sigma^2 = 1$	0.6229	0.03775
$\alpha = 0$	0.05955	0.1143
$\mu_\beta = 0.6$	0.4991	0.2808
$\sigma_\beta^2 = 0.64$	0.2475	0.1619
$\sigma^2 = 1$	0.1588	0.009618
$\alpha = 0$	-0.04871	0.09369
$\mu_\beta = 0.995$	0.5122	0.2771
$\sigma_\beta^2 = 0.01$	0.2469	0.1601
$\sigma^2 = 1$	0.2391	0.01445
$\alpha = 0$	0.01513	0.2194
$\mu_\beta = 0.1$	0.4938	0.2834
$\sigma_\beta^2 = 0.99$	0.2432	0.16
$\sigma^2 = 1$	0.04403	0.002695
$\alpha = 0$	-0.008592	0.1185
$\mu_\beta = -0.6$	0.4598	0.2792
$\sigma_\beta^2 = 0.64$	0.2391	0.1601
$\sigma^2 = 1$	0.1528	0.00931
$\alpha = 0$	-0.08823	0.0808
$\mu_\beta = -0.995$	0.4546	0.266
$\sigma_\beta^2 = 0.01$	0.2308	0.1446
$\sigma^2 = 1$	0.3142	0.01945
$\alpha = 0$	-0.05791	0.1778
$\mu_\beta = -0.1$	0.488	0.2835
$\sigma_\beta^2 = 0.99$	0.24	0.1582
$\sigma^2 = 1$	0.06795	0.004204
$\alpha = 0$	0.09529	0.0661
$\mu_\beta = 1$	0.5149	0.2767
$\sigma_\beta^2 = 0$	0.2468	0.1594
$\sigma^2 = 1$	0.9705	0.05915

Bibliographie

[1] Patrick Billingsley. The lindeberg-levy theorem for martingales. *Proceedings of the American Mathematical Society*, 12(5) :788–792, 1961.

[2] Des F Nicholls and Barry G Quinn. *Random Coefficient Autoregressive Models : An Introduction : An Introduction*, volume 11. Springer Science & Business Media, 2012.

[3] BG Quinn. A note on the existence of strictly stationary solutions to bilinear equations. *Journal of Time Series Analysis*, 3(4) :249–252, 1982.

[4] Anton Schick. $\sqrt{n}$-consistent estimation in a random coefficient autoregressive model. *Australian Journal of Statistics*, 38(2) :155–160, 1996.

[5] Dazhe Wang. Frequentist and bayesian analysis of random coefficient autoregressive models. 2004.