Remerciments

Je remercie Dieu tout puissant de m'avoir donné les forces qui m'ont permis de dépasser toutes les difficultés pour aboutir enfin à la réalisation de ce modeste travail.

Je tiens en particulier à remercier monsieur Tahar MOURID de m'avoir soutenu en acceptant d'être mon président du jury.

Je remercie aussi mon encadreur monsieur LABBAS Ahmed, ainsi que l'examinateur monsieur Abdelaziz.ALLAM .

Table des matières

Introduction

Les processus moyenne mobile (MM) abordés dans ce mémoire trouvent leurs applications dans plusieurs domaines: économie, biologie, medecine ect...Un processus X_t moyenne mobile d'ordre un $MM(1)$ est défini par

$$X_t = \varepsilon_t + \beta\varepsilon_{t-1} \quad (t \in \mathbb{Z})$$

où (ε_t) est une suite de variables aléatoires indépendantes et identiquement distribuées centrées et de variance σ_ε^2 (bruit blanc) et β est une constante en module inférieure à un.

Parmis les problèmes abordés dans cette classe de processus nous relevons l'importance de l'estimation des paramètres, en particulier le paramètre β définissant le modèle MM .

Dans ce mémoire nous développons les résultats traitants l'estimation des paramètres dans un modèle MM par la méthode du RIV (Random initial values).

Nous développons les résultats établis dans les articles

"A new preliminary estimator for $MM(1)$ models "Anna Clara Monti in Computationnal Statistics Data Analysis 21(1996).

"A new preliminary estimator for $MM(1)$ models "Anna Clara Monti in Journal of time series analysis, vol 19, No 2 (1998).

Dans le chapitre 1, nous rappellons les définitions et propriétés générales des processus moyenne mobile de dimensions finis.

Dans le chapitre 2, nous citons d'abord les théorèmes de convergence

(T.C.L) pour les variables indépendantes, et les variables dépendantes ainsi que leurs applications pour les processus moyenne mobile d'ordre 1. Nous étudions ensuite les différentes méthodes d'estimations : La méthode des moments, la méthode des moindres carrés et la méthode du maximum de vraissemblace dans le cas conditionnel et non conditionnel, pour un $MM(1)$.

Dans le chapitre 3 et 4 nous développons les calculs dans [1] et [2]. les deux articles de C.Monti, traitent de la construction et des propriétés asymptotiques (convergence en proba et convergence en loi) de l'estimateur RIV (Random Initial Values) d'une MM dans les cas scalaire et multivarié.

Chapitre 1

Généralités sur les processus moyenne mobile de dimension finie

Un processus stochastique $(X_t)_{t\in T}$ est une famille de variables aléatoires réelles indéxée par T, où T un ensemble non vide

Lorsque T est fini $(X_t)_{t\in T}$ est un vecteur aléatoire.

Lorsque $T\subsetneq R^d$ $(d\geq 2)$ on dit que $(X_t)_{t\in T}$ est un champs aléatoire.

Pour t fixé t$\rightarrow X_t(\omega)$est une variable aléatoire réelle.

Pour ω fixé t$\rightarrow X_t(\omega)$ est une trajéctoire.

1.1 Processus moyenne mobile réel

1.1.1 Définitions

Tout d'abord, on définit les processus bruit blanc réel, moyenne mobile et son dual, le modèle autorégréssif.

Définition 1.1.1. Un processus $\varepsilon = (\varepsilon_t, t\in \mathbb{Z})$est un bruit blanc réel, si c'est une suite de variables aléatoires réelles de moyenne nulle et de variance σ^2_ε, non corrélées, ceci indépendamment de l'instant t, noté $BBR(0,\sigma_\varepsilon^2)$. Si, de plus, les variables aléatoires sont indépendantes et identiquement distribuées, on dira que ε est un bruit blanc réel fort, noté $IID(0,\sigma_\varepsilon^2)$.

Remarque 1.1.2. Un bruit blanc multivarié est une suite $\varepsilon = (\varepsilon_t, t \in \mathbb{Z})$ non corrélée si $t \neq t'$ de vecteurs aléatoires ε_t de moyenne nulle, de même variance notée Σ, noté $BBM(0,\Sigma)$. De ce fait, ε est centré et a pour fonction d'autocovariance

$$\Gamma(h) = \begin{cases} \Sigma \ si & h = 0 \\ 0 & sinon \end{cases}$$

Si de plus, les vecteurs aléatoires ε_t sont indépendants et identiquement distribués, nous écrirons ε est $IID(0,\Sigma)$. Σ est une matrice symétrique définie positive.

Définition 1.1.3. On dit qu'un processus est stationnaire si ses moments ne varient pas au cours du temps.

La suite $X = (X_t,\ t \in \mathbb{Z})$ de vecteurs aléatoires X_t dont la moyenne et la matrice de covariance sont définies par (1.1.5) et (1.1.7), est stationnaire si μ_t et $\Gamma(t+h,t)$, $h = 0, 1, ...$ne depend pas de t .

Sous cette hypothèse, nous notons $E(X_t) = \mu$, où X un vecteur et

$$\begin{aligned} \Gamma(h) : &= E[(X_t - \mu)(X_{t+h} - \mu)'] \\ &= \begin{pmatrix} \gamma_{11}(h) & \gamma_{12}(h) & . & . & . & \gamma_{1m}(h) \\ \gamma_{21}(h) & \gamma_{22}(h) & . & . & . & \gamma_{2m}(h) \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ \gamma_{m1}(h) & \gamma_{m2}(h) & . & . & . & \gamma_{mm}(h) \end{pmatrix} \end{aligned}$$

où $\gamma_{ij}(h) = E(X_{it} - \mu_i)(X_{j,t+h} - \mu_j)$ pour $t, h \in \mathbb{Z}$, et $i, j = 1, 2, ..., m$. Les éléments $\gamma_{ii}(h)$ de la diagonale de $\Gamma(h)$ sont les autocovariances des $X_{it}, i = 1, ..., m$, et les éléments $\gamma_{ij}(h)$ en dehors de la diagonale de $\Gamma(h)$ indiquent la covariance-croisée entre X_{it} et X_{jt}.

Remarque 1.1.4. La matrice $\Gamma(0)$ est la matrice d'autocovariance du processus.

Définition 1.1.5. Un processus X=(X_t, $t\in \mathbb{Z}$) moyenne mobile réelle d'ordre q, noté $MM(q)$, est défini par

$$X_t = \mu_t + \varepsilon_t + \theta_1\varepsilon_{t-1} + ... + \theta_q\varepsilon_{t-q} \tag{1.1.1}$$

où $\varepsilon = (\varepsilon_t,t\in Z)$ est un $BBR(0,\sigma_\varepsilon^2)$, et $\forall\ j = 1, ..., q$, on a $\theta_j\in \mathbb{R}$ avec $\theta_q\neq 0$.

Remarque 1.1.6. Un processus $X = (X_t,\ t\in \mathbb{Z})$ centré a une représentation moyenne mobile multivariée d'ordre (q), $MMM(q)$, s'il s'écrit

$$X_t = \sum_{s=0}^{q} L_s\varepsilon_{t-s} \tag{1.1.2}$$

où

-les L_s sont des matrices de taille $m\times m$ des coefficients moyennes mobiles,

-$L_0 = I_m$ avec I_m la matrice identité de R^m,

-ε_t est un $BBM(0, \Sigma)$.

Une notation équivalente à (1.1.2) est d'utiliser l'operateur de retard $B^s\varepsilon_t = \varepsilon_{t-s}$ pour écrire

$$X_t = L(B)\varepsilon_t$$

où $L(B) = \sum_{s=0}^{q} L_sB^s$, avec $L_s = [l_{ij,s}]$ pour $i, j = 1, ..., m$.

Pour que le processus soit stationnaire, il faut que les coefficients de la matrice L_s soient de carré sommable, Pour qu'un $MMM(q)$ stationnaire soit inversible, L doit verifier $det(L(B)) \neq 0$ pour $|B| \leq 1$.

Nous définissons la représentation duale du processus moyenne mobile, le processus autorégressif.

Définition 1.1.7. Un processus $X = (X_t, t \in \mathbb{Z})$ autorégressif réel d'ordre p, noté $ARR(p)$, est défini par

$$X_t = \mu_t + \psi_1 X_{t-1} + ... + \psi_p X_{t-p} + \varepsilon_t \tag{1.1.3}$$

où $\varepsilon = (\varepsilon_t, t \in \mathbb{Z})$ est un $BBR(0, \sigma_\epsilon^2)$, et $\forall\ j = 1, ..., p$, on a $\psi_j \in \mathbb{R}$ avec $\psi_p \neq 0$.

Les deux processus définis précédemment peuvent s'écrire sous une autre forme.

Si B est l'opérateur retard tel que $B^j(X_n) = X_{n-j}$ pour tout j=1,2,..., alors on note pour tout $t \in \mathbb{Z}$:

Un $MMR(q)$ par

$X_t = \ominus(B)\varepsilon_t$ où $\ominus(B) = 1 + \theta_1 B + \theta_2 B^2 + ... + \theta_q B^q$

Un $ARR(P)$ par

$\psi(B)X_t = \varepsilon_t$ où $\psi(B) = 1 - \psi_1 B - \psi_2 B^2 - ... \psi_p B^p$.

Remarque 1.1.8. Un processus $X = (X_t, t \in \mathbb{Z})$ est centré et admet une représentation autorégressive multivariée d'ordre (p), $ARM(p)$, s'il vérifie

$$X_t = \sum_{s=1}^{p} A_s X_{t-s} + \varepsilon_t \tag{1.1.4}$$

Où

-Les A_s sont des matrices de taille $m \times m$ des coefficients autoregréssifs,

-ε_t est un $BBM(0, \Sigma)$.

En d'autre terme, la relation (1.1.4) peut s'écrire

$$A(X_t) = \varepsilon_t$$

où $A(B) = I_m - \sum_{s=1}^{p} A_s B^s$, avec $A_s = [a_{ij,s}]$ pour $i, j = 1, ..., m$.

Sous certaines conditions, le processus est inversible.

Maintenant nous allons étudier les premiers moments d'un processus .

Définition 1.1.9. La moyenne du processus $X = (X_t,\ t \in \mathbb{Z})$ est définie par μ_t=$E(X_t)$

On dit que le processus X est centré, si μ_t=μ= 0, $\forall t \in \mathbb{Z}$.

Remarque 1.1.10. La moyenne de X_t, notée μ_t est le vecteur de $\mathbb{R}^m$ donné par

$$E(X_t) = \begin{pmatrix} E(X_{1t}) \\ E(X_{2t}) \\ . \\ . \\ . \\ E(X_{mt}) \end{pmatrix} = \begin{pmatrix} \mu_{1t} \\ \mu_{2t} \\ . \\ . \\ . \\ \mu_{mt} \end{pmatrix} = \mu_t. \tag{1.1.5}$$

Définition 1.1.11. la fonction d'autocovariance du processus $X = (X_t, t \in \mathbb{Z})$ est définie par

$$\gamma(s,t) = cov(X_s, X_t) = E[(X_s - E(X_s))(X_t - E(X_t))] \forall s,\ t \in \mathbb{Z}. \tag{1.1.6}$$

Remarque 1.1.12. La matrice de covariance de X_t de retard h, notée $\Gamma(h)$ stationnaire, vérifie

$$\Gamma(t+h,t) = cov(X_{t+h}, X_t) = E[(X_{t+h} - \mu_{t+h})(X_t - \mu_t)']. \tag{1.1.7}$$

Définition 1.1.13. Un processus $X = (X_t,\ t \in \mathbb{Z})$ est du second ordre si

$$E\left|X_t\right|^2 < \infty,$$

de plus il est stationnaire si pour tout $s,\ t,\ h \in \mathbb{Z}$

$$E(X_t) = \mu \text{ et } cov(X_s, X_t) = cov(X_{s+h}, X_{t+h}).$$

Définition 1.1.14. Un processus $MMR(q)$ défini, pour tout $t \in \mathbb{Z}$, par l'équation $X_t = \ominus(B)\varepsilon_t$, est inversible s'il existe une suite de valeurs réelles $\{a_j\}$ telles que $\sum |a_j| < \infty$ et

$$\varepsilon_t = \sum a_j X_{t-j}.$$

Définition 1.1.15. Une suite strictement stationnaire (X_n) de variables aléatoires de $\mathbb{R}$ est dite m-dépendante, avec $m \in \mathbb{N}^*$, si pour tous les ensembles de variables $\{X_j,\ j \leq t\}$ et $\{X_j,\ j \geq t+m+1\}$ sont indépendants.

Remarque 1.1.16. La m-dépendance se traduit par l'indépendance des observations dés qu'elles sont séparées de $m+1$ unités de temps.

Un $MMR(q)$ est un processus q-dépendant.

Nous complétons maintenant l'étude des moments d'un processus en détaillant les fonctions d'autocovariances, d'autocorrélation et d'autocorrélation partielle, puis en donnant quelques unes de leurs propriétés.

Définition 1.1.17. La fonction d'autocorrélation est pour un processus stationnaire définie à partir de la fonction d'autocovariance par

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)}.$$

Remarque 1.1.18. La matrice de corrélation de retard h de X_t est définie par

$$R(h) = \Delta^{-\frac{1}{2}}\Gamma(h)\Delta^{-\frac{1}{2}} = [\rho_{ij}(h)]$$

Pour $i,j = 1,2,...,m$ avec $\Delta = diag(\gamma_{11}(0), \gamma_{22}(0), ..., \gamma_{mm}(0)$ et $R = [\rho_{ij}(h)]_{i,j=1,2...,m}$ où $\rho_{ii}(h)$ est la corrélation de X_{it}, et ρ_{ij} est la corrélation croisée entre X_{it} et $X_{j,t+h}$ vérifiant

$$\rho_{ij}(h) = \frac{\gamma_{ij}(h)}{[\gamma_{ii}(0)\gamma_{jj}(0)]^{\frac{1}{2}}}.$$

Si μ est le vecteur nul de $\mathbb{R}^m$, alors le processus est centré.

Nous supposons dorénavant que X est un processus m-vectoriel stationnaire, centré. Ainsi ses matrices de covariance et de corrélation vérifient :

i)$\Gamma(h) = \Gamma^t(-h)$ et donc $R(h) = R'(-h)$;

ii)$|\gamma_{i,j}(h)| \leq [\gamma_{ii}(0)\gamma_{jj}(0)]^{\frac{1}{2}}, i, j = 1, .., m;$

iii)$\Gamma(.)$ est une fonction définie positive, c'est-à-dire que pour tout point $t_1, ..., t_n$ et tout vecteur $(\alpha_1, ..., \alpha_n)$ de $\mathbb{R}^m$,

on a

$$\sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i^t\Gamma(t_i - t_j)\alpha_j \geq 0;$$

iv)$\rho_{ii}(0) = 1$ pour $i = 1, ..., n$.

Nous résumons les notions vues dans cette sous-section pour un processus $MMR(1)$.

Etude des moments et propriétés d'un $MMR(1)$

De la relation (1.1.1), un $MMR(1)$ centré est défini par

$$X_t = \varepsilon_t + \theta\varepsilon_{t-1}, \quad t \in \mathbb{Z}. \tag{1.1.8}$$

Où ε_t est un $IID(0,\sigma_\epsilon^2)$, $\theta \in \mathbb{R}$ tel que —$\theta| < 1$. Cette dernière condition assure l'inversibilité du processus.

Sa fonction d'autocovariance vérifiant

$$\gamma(h) = \begin{cases} (1+\theta^2)\sigma_\varepsilon^2 & si \quad h = 0 \\ \theta\sigma_\varepsilon^2 & si \quad |h| = 1 \\ 0 & si \quad |h| > 1 \end{cases} . \tag{1.1.9}$$

On peut en déduire sa fonction d'autocorrélation par

$$\rho(h) = \begin{cases} 1 & si \quad h = 0 \\ \frac{\theta}{(1+\theta^2)} & si \quad |h| = 1 \\ 0 & si \quad h > 1 \end{cases} . \tag{1.1.10}$$

Par définition d'un $MMR(1)$, sa stationnarité est entièrement établie.

Et pour l'inversibilité du processus, sous l'hypothèse $|\theta| < 1$, on peut écrire $(1 + \theta B)^{-1} X_t = \varepsilon_t$

et comme

$$(1 + \theta B)^{-1} = \sum \theta^j B^j.$$

On peut écrire $\varepsilon_t = \sum \theta^j X_{t-j}$ et en déduire que le processus $MMR(1)$ est inversible.

Le processus $MMR(1)$ est 1-dépendant. En effet, pour tout $t \in \mathbb{Z}$, X_t et X_{t+1} sont des variables aléatoires dépendantes, alors que pour tout $j \geq t+2$, X_t et X_j sont indépendantes.

Remarque 1.1.19. Si on a 2 processus : $X_t^1 = \varepsilon_t + \theta\varepsilon_{t-1}$ et $X_t^2 = \varepsilon_t + \frac{1}{\theta}\varepsilon_{t-1}$, ils auront la même autocorrélation. Cepandant, si la racine de $(1 - \theta B)$ n'est pas à l'intérieur du cercle unité, alors la racine de $(1 - \frac{1}{\theta}B)$ y est. Parmis X_t^1 et X_t^2 un seul processus est inversible.

1.2 Estimation des moments d'un processus stationnaire

Nous donnons maintenant une définition concernant le comportement asymptotique d'une suite de variables aléatoires.

Définition 1.2.1. Une suite de variables aléatoires $X = (X_n, n \in \mathbb{Z})$ est asymptotiquement normale de moyenne μ_n et d'écart-type σ_n, si $\sigma_n > 0$ pour

n grand, et

$$\frac{X_n - \mu_n}{\sigma_n} \xrightarrow{L} Z, \qquad \text{avec } Z \sim N(0,1),$$

où $N(0,1)$ indique une loi Gaussienne de moyenne nulle et d'écart-type 1.

Chapitre 2

Méthodes d'estimation

2.1 Théorie asymptotique pour les processus moyenne mobile

Pour évaluer le comportement asymptotique des séries tomporelles, nous avons besoin de connaitre la distribution de statistique (telles que la moyenne, la fonction d'autocovariance...) à partir des données dont nous disposons. Cependant, même pour un nombre fini d'observations la distribution exacte n'est pas toujours facile à déterminer. Dans ces cas-là, nous nous basons sur l'inférence statistique obtenue à partir des grands échantillons pour obtenir les résultats.

Théorèmes limites centraux et résultats associés

Nous commençons par énoncer le théorème de la limite centrale

Théorème 2.1.1. *Soit* $X_n \sim IID(\mu, \sigma^2)$ *et* $\bar{X_n} = (X_1 + X_2 + ... + X_n)n^{-1}$*, alors*

$$\frac{\bar{X_n} - \bar{\mu}}{\sigma^2} \sim N(0,1)$$

Ce théorème s'applique aux variables indépendantes, or nous avons vu que le modèle $MMR(1)$ est un processus 1-dépendant. Nous énonçons donc des résultats qui étendent le théorème de la limite centrale aux processus m-dépendants, dùs à Hoeffding et Robbins (1948).

Théorème 2.1.2. *Soit $X = (X_t,\ t \in \mathbb{Z})$ une suite strictement stationnaire de variables aléatoires centrées et de fonction d'autocovariance $\gamma(.)$, si $v_m = \gamma(0) + \sum 2\gamma(j) \neq 0$, alors*

(i) $\lim\limits_{n\to\infty} n\ var(\bar{X}_n) = v_m$.

(ii) $\bar{X}_n \sim N(0, \frac{v_m}{n})$.

L'application de ce théorème au processus $MMR(1)$ donne :

Corollaire 2.1.3. *Soit $X = (X_t,\ t \in \mathbb{Z})$ une suite strictement stationnaire de variables aléatoires 1-dépendantes définies par une $MMR(1)$, tel que*

$$X_t = \varepsilon_t + \theta\varepsilon_{t-1} \qquad avec\, \varepsilon_t \sim IID(0, \sigma_\varepsilon^2).$$

Ces variables sont centrées et ont $\gamma(.)$ pour fonction d'autocovariance, comme $v_1 = \gamma(0) + 2\gamma(1) \neq 0$, alors

(i) $\lim\limits_{n\to\infty} n\ var(\bar{X}_n) = (1+\theta)^2\sigma_\varepsilon^2$.

(ii) $(1+\theta)^2\bar{X}_n \sim N(0, n^{-1}\sigma_\varepsilon^2)$.

Donc d'aprés la définition 1.2.1 de la normalité asymptotique, la condition $v_m \neq 0$ du théorème 2.1.2 est indispensable pour l'obtention de (ii). Dans le cas où $v_m = 0$, on peut montrer que

$n^{\frac{1}{2}}\bar{X}_n \xrightarrow{P} 0$ et $n\ var(\bar{X}_n) \to 0$ quand $n \to \infty$

2.2 Estimation du paramètre d'un processus $MMR(1)$

Nous considérons ici uniquement le cas d'un processus $MMR(1)$, centré et inversible vérifiant

$$X_t = \varepsilon_t + \theta\varepsilon_{t-1}, \tag{2.2.1}$$

où ε_t est un $IID(0,\sigma_\varepsilon^2)$, $\theta \in R$ tel que $|\theta| < 1$. Notre problème consiste à estimer le paramètre θ à partir des observations $X_1,X_2,...,X_n$.

Dans le cas où nous avons supposé que $\mu = 0$ nous n'avons pas ce paramètre à estimer, cepandant si X n'est pas un processus centré, μ est aussi un paramètre à estimer. De même, le paramètre σ_ε^2 est a estimer, que le processus soit centré ou non.

Par la suite, nous nous limitons au problème de l'estimation du paramètre θ.

Bien qu'il existe plusieurs méthodes d'estimations, nous présentons les trois méthodes les plus courantes, à savoir

-La méthode des moments.

-La méthode des moindres carrées.

-La méthode du maximum de la fonction de vraisemblance.

Pour chacune de ces méthodes, nous éxpliquons l'idée de la démarche de l'estimation et nous donnons pour chacune les propriétés asymptotiques de l'estimateur obtenu.

2.2.1 L'estimateur par la méthode des moments

Cette méthode est basée sur la fonction d'autocorrélation, elle consiste à substituer les moments empiriques aux moments théoriques et à résoudre les équations obtenues.

Pour un $MMR(1)$, l'éstimateur de θ noté $\hat{\theta}$, doit être solution de

$$\hat{\rho}(1) = \frac{\hat{\gamma}(1)}{\hat{\gamma}(0)}.$$

Soit à partir de (1.1.10)

$$\hat{\rho}(1) = \frac{\hat{\theta}}{1+\hat{\theta}^2},$$

$\hat{\theta}$ est alors la solution de l'équation

$$\hat{\rho}(1)\hat{\theta}^2 - \hat{\theta} + \hat{\rho}(1) = 0, \tag{2.2.2}$$

et vérifie

$$\hat{\theta} = \frac{1 \pm \sqrt{1-4\hat{\rho}(1)^2}}{2\hat{\rho}(1)}.$$

Plusieurs cas se présentent :

1.Si $|\hat{\rho}(1)| < \frac{1}{2}$, alors l'équation (2.2.2) à 2 solutions. Nous choisissons celle qui donne un modèle inversible, c'est-à-dire vérifiant $|\hat{\theta}| < 1$. On obtient donc

$$\hat{\theta} = \frac{1 - \sqrt{1-4\hat{\rho}(1)^2}}{2\hat{\rho}(1)}$$

2.Si $|\hat{\rho}(1)| = \frac{1}{2}$, alors l'équation (2.2.2) admet une unique solution $|\hat{\theta}| = 1$. Le modèle n'est pas inversible.

3.Si $|\hat{\rho}(1)| > \frac{1}{2}$, alors les racines ne sont pas réelles et donc l'équation (2.2.2) n'a pas de solution réelle.

En résumé, sous l'hypothèse de l'existence d'un modèle $MMR(1)$ inversible défini par (2.2.1), l'estimateur des moments se limite au cas où $|\hat{\rho}(1)| < \frac{1}{2}$.

Remarque 2.2.1. Dans le cas général, la méthode des moments est souvent difficile à résoudre pour un $MMR(q)$, car les autocorrélations $\rho(h)$ sont des fonctions non linéaires des paramètres $\theta_1, ..., \theta_q$.

Du fait de la sensibilité de cette méthode face aux erreurs d'arrondi, la méthode des moments est davantage utilisée pour calculer les valeurs initiales des paramètres avant l'utilisation d'autres méthodes d'estimation, que pour estimer le paramètre en lui même.

Fuller [8] insiste sur le fait que l'estimation de θ ne peut se faire qu'à partir de $\rho(1)$.

Cepandant, pour ne pas se limiter au cas où $|\hat{\rho}(1)| < \frac{1}{2}$, il propose d'étendre l'estimateur des moments en définissant

$$\hat{\theta}_F = \begin{cases} \frac{1-\sqrt{1-4\rho(\hat{1})^2}}{2\rho(\hat{1})} & si \quad |\hat{\rho}(1)| < 0.5 \\ -1 & si \quad \hat{\rho}(1) < -0.5 \\ 1 & si \quad \hat{\rho}(1) > 0.5 \\ 0 & si \quad \hat{\rho}(1) = 0 \end{cases}$$

$\hat{\rho}(1)$ estime $\rho(1)$ avec une erreur de l'ordre de $O(n^{-\frac{1}{2}})$, et $\hat{\theta}_F$ est un estimateur consistant dont l'erreur asymptotique est du même ordre.

2.2.2 L'estimateur des moindres carrés

Le principe des moindres carrés repose sur la recherche de la valeur de θ qui minimise la somme des carrés des erreurs commises lors de l'estimation. Soit $S(\theta)$ cette somme, définie par

$$S(\theta) = \sum_{t=-\infty}^{n} E(\varepsilon_t, \theta, X)^2 = \sum_{t=-\infty}^{n} \hat{\varepsilon}_t^2 \tag{2.2.3}$$

Cette méthode est basée sur une relation de récurrence nécessitant des valeurs initiales de $\hat{\varepsilon}_t^2$.

Selon la méthode d'initialisation de $\hat{\varepsilon}_t$, pour $t \leq 0$, on distingue deux types d'estimateurs :

- L'estimateur des moindres carrés conditionnel.

- L'estimateur des moindres carrés non- conditionnel.

2.2.2.1 L'estimateur des moindres carrés conditionnel

On considère que les valeurs initiales des bruits blancs sont des valeurs fixées. Pour un $MMR(q)$ les valeurs $\hat{\varepsilon}_{1-q}$,$\hat{\varepsilon}_{2-q}$, ...,$\hat{\varepsilon}_{-1}$, $\hat{\varepsilon}_0$ sont supposées nulles. Nous limitant au $MMR(1)$, nous présumons que $\hat{\varepsilon}_0 = 0$, et cherchons à minimiser

$$SC(\theta) = \sum_{t=1}^{n} \varepsilon_t^2.$$

La relation (2.2.1) s'écrit,

$$\varepsilon_t = X_t - \theta\varepsilon_{t-1}.$$

Lemme 2.2.2. *Pour tout* $t \geq 2$, $\varepsilon_t = X_t - \theta\varepsilon_{t-1}$ *s'écrit sous la forme suivante*

$$\varepsilon_t = \sum_{i=0}^{t-1} (-\theta)^i X_{t-i}.$$

Démonstration. On a

$$\begin{aligned} \varepsilon_0 &= 0 \\ \varepsilon_1 &= X_1. \end{aligned}$$

Donc

$$\begin{aligned} \varepsilon_2 &= X_2 - \theta\varepsilon_1 \\ &= X_2 - \theta X_1 \end{aligned}$$

$$\begin{aligned} \varepsilon_3 &= X_3 - \theta\varepsilon_2 \\ &= X_3 - \theta(X_2 - \theta X_1) \\ &= X_3 - \theta X_2 + \theta^2 X_1 \end{aligned}$$

$$\begin{aligned}\varepsilon_4 &= X_4 - \theta\varepsilon_3 \\ &= X_4 - \theta(X_3 - \theta X_2 + \theta^2 X_1) \\ &= X_4 - \theta X_3 + \theta^2 X_2 - \theta^3 X_1\end{aligned}$$

.

.

.

$$\varepsilon_t = \sum_{i=0}^{t-1}(-\theta)^i X_{t-i}.$$

$\square$

Donc d'aprés le lemme 2.2.2, on a

$$\varepsilon_t = \sum_{i=0}^{t-1}(-\theta)^i X_{t-i}.$$

Ainsi, la somme des carrés conditionnellement à $\hat{\varepsilon}_0 = 0$ devient

$$SC(\theta) = \sum_{t=1}^{n}\varepsilon_t^2 = \sum_{t=1}^{n}\sum_{i=1}^{t-1}((-\theta)^i X_{t-i})^2.$$

L'estimateur $\hat{\theta}_{SC}$ est la valeur qui minimise la fonction $SC(\theta)$. Celle-ci n'étant pas quadratique en θ, l'équation permettant d'estimer θ n'est donc pas quadratique.

2.2.2.2 L'estimateur des moindres carrés non conditionnel

Le problème tout d'abord est de minimiser

$$SC(\theta) = \sum_{t=1}^{n} \varepsilon_t^2.$$

Mais minimiser $\sum_{t=1}^{n} \varepsilon_t^2$ revient à minimiser $\varepsilon'\varepsilon$,
et ε s'écrit sous la forme suivante

$$\varepsilon = MX + W\varepsilon^*$$

ε est un paramètre, M, W deux matrices dim $M(n+1, n)$ et dim $W(n+1, 1)$.

Car

$$\begin{cases} \varepsilon_0 = \varepsilon_0 \\ \varepsilon_1 = X_1 - \theta_1\varepsilon_0 \\ \varepsilon_2 = X_2 - \theta_1\varepsilon_1 \\ \varepsilon_3 = X_3 - \theta_1\varepsilon_2 \\ . \\ . \\ . \\ \varepsilon_n = X_n - \theta_1\varepsilon_{n-1} \end{cases} \iff \begin{cases} \varepsilon_0 = \varepsilon_0 \\ \varepsilon_1 = X_1 - \theta_1\varepsilon_0 \\ \varepsilon_2 = X_2 - \theta_1 X_1 + \theta_1^2\varepsilon_0 \\ \varepsilon_3 = X_3 - \theta_1 X_2 + \theta_1^2 X_1 - \theta_1^3\varepsilon_0 \\ . \\ . \\ . \\ \varepsilon_n = X_n - \theta_1 X_{n-1} + (-\theta_1)^2 X_n + ... + (-\theta_1)^{n-1} X_1 + (-\theta_1)^n \varepsilon_0 \end{cases}$$

$$\varepsilon_i = X_i + \sum_{k=1}^{i-1} (-1)^k \theta_1^k X_{i-k} + (-1)^i \theta_1^i \varepsilon_0.$$

Avec $\begin{cases} MX = \sum\limits_{k=1}^{i-1}(-1)^k\theta_1^k X_{i-k} \\ W\varepsilon_0 = (-1)^i\theta_1^i\varepsilon_0. \end{cases}$

$$\begin{pmatrix} \varepsilon_0 \\ . \\ . \\ . \\ \varepsilon_n \end{pmatrix} = \begin{pmatrix} 1 & 0 & . & 0 \\ . & . & . & . \\ . & . & . & . \\ . & . & . & . \\ (-\theta_1)^{n-1} & . & & 1 \end{pmatrix} X + \begin{pmatrix} 1 \\ -\theta_1 \\ . \\ . \\ (-\theta_1)^n \end{pmatrix} \varepsilon_0,$$

Commençons d'abord par estimer les valeurs initiales pour améliorer l'estimation.

Soit $\varepsilon = (\varepsilon_{1-q}, ..., \varepsilon_n)$.

On prend $MM(1)$ donc $\varepsilon = (\varepsilon_0, ..., \varepsilon_n)$. Et $\varepsilon^* = \varepsilon_0$.

Comme $q = 1$ donc $MM(1)$

$$\varepsilon^* = \varepsilon_0.$$

$$\hat{\varepsilon}^* = -(W'W)^{-1}W'MX.$$

D'aprés $Y = Z\theta + V$

On a $\hat{\theta} = (Z'Z)^{-1}Z'Y$

On a la proposition suivante

Proposition 2.2.3. *Soit* ε^**=* $(\varepsilon_{1-q},...,\varepsilon_0)$*, l'estimateur des moindres carrés* $\hat{\varepsilon}^*$ *de* ε^* *est le suivant*

$$\hat{\varepsilon}^* = -(W^{'}W)^{-1}W'MX$$

Démonstration. On a Soit $Y = Z\theta + V$ tel que $V_t \hookrightarrow N(0, \sigma^2)$

$$\theta_{OLS} = (\hat{Z'}Z)^{-1}Z'Y$$

Mais nous, on a $\varepsilon = MX + W\varepsilon^*$ tel que $\varepsilon_t \hookrightarrow N(0, \sigma^2)$

$\varepsilon = MX + W\varepsilon^* \Longleftrightarrow -MX = W\varepsilon^* - \varepsilon$

$$\begin{aligned} \hat{\varepsilon}^* &= (W'W)^{-1}W'(-MX) \\ &= -(W'W)^{-1}W'MX \end{aligned}$$

$$\begin{cases} W = Z \\ \varepsilon = V \\ Y = MX. \end{cases} \qquad \square$$

Or

$$S(\theta) = \varepsilon'\varepsilon = (MX + W\varepsilon^*)'(MX + W\varepsilon^*)$$

et

$$\hat{S}(\theta) = (MX + W\hat{\varepsilon}^*)'(MX + W\hat{\varepsilon}^*)$$

Soit

$$\begin{aligned} MX + W\varepsilon^* &= MX + W\hat{\varepsilon}^* - W\hat{\varepsilon}^* + W\varepsilon^* \\ &= MX + W\hat{\varepsilon}^* + W(\varepsilon^* - \hat{\varepsilon}^*) \end{aligned}$$

donc

$$\begin{aligned} S(\theta) &= (MX + W\hat{\varepsilon}^* + W(\varepsilon^* - \hat{\varepsilon}^*))'(MX + W\hat{\varepsilon}^* + W(\varepsilon^* - \hat{\varepsilon}^*)) \\ &= [(MX + W\hat{\varepsilon}^*) + (\varepsilon^* - \hat{\varepsilon}^*)'W'][MX + W\hat{\varepsilon}^* + W(\varepsilon^* - \hat{\varepsilon}^*)] \\ &= (MX + W\hat{\varepsilon}^*)'(MX + X\hat{\varepsilon}^*) + (MX + W\hat{\varepsilon}^*)'W(\varepsilon^* - \hat{\varepsilon}^*) + (\varepsilon^* - \hat{\varepsilon}^*)'W'(MX + W\hat{\varepsilon}^*) \end{aligned}$$

$+'(\varepsilon^* - \hat{\varepsilon}^*)W'W(\varepsilon^* - \hat{\varepsilon}^*)$

De plus

$$(MX + W\hat{\varepsilon}^*)'W(\varepsilon^* - \hat{\varepsilon}^*) = 0$$

$$(\varepsilon^* - \hat{\varepsilon}^*)' W' (MX + W\hat{\varepsilon}^*) = 0,$$

donc

$$S(\theta) = S(\theta) + (\varepsilon^* - \hat{\varepsilon}^*) W' W (\varepsilon^* - \hat{\varepsilon}^*)$$

et enfin trouver l'estimateur des moindres carrés non conditionnel revient a minimiser le terme $S(\theta)$.

2.2.3 L'estimateur par la méthode du maximum de la vraisemblance

Là encore, nous étudions deux types d'éstimateurs selon le choix des valeurs initiales, à savoir

-L'estimateur du maximum de la vraisemblance conditionnelle.

-L'estimateur du maximum de la vraisemblance non conditionnelle ou vraisemblance exacte .

2.2.3.1 L'estimateur du maximum de la vraissemblance conditionnelle

Ayant une relation entre X_t et ε_t, la fonction de vraissemblance des X_t est déterminée à partir de celle de ε_t. On a donc besoin d'une hypothèse sur la distribution liée des ε_t : nous les supposons indépendantes de loi $g(0, \sigma_\varepsilon^2)$. Partant de la densité des probabilités conditionnelles des bruits blancs, on a

$$f(X, \theta, \sigma_\varepsilon^2) = (2\pi\sigma_\varepsilon^2)^{-\frac{1}{2}} exp(-\frac{\varepsilon_t^2}{2\sigma_\varepsilon^2}) \tag{2.2.4}$$

En réecrivant ε_t à partir de (2.2.1) , on peut écrire la fonction de vraisemblance en fonction des paramètres $(\theta, \sigma_\varepsilon^2)$ en supposant que les valeurs initiales X_0, ε_{-1}, ε_0 sont fixées à 0.

La fonction de vraisemblance conditionnelle est alors

$$L_C(\theta,\sigma_\varepsilon^2)=(2\pi\sigma_\varepsilon^2)^{-\frac{n}{2}}exp(-\frac{1}{2\sigma_\varepsilon^2}\sum_{t=1}^{n}\varepsilon_t^2),$$

d'où la fonction de la log-vraissemblance

$$LogL_C(\theta,\sigma_\varepsilon^2)=-\frac{n}{2}Log(2\pi\sigma_\varepsilon^2)-\frac{1}{2\sigma_\varepsilon^2}SC(\theta) \tag{2.2.5}$$

où

$$SC(\theta)=\sum_{t=1}^{n}\varepsilon_t^2(\theta) \tag{2.2.6}$$

Les premiers termes de $Log\ L_C(\theta,\sigma_\varepsilon^2)$ étant négligeables lorsque le nombre d'observations est grand, il en découle que $Log\ L_C(\theta,\sigma_\varepsilon^2)$ est dominée par $SC(\theta)$, et de ce fait, rechercher le maximum du logarithme de la vraisemblance conditionnelle coincide avec l'estimateur conditionnel des moindres carrés étudié précédemment.

2.2.3.2 L'estimateur du maximum de la vraissemblace non conditionnelle

Comme on a déjà vu dans la partie conditionnelle on connait la fonction de vraisemblance des ε_t

$$L_C(\theta,\sigma_\varepsilon^2)=(2\pi\sigma_\varepsilon^2)^{-\frac{n}{2}}exp(-\frac{1}{2\sigma_\varepsilon^2}\sum_{t=1}^{n}\varepsilon_t^2)$$

et pour obtenir la fonction de vraisemblance, nous utilisons la densité de probabilité d'une série d'observation $X_1,...,X_n$, en supposant que chaque observation de cette série est générée par une $MMR(1)$ définie par (2.2.1).

Sous l'hypothèse de normalité des ε_i , et par conséquent des X_i , on obtient la fonction de densité jointe

$$f(X,\theta,\sigma_\varepsilon^2)=(2\pi\sigma_\varepsilon^2)^{-\frac{n}{2}}(det\Gamma)^{-\frac{n}{2}}exp(-\frac{X'\Gamma^{-1}X}{2\sigma_\varepsilon^2}) \tag{2.2.7}$$

et chaque $X_i = \varepsilon_i + \theta\varepsilon_{i-1}$ pour chaque $i = 1, ..., n$

où X est le vecteur des observations $(X_1, ..., X_n)$ et Γ est la matrice de covariance de $\frac{X}{\sigma_\varepsilon^2}$, définie positive .

Mais là nous somme confrontés au problème de l'inversion de la matriceΓ

donc on passe à une autre méthode qui nous facilite les calculs.

On commence d'abord par,

Proposition 2.2.4. *Soient $\varepsilon = (\varepsilon_{1-q},...,\varepsilon_n)$, $X = (X_1..., X_n)$, $\varepsilon^* = (\varepsilon_{1-q},...,\varepsilon_0)$ et M, W deux matrices de dimension $(n+q)\times n$ et $(n+q)\times(n+q)\times q$ respectivement et leurs éléments sont en fonction de θ seulement. Alors notre ε va s'écrire sous cette forme*

$$\varepsilon = MX + W\varepsilon^* \tag{2.2.8}$$

Pour $q = 1$

$$\varepsilon = (\varepsilon_0, ..., \varepsilon_n), X = (X_1..., X_n),$$

$$\varepsilon^* = \varepsilon_0$$

et M matrice, dim $(n+1)\times n$

W matrice, dim $(n+1)\times 1$

On a

$$\begin{cases} \varepsilon_0 = \varepsilon_0 \\ \varepsilon_1 = X_1 - \theta_1 \varepsilon_0 \\ \varepsilon_2 = X_2 - \theta_1 \varepsilon_1 \\ \varepsilon_3 = X_3 - \theta_1 \varepsilon_2 \\ . \\ . \\ . \\ \varepsilon_n = X_n - \theta_1 \varepsilon_{n-1} \end{cases} \Longleftrightarrow \begin{cases} \varepsilon_0 = \varepsilon_0 \\ \varepsilon_1 = X_1 - \theta_1 \varepsilon_0 \\ \varepsilon_2 = X_2 - \theta_1 X_1 + \theta_1^2 \varepsilon_0 \\ \varepsilon_3 = X_3 - \theta_1 X_2 + \theta_1^2 X_1 - \theta_1^3 \varepsilon_0 \\ . \\ . \\ . \\ \varepsilon_n = X_n - \theta_1 X_{n-1} + (-\theta_1)^2 X_n + ... + (-\theta_1)^{n-1} X_1 + (-\theta_1)^n \varepsilon_0 \end{cases}$$

donc

$$\varepsilon_i = X_i + \sum_{k=1}^{i-1} (-1)^k \theta_1^k X_{i-k} + (-1)^i \theta_1^i \varepsilon_0.$$

$$\begin{pmatrix} \varepsilon_0 \\ . \\ . \\ . \\ \varepsilon_n \end{pmatrix} = \begin{pmatrix} m_{1,0} & . & . & m_{1,n} \\ . & . & . & . \\ . & . & . & . \\ . & . & . & . \\ m_{n,0} & . & . & m_{n+1,n} \end{pmatrix} \begin{pmatrix} X_1 \\ . \\ . \\ . \\ X_n \end{pmatrix} + \begin{pmatrix} W_1 \\ . \\ . \\ . \\ W_{n+1} \end{pmatrix} \varepsilon_0, \qquad (2.2.9)$$

avec

$$M=\begin{pmatrix} 0 & . & . & . & . & . & 0\\ 1 & 0 & . & . & . & . & 0\\ -\theta_1 & 1 & 0 & . & . & . & 0\\ \theta_1^2 & -\theta_1 & 1 & 0 & . & . & 0\\ . & . & . & 1 & 0 & . & 0\\ . & . & . & . & 1 & 0 & 0\\ . & . & . & . & . & 1 & 0\\ (-\theta_1)^{n-1} & (-\theta_1)^{n-2} & . & . & . & . & 1 \end{pmatrix},$$

et

$$W=\begin{pmatrix} 1\\ -\theta_1\\ \theta_1^2\\ -\theta_1^3\\ .\\ .\\ .\\ (-\theta_1)^n \end{pmatrix}$$

(2.2.9) $\Leftrightarrow \varepsilon = MX + W\varepsilon^{*}$,
avec ε^{*}=ε_0 , $i = 1, ..., n + q$ et $j = 1, ..., n$.

Proposition 2.2.5. *Soit* ε^*= (ε_{1-q},...,ε_0)*, l'estimateur de maximum de vraisemblance* $\hat{\varepsilon}^*$ *de* ε^* *est le suivant*

$$\hat{\varepsilon}^* = -(W'W)^{-1}W'MX \tag{2.2.10}$$

Démonstration. Soit $Y = Z\theta + V$ tel que $V_t \hookrightarrow N(0, \sigma^2)$

$$\theta_{OLS} = (\hat{Z'}Z)^{-1}Z'Y$$

Mais, on a $\varepsilon = MX + W\varepsilon^*$ tel que $\varepsilon_t \hookrightarrow N(0, \sigma^2)$

$\varepsilon = MX + W\varepsilon^* \Longleftrightarrow -MX = W\varepsilon^* - \varepsilon$

$$\begin{aligned}
\hat{\varepsilon}^* &= (W^{'}W)^{-1}W'(-MX) \\
&= -(W^{'}W)^{-1}W'MX
\end{aligned}$$

$$\begin{cases} W = Z \\ \varepsilon = V \\ Y = MX. \end{cases}$$ □

Théorème 2.2.6. *Soient W une matrice de dimention $(n+q)\times q$ ses éléments sont en fonction de θ seulement et $S(\theta) = \varepsilon'\varepsilon$ alors :*

$$l(\theta, \sigma^2) = -\frac{n}{2}log2\pi - \frac{n}{2}log\sigma^2 - \frac{1}{2}log|W'W| - \frac{1}{2\sigma^2}S(\theta) \qquad (2.2.11)$$

Démonstration. On a

$$S(\theta) = \varepsilon'\varepsilon = (MX + W\varepsilon^*)'(MX + W\varepsilon^*)$$

e

$$\hat{S}(\theta) = (MX + W\hat{\varepsilon}^*)^{'}(MX + W\hat{\varepsilon}^*)$$

Soit

donc

$$\begin{aligned}
S(\theta) &= (MX + W\hat{\varepsilon}^* + W(\varepsilon^* - \hat{\varepsilon}^*))^{'}(MX + W\hat{\varepsilon}^* + W(\varepsilon^* - \hat{\varepsilon}^*)) \\
&= [(MX + W\hat{\varepsilon}^*) + (\varepsilon^* - \hat{\varepsilon}^*)^{'}W^{'}][MX + W\hat{\varepsilon}^* + W(\varepsilon^* - \hat{\varepsilon}^*)] \\
&= (MX + W\hat{\varepsilon}^*)(MX + X\hat{\varepsilon}^*) + (MX + W\hat{\varepsilon}^*)W(\varepsilon^* - \hat{\varepsilon}^*) + (\varepsilon^* - \hat{\varepsilon}^*)^{'}W^{'}(MX + W\hat{\varepsilon}^*) \\
&+ (\varepsilon^* - \hat{}^*)^{'}\hat{W}^{'}W(\varepsilon^* - \hat{\varepsilon}^*) \\
&= (MX + W\hat{\varepsilon}^*)'(MX + X\hat{\varepsilon}^*) + (MX + W\hat{\varepsilon}^*)'W(\varepsilon^* - \hat{\varepsilon}^*) + (\varepsilon^* - \hat{\varepsilon}^*)^{'}W^{'}(MX + W\hat{\varepsilon}^*) \\
&+ (\varepsilon^* - \hat{\varepsilon}^*)^{'}W^{'}W(\varepsilon^* - \hat{\varepsilon}^*)
\end{aligned}$$

De plus

$$(MX + W\hat{\varepsilon}^*)'W(\varepsilon^* - \hat{\varepsilon}^*) = 0$$

$$(\varepsilon^* - \hat{\varepsilon}^*)^{'}W^{'}(MX + W\hat{\varepsilon}^*) = 0,$$

donc

$$S(\theta) = S(\theta) + (\varepsilon^* - \hat{\varepsilon}^*)W^{'}W(\varepsilon^* - \hat{\varepsilon}^*) \qquad (2.2.12)$$

car

$$(MX+W\hat{\varepsilon}^*)^{'}W(\varepsilon^*-\hat{\varepsilon}^*) = (MX)^{'}W\varepsilon^* - (MX)^{'}W\hat{\varepsilon}^* + (W\hat{\varepsilon}^*)^{'}W\varepsilon^* - (W\hat{\varepsilon}^*)^{'}W\hat{\varepsilon}^*$$

Sachant déjà que $W^{'}W\hat{\varepsilon}^* = W^{'}MX \iff \begin{cases} (W^{'}W\hat{\varepsilon}^*)^{'} & = -(W^{'}MX)^{'} \\ (W\hat{\varepsilon}^*)^{'}W & = -(MX)\text{'}W \\ (W\hat{\varepsilon}^*)^{'}W\varepsilon^* & = -(MX)\text{'}W\varepsilon^* \\ (W\hat{\varepsilon}^*)^{'}W\hat{\varepsilon}^* & = -(MX)\text{'}W\hat{\varepsilon}^* \end{cases}$

de ça on obtient

$$(MX + W\hat{\varepsilon}^*)'W(\varepsilon^* - \hat{\varepsilon}^*) = 0.$$

On a :

$$f(\varepsilon;\theta,\sigma) = (2\pi\sigma^2)^{-\frac{(n+q)}{2}} e^{-\frac{\varepsilon^{'}\varepsilon}{2\sigma^2}}$$

et ε est une fonction de (ε^*,X) car $\varepsilon = MX + W\varepsilon^*$

et de plus

$$f(x,\varepsilon^*) = f(h^{-1}(x,\varepsilon^*))|J_{h^{-1}}(x,\varepsilon^*)|$$

avec

$$\begin{cases} (X,\varepsilon^*) & = h(\varepsilon) \\ |J_{h^{-1}}(x,\varepsilon^*)| & = 1 \end{cases} \iff \begin{cases} \varepsilon & = h^{-1}(X,\varepsilon^*) = MX + W\varepsilon^* \end{cases}$$

donc

$$f(x,\varepsilon^*;\theta,\sigma) = (2\pi\sigma^2)^{-\frac{(n+q)}{2}} |\frac{D(\varepsilon)}{D(x,\varepsilon^*)}| e^{-\frac{S(\theta)}{2\sigma^2}}$$

en remplaçant $S(\theta)$ par $\hat{S}(\theta)+(\varepsilon^*-\hat{\varepsilon}^*)'W'W(\varepsilon^*-\hat{\varepsilon}^*)$

on aura

$$f(x,\varepsilon^*;\theta,\sigma)=(2\pi\sigma^2)^{-\frac{(n+q)}{2}}|\frac{D(\varepsilon)}{D(x,\varepsilon^*)}|e^{-\frac{S(\theta)}{2\sigma^2}}$$

devient

$$\begin{aligned} f(x,\varepsilon^*;\theta,\sigma) &= (2\pi\sigma^2)^{-\frac{n}{2}}(2\pi\sigma^2)^{-\frac{q}{2}}|W'W|^{-\frac{1}{2}}|W'W|^{\frac{1}{2}}e^{-\frac{\hat{S}(\theta)}{2\sigma^2}}e^{-\frac{(\varepsilon^*-\hat{\varepsilon}^*)'W'W(\varepsilon^*-\hat{\varepsilon}^*)}{2\sigma^2}} \\ &= (2\pi\sigma^2)^{-\frac{n}{2}}|W'W|^{-\frac{1}{2}}e^{-\frac{\hat{S}(\theta)}{2\sigma^2}}(2\pi\sigma^2)^{-\frac{q}{2}}|W'W|^{\frac{1}{2}}e^{-\frac{(\varepsilon^*-\hat{\varepsilon}^*)'W'W(\varepsilon^*-\hat{\varepsilon}^*)}{2\sigma^2}} \end{aligned}$$

cepandant

$$f(x,\varepsilon^*;\theta,\sigma)=f(x;\theta,\sigma)f(\varepsilon^*\diagdown x;\theta,\sigma),$$

avec $L(\theta,\sigma)=f(x,\theta,\sigma)=(2\pi\sigma^2)^{-\frac{n}{2}}|W'W|^{\frac{-1}{2}}e^{-\frac{\hat{S}(\theta)}{2\sigma^2}}$,

et

$f(\varepsilon^*\diagdown x;\theta,\sigma)=(2\pi\sigma^2)^{-\frac{q}{2}}|W'W|^{\frac{1}{2}}e^{-\frac{(\varepsilon^*-\hat{\varepsilon}^*)'W'W(\varepsilon^*-\hat{\varepsilon}^*)}{2\sigma^2}}$

avec $L(\theta,\sigma)=f(x,\theta,\sigma)=(2\pi\sigma^2)^{-\frac{n}{2}}|W'W|^{\frac{-1}{2}}e^{-\frac{\hat{S}(\theta)}{2\sigma^2}}$,

et maintenant on fait l'identification avec

$L(\theta$,$\sigma)$ la fonction de vraissemblance précedente

$$L(\theta,\sigma)=f(x,\theta,\sigma)=(2\pi\sigma^2)^{-\frac{n}{2}}|V|^{\frac{-1}{2}}e^{-\frac{X'V^{-1}X}{2\sigma^2}}$$

on aura $\begin{cases} |V| & =|W'W| \\ \hat{S}(\theta) & =X'V^{-1}X \end{cases}$

d'où le résultat final.

$$l(\theta,\sigma^2)=-\frac{n}{2}log2\pi-\frac{n}{2}log\sigma^2-\frac{1}{2}log|W'W|-\frac{1}{2\sigma^2}S(\theta)$$

avec

$$S(\theta)=\sum_{t=1-q}^{n}\varepsilon_t^2.$$

□

Trouver l'estimateur du maximum de la vraisemblance non conditionnelle revient a minimiser le terme $S(\theta)$.

Chapitre 3

L'estimateur du RIV (Cas scalaire)

3.1 Introduction

Dans ce chapitre, nous étudions la méthode RIV proposée par Anna Clara Monti, c'est une méthode itérative d'estimation des paramètres d'un processus moyenne mobile

Soit $(X_t, t \in Z)$ un processus mobile d'ordre 1 $MM(1)$ défini par

$$X_t = \varepsilon_t - \beta\varepsilon_{t-1} \tag{3.1.1}$$

avec $|\beta| \leq 1$ et $(\varepsilon_t, t \in \mathbb{Z}$ un bruit blanc faible

Le processus $(X_t, t \in \mathbb{Z})$ est inversible si $|\beta| < 1$

On a

$$X_t = \varepsilon_t - \beta\varepsilon_{t-1}$$

donc

$$\begin{cases} X_1 = \varepsilon_1 - \beta\varepsilon_0 \\ . \\ . \\ . \\ X_n = \varepsilon_n - \beta\varepsilon_{n-1} \end{cases}$$

On fait la somme

$$\sum_{t=1}^{n} X_t = \sum_{t=1}^{n} \varepsilon_t - \beta\varepsilon_{t-1}$$

On multiplie par ε_{t-1}

$$\sum_{t=1}^{n} X_t\varepsilon_{t-1} = \sum_{t=1}^{n} \varepsilon_t\varepsilon_{t-1} - \beta\sum_{t=1}^{n} \varepsilon_{t-1}^2$$

On passe à l'éspérance

$$E(\sum_{t=1}^{n} X_t\varepsilon_{t-1}) = E(\sum_{t=1}^{n} \varepsilon_t\varepsilon_{t-1} - \beta\sum_{t=1}^{n} \varepsilon_{t-1}^2)$$

On prend $\frac{1}{n}\sum\limits_{t=1}^{n} \varepsilon_t\varepsilon_{t-1} = 0$

$$\frac{1}{n}\sum_{t=1}^{n} X_t\varepsilon_{t-1} = -\beta\frac{1}{n}\sum_{t=1}^{n} \varepsilon_{t-1}^2$$

$$\Rightarrow \hat{\beta} = \frac{\frac{1}{n}\sum\limits_{t=1}^{n} X_t\varepsilon_{t-1}}{\frac{1}{n}\sum\limits_{t=1}^{n} \varepsilon_{t-1}^2} \tag{3.1.2}$$

3.2 Principe de la méthode RIV

On génère un échantillon de taille n, v.a.i.i.d, centrées, réduites, de loi quelconque b_0^0,b_1^0, ...,b_n^0

Pour $i = 1$

$$\begin{cases} X_1 = -\beta b_0^{(0)} \\ X_2 = -\beta b_1^{(0)} \\ . \\ . \\ . \\ X_n = -\beta b_{t-1}^{(0)} \end{cases}$$

$$\begin{cases} X_1 b_0^{(0)} = -\beta b_0^{(0)} b_0^{(0)} \\ X_2 b_1^{(0)} = -\beta b_1^{(0)} b_1^{(0)} \\ . \\ . \\ . \\ X_n b_{t-1}^{(0)} = -\beta b_{t-1}^{(0)} b_{t-1}^{(0)} \end{cases}$$

On sommant, terme à terme, on obtient

$$\sum_{t=1}^{n} b_{t-1}^{(0)} X_t = -\beta \sum_{t=1}^{n} {b_{t-1}^{(0)}}^2$$

et on multiplie par $\frac{1}{n}$, on aura

$$\frac{1}{n}\sum_{t=1}^{n} X_t b_{t-1}^{(0)} = -\hat{\beta}\frac{1}{n}\sum_{t=1}^{n} {b_{t-1}^{(0)}}^2$$

Donc

$$\hat{\beta}_n^{(1)} = -\frac{\frac{1}{n}\sum\limits_{t=1}^{n} X_t b_{t-1}^{(0)}}{\frac{1}{n}\sum\limits_{t=1}^{n} (b_{t-1}^{(0)})^2}. \tag{3.2.1}$$

De nouveau, on définit $b_t^{(i)}$ par
$b_t^{(i)} = X_t + \hat{\beta}^{(i)} b_{t-1}^{(i-1)}$, $t = 1, ..., n$
Aprés pour $i = i + 1 (i = 2)$

$$\hat{\beta}_n^{(2)} = -\frac{\frac{1}{n}\sum\limits_{t=1}^{n} X_t b_{t-1}^{(1)}}{\frac{1}{n}\sum\limits_{t=1}^{n} (b_{t-1}^{(1)})^2}.$$

On répète cette opération jusqu'a ce que

$$|\hat{\beta}_n^{(i)} - \hat{\beta}_n^{(i-1)}| < \varepsilon,$$

$$\hat{\beta}_n^{(i)} = -\frac{\frac{1}{n}\sum\limits_{t=1}^{n} X_t b_{t-1}^{(i-1)}}{\frac{1}{n}\sum\limits_{t=1}^{n} (b_{t-1}^{(i-1)})^2}, \tag{3.2.2}$$

avec ε quelconque $(\varepsilon > 0)$.
On définit

$$\hat{\beta}_{RIVE} = \lim_{i \to \infty} \hat{\beta}_n^{(i)} \tag{3.2.3}$$

et aussi

$$\hat{\sigma}_{RIV}^2 = \frac{1}{n}\sum_{t=1}^{n} b_t^2 \tag{3.2.4}$$

et $b_t = \lim\limits_{i \to \infty} b_t^{(i)}$

donc

$$\hat{\sigma}^2_{RIV} = \frac{1}{n}\sum_{t=1}^{n}(\lim_{i\to\infty} b_t^{(i)})^2. \qquad (3.2.5)$$

Remarque 3.2.1. 1.Les nombres b_0^0,b_1^0, ...,b_{n-1}^0 peuvent ètre générés, par exemple en utilisant une distribution gausienne.

2.Quand le nombre d'itération i augmente la variable aléatoire $b_t^{(i)}$ converge vers le bruit ε_t en probabilité. Les propriétées des estimateurs $\hat{\beta}_{RIVE}$ et σ^2_{RIVE} pour un processus moyenne mobile d'ordre un, sont données par le théorème suivant.

3.3 Théorème

Soient $\hat{\beta}_{RIVE}$ et σ^2_{RIVE} définient par (3.2.3) et (3.2.4) respectivement, $\beta \in]-1, 1[$, l'éstimateur RIVE à les propriétées suivantes

i) $\lim\limits_{n\to+\infty} \hat{\beta}_{RIVE} \to \beta$ en probabilité.

ii) $\lim\limits_{n\to+\infty} \hat{\sigma}^2_{RIVE} \to \sigma^2$ en probabilité.

iii) $\lim\limits_{n\to+\infty} b_t \to \varepsilon_t$ en probabilité.

iv) $\sqrt{n}(\hat{\beta}_{RIVE} - \beta) \Rightarrow N(0, 1)$.

Pour la démonstration de ce théorème on a besoin des lemmes suivants

Lemme 3.3.1. *On a pour j< i* $(i > 1)$ *l'estimateur* $\hat{\beta}^{(j)}$ *admet la décomposition suivante*

$$\hat{\beta}^{(j)} = \nu(j) + O_p(n^{-\frac{1}{2}}) pour\ \ j < i. \qquad (3.3.1)$$

Démonstration. Pour prouver cela il faut montrer que , $E(\hat{\beta}^{(j)}) = \nu(j)$ et que $\hat{\beta}^{(j)} \sim N(\nu(j), \frac{v_1}{n})$.

On sait que

$$\hat{\beta}_n^{(i)} = -\frac{\frac{1}{n}\sum\limits_{t=1}^{n} X_t b_{t-1}^{(i-1)}}{\frac{1}{n}\sum\limits_{t=1}^{n}(b_{t-1}^{(i-1)})^2}.$$

On pose

$$Y_t = X_t b_{t-1}.$$

On passe à l'éspérance de Y_t

$$E(Y_t) = E(X_t b_{t-1}) = \nu(j),$$

et

$$E(b_{t-1}^{j-1 2}) = Var(b_{t-1}^{j-1}) = 1$$

Y_t et Y_{t+1} ne sont pas indépendants,

car $E(Y_t Y_{t+1}) \neq 0$

($Y_t = X_t b_{t-1}$, $Y_{t+1} = X_{t+1} b_t$

$X_t = \varepsilon_t - \theta\varepsilon_{t-1}$; $X_t = (\varepsilon_t,\ \theta,\ \varepsilon_{t-1}$) et $X_{t+1} = \varepsilon_{t+1} - \theta\varepsilon_t$;$X_{t+1} = (\varepsilon_{t+1}, \theta,\ \varepsilon_t$)

donc $X_t$et X_{t+1} sont 1- dépendants)

Par contre

Y_t et Y_{t+2} sont indépendants,

car ($Y_t = X_t b_{t-1}$ et $X_t = \varepsilon_t - \theta\varepsilon_{t-1}$ donc X_t est en fonction de ε_t, θ, ε_{t-1}

et $y_{t+2} = X_{t+2} b_{t+1}$ et $X_{t+2} = \varepsilon_{t+2} - \theta\varepsilon_{t+1}$ donc X_{t+2} est en fonction de $\varepsilon_{t+2}, \theta, \varepsilon_{t+1}$

donc X_t et X_{t+2} sont indépendants

et par suite

Y_t et Y_{t+2} sont indépendants

Donc les Y_t sont 1-dépendants d'aprés la définition 1.1.15

Et par le T.C.L 2.1.2

$$\frac{1}{n}\sum_{t=1}^{n} Y_t \sim N(\nu(j), \frac{v_1}{n}).$$

$$\begin{aligned} \frac{1}{n}\sum_{t=1}^{n}(b_{t-1}^{(i-1)})^2\hat{\beta}_n^{(i)} &= \frac{1}{n}\sum_{t=1}^{n}X_t b_{t-1}^{(i-1)} \\ &= \frac{1}{n}\sum_{t=1}^{n}Y_t. \end{aligned}$$

Mais $\frac{1}{n}\sum\limits_{t=1}^{n}Y_t \sim N(\nu(j), \frac{v_1}{n})$ donc $\frac{1}{n}\sum\limits_{t=1}^{n}(b_{t-1}^{(i-1)})^2\hat{\beta}_n^{(i)} \sim N(\nu(j), \frac{v_1}{n})$, et puis $\hat{\beta}_n^{(i)} \sim N(\nu(j), \frac{v_1}{n})$. □

Lemme 3.3.2. *On convient pour $j = 1$, on prend $\nu(1) = 0$.*

Démonstration. D'aprés le lemme 3.3.1 on a

$$\hat{\beta}^{(1)} = \nu(1) + O_p(n^{-\frac{1}{2}})$$

on doit montrer que $\nu(1) = 0$ ç-a-d que $E(\hat{\beta}^{(1)}) = 0$.

Or

$$\hat{\beta}^{(1)} = -\frac{\frac{1}{n}\sum\limits_{t=1}^{n}X_t b_{t-1}^{(0)}}{\frac{1}{n}\sum\limits_{t=1}^{n}(b_{t-1}^{(0)})^2}.$$

On pose

$$Y_t = X_t b_{t-1}^{(0)}$$

$$E(Y_t) = E(X_t b_{t-1}^{(0)})$$

$$X_t = \varepsilon_t - \theta\varepsilon_{t-1},$$

donc

$$E(X_t b_{t-1}^{(0)}) = E((\varepsilon_t - \theta\varepsilon_{t-1})b_{t-1}^{(0)}) = 0,$$

et enfin

$$\hat{\beta}^{(1)} = O_p(n^{-\frac{1}{2}}).$$

□

Lemme 3.3.3. *On a pour $k \geq 2$*

$$\frac{1}{n}\sum_{t=1}^{n} X_t X_{t-k} = O_p(n^{-\frac{1}{2}}).$$

Démonstration. On a

$$E(X_t X_{t-k}) = 0,$$

pour $k \geq 2$, d'aprés (1.1.9) ,les X_t sont 1- dépendants donc d'aprés le T.C.L

$$\frac{1}{n}\sum_{t=1}^{n} X_t X_{t-k} \sim N(\nu(j), \frac{v_1}{n}).$$

Ici $\nu(j) = 0$,

donc

$$\frac{1}{n}\sum_{t=1}^{n} X_t X_{t-k} = O_p(n^{-\frac{1}{2}}).$$

□

Lemme 3.3.4. *On a*

$$\frac{1}{n}\sum_{t=1}^{n} X_t X_{t-1} = -\beta\sigma^2 + O_p(n^{-\frac{1}{2}}).$$

Démonstration. On doit montrer que

$$E(X_t X_{t-1}) = -\beta\sigma^2,$$

et on là d'aprés (1.1.9) pour $h = 1$.

Et que

$$X_t X_{t-1} \sim N(-\beta\sigma^2, \frac{v_1}{n}).$$

Les X_t sont 1-dépendants donc d'aprés le T.C.L

$$\frac{1}{n}\sum_{t=1}^{n} X_t X_{t-1} \sim N(\nu(j), \frac{v_1}{n}).$$

Ici

$$\nu(j) = -\beta\sigma^2$$

Donc on a

$$\frac{1}{n}\sum_{t=1}^{n} X_t X_{t-1} = -\beta\sigma^2 + O_p(n^{-\frac{1}{2}}).$$

$\square$

3.3.1 Preuve du théorème

D'aprés le lemme 3.3.1 on a :

Supposons qu'à l'étape $j< i$ $(i > 1)$, l'éstimateur $\hat{\beta}^{(j)}$ admet la décomposition suivante

$$\hat{\beta}^{(j)} = \mu(j) + O_p(n^{-\frac{1}{2}}) pour\, j < i \tag{3.3.2}$$

où $\mu(j)$ est un terme non aléatoire.

On a

$$b_t^{(j)} = X_t + \hat{\beta}^{(i)} b_{t-1}^{(j-1)}, t = 1, ..., n. \tag{3.3.3}$$

Nous remplaçons $b_{t-1}^{(j-1)}$ dans cette expression et par itération nous avons,

$$
\begin{aligned}
b_t^{(j)} &= X_t + \hat{\beta}^{(i)}(X_{t-1} + \hat{\beta}^{(j-1)} b_{t-2}^{(j-2)}) \\
&= X_t + \hat{\beta}^{(j)} X_{t-1} + \hat{\beta}^{(j)}\hat{\beta}^{(j-1)} b_{t-2}^{(j-2)}) \\
&= X_t + \hat{\beta}^{(j)} X_{t-1} + \hat{\beta}^{(j)}\hat{\beta}^{(j-1)} X_{t-2} + \hat{\beta}^{(j)}\hat{\beta}^{(j-1)}\hat{\beta}^{(j-2)} b_{t-3}^{(j-3)} \\
&= . \\
&= . \\
&= . \\
&= X_t + \hat{\beta}^{(j)} X_{t-1} + \hat{\beta}^{(j)}\hat{\beta}^{(j-1)} X_{t-2} + ... + \hat{\beta}^{(j)}\hat{\beta}^{(j-1)}\hat{\beta}^{(j-j+1)} b_{t-j}^{(j-j)} \\
&= X_t + \hat{\beta}^{(j)} X_{t-1} + \hat{\beta}^{(j)}\hat{\beta}^{(j-1)} X_{t-2} + ... + \hat{\beta}^{(j)}\hat{\beta}^{(j-1)}\hat{\beta}^{(1)} b_{t-j}^{(0)},
\end{aligned}
$$

par l'expression (3.3.3) nous avons pour $j < i$

$$
b_t^{(j)} = X_t + \mu(j)X_{t-1} + \mu(j)\mu(j-1)X_{t-2} + ... + \mu(j)...\mu(2)X_{t-j+1} + \mu(j)...\mu(1)b_{t-j}^{(0)} + O_p(n^{-\frac{1}{2}}).
$$

Comme $\mu(1) = 0$ d'aprés le lemme 3.3.2, et $O_p(n^{-\frac{1}{2}})X_{t-1} = O_p(n^{-\frac{1}{2}})$
on a

$$
b_t^{(j)} = X_t + \mu(j)X_{t-1} + \mu(j)\mu(j-1)X_{t-2} + ... + \mu(j)...\mu(2)X_{t-j+1} + O_p(n^{-\frac{1}{2}}).
$$

Par suite

$$
b_{t-1}^{(i-1)} = X_{t-1} + \mu(i-1)X_{t-1} + \mu(i-1)\mu(i-2)X_{t-3} + ... + \mu(i-1)...\mu(2)X_{t-i+1} + O_p(n^{-\frac{1}{2}}). \tag{3.3.4}
$$

Multiplions l'expression (3.3.4) par X_t

$$
X_t b_{t-1}^{(i-1)} = X_t X_{t-1} + \mu(i-1)X_t X_{t-1} + \mu(i-1)\mu(i-2)X_t X_{t-3} + ... + \mu(i-1)...\mu(2)X_t X_{t-i+1} + O_p(n^{-\frac{1}{2}})
$$

Par conséquent

$$\begin{aligned} n^{-1}\sum_{t=1}^{n} X_t b_{t-1}^{(i-1)} &= n^{-1}\sum_{t=1}^{n} X_t X_{t-1} + \mu(i-1) n^{-1}\sum_{t=1}^{n} X_t X_{t-1} + \mu(i-1)\mu(i-2) n^{-1}\sum_{t=1}^{n} X_t X_{t-3} \\ ...+ &\ \mu(i-1)...\mu(2) n^{-1}\sum_{t=1}^{n} X_t X_{t-i+1} + O_p(n^{-\frac{1}{2}}). \qquad (3.3.6) \end{aligned}$$

Nous avons aussi pour $k \geq 2$,d'aprés le lemme 3.3.3

$$\frac{1}{n}\sum_{t=1}^{n} X_t X_{t-k} = O_p(n^{-\frac{1}{2}}),$$

et d'aprés le lemme 3.3.4

$$\frac{1}{n}\sum_{t=1}^{n} X_t X_{t-1} = -\beta\sigma^2 + O_p(n^{-\frac{1}{2}}),$$

par suite nous avons

$$n^{-1}\sum_{t=1}^{n} X_t b_{t-1}^{(i-1)} = -\beta\sigma^2 + O_p(n^{-\frac{1}{2}}). \qquad (3.3.7)$$

De la définition

$$b_t^{(i-1)2} = (X_t + \beta^{(i-1)} b_{t-1}^{(i-2)})^2.$$

Et par suite en sommant on a

$$\begin{aligned}
n^{-1}\sum_{t=1}^{n}(b_t^{(i-1)})^2 &= n^{-1}\sum_{t=1}^{n}(X_t^2+2X_t\hat{\beta}^{(i-1)}b_{t-1}^{(i-2)}+(\hat{\beta}^{i-1})^2(b_{t-1}^{(i-2)})^2)\\
&= n^{-1}\sum_{t=1}^{n}X_t^2+2\hat{\beta}^{(i-1)}\sum_{t=1}^{n}X_tb_{t-1}^{(i-2)}+(\hat{\beta}^{i-1})^2n^{-1}\sum_{t=1}^{n}(b_{t-1}^{(i-2)})^2)\\
&= (1+\beta^2)\sigma^2+2(\mu(i-1)+O_p(n^{-\frac{1}{2}}))n^{-1}\sum_{t=1}^{n}X_tb_{t-1}^{(i-2)}+(\hat{\beta}^{i-1})^2n^{-1}\sum_{t=1}^{n}(b_{t-1}^{(i-2)})^2\\
&= (1+\beta^2)\sigma^2+2\mu(i-1)n^{-1}\sum_{t=1}^{n}X_tb_{t-1}^{(i-2)}+O_p(n^{-\frac{1}{2}})+(\hat{\beta}^{i-1})^2n^{-1}\sum_{t=1}^{n}(b_{t-1}^{(i-2)})^2\\
&= (1+\beta^2)\sigma^2+2\mu(i-1)(-\beta\sigma^2+O_p(n^{-\frac{1}{2}}))+O_p(n^{-\frac{1}{2}})+(\mu(i-1)\\
&+ O_p(n^{-\frac{1}{2}}))\hat{\beta}^{i-1}n^{-1}\sum_{t=1}^{n}(b_{t-1}^{(i-2)})^2.
\end{aligned}$$

Donc

$$n^{-1}\sum_{t=1}^{n}(b_t^{(i-1)})^2=(1+\beta^2)\sigma^2-2\beta\mu(i-1)\sigma^2+\mu(i-1)\hat{\beta}^{(i-1)}n^{-1}\sum_{t=1}^{n}(b_{t-1}^{(i-2)})^2+O_p(n^{-\frac{1}{2}}). \tag{3.3.8}$$

Nous remplaçons la relation (3.3.7) dans l'expression (3.2.2) et nous avons

$$\hat{\beta}^{(i-1)}=-\frac{\sum_{t=1}^{n}X_tb_{t-1}^{(i-2)}}{\sum_{t=1}^{n}(b_{t-1}^{(i-2)})^2}$$

$$\hat{\beta}^{(i-1)}=\frac{\beta\sigma^2+O_p(n^{-\frac{1}{2}})}{n^{-1}\sum_{t=1}^{n}(b_{t-1}^{(i-2)})^2}$$

$$\hat{\beta}^{(i-1)}=\frac{\beta\sigma^2}{n^{-1}\sum_{t=1}^{n}(b_{t-1}^{(i-2)})^2}+O_p(n^{-\frac{1}{2}}).$$

Alors l'expression (3.3.8) devienne

$$
\begin{aligned}
n^{-1}\sum_{t=1}^{n}(b_t^{(i-1)})^2 &= (1+\beta^2)\sigma^2 - 2\beta\mu(i-1)\sigma^2 + \mu(i-1)\frac{\beta\sigma^2}{n^{-1}\sum\limits_{t=1}^{n}(b_{t-1}^{(i-2)})^2} n^{-1}\sum_{t=1}^{n}(b_{t-1}^{(i-2)})^2 + \\
&+ \mu(i-1)n^{-1}\sum_{t=1}^{n}(b_{t-1}^{(i-2)})^2 O_p(n^{-\frac{1}{2}}) + O_p(n^{-\frac{1}{2}}).
\end{aligned}
$$

On en déduit

$$
n^{-1}\sum_{t=1}^{n}(b_{t-1}^{(i-2)})^2 = (1+\beta^2-\beta\mu(i-1))\sigma^2 + O_p(n^{-\frac{1}{2}}). \tag{3.3.9}
$$

Nous remplaçons la relation (3.3.7) et (3.3.9) dans l'expression (3.2.2) et nous obtenons

$$
\hat{\beta}^{(i)} = \frac{\beta}{1+\beta^2-\beta\mu(i-1)} + O_p(n^{-\frac{1}{2}}).
$$

Ainsi pour $i > 1$, par la décomposition (3.3.2) on a

$$
\begin{aligned}
\mu(i) &= \hat{\beta}^{(i)} + O_p(n^{-\frac{1}{2}}) \\
&= \frac{\beta}{1+\beta^2-\beta\mu(i-1)} \\
&= \frac{\beta(1+\beta^2-\beta\mu(i-1)) - \beta(1+\beta^2-\beta\mu(i-1)) + \beta}{1+\beta^2-\beta\mu(i-1)} \\
&= \beta - \frac{\beta^3-\beta^2\mu(i-1)}{1+\beta^2-\beta\mu(i-1)} \\
&= \beta - \frac{\beta^3-\beta^2\frac{\beta}{1+\beta^2-\beta\mu(i-2)}}{1+\beta^2-\beta\frac{\beta}{1+\beta^2-\beta\mu(i-1)}} \\
&= \beta - \frac{\beta^5-\beta^4\mu(i-2)}{1+\beta^2+\beta^4-(\beta+\beta^3)\mu(i-2)} \\
&= \beta - \frac{\beta^7-\beta^6\mu(i-3)}{1+\beta^2+\beta^4+\beta^6-(\beta+\beta^3+\beta^5)\mu(i-3)} \\
&= \beta - \frac{\beta^{2i-1}-\beta^{2i-2}\mu(i-(i-1))}{1+\beta^2+...+\beta^{2i-2}-(\beta+...+\beta^{2i-3})\mu(i-(i-1))}.
\end{aligned}
$$

Donc

$$\mu(i) = \beta - \frac{\beta^{2i-1}}{1+\beta^2+...+\beta^{2i-2}}. \tag{3.3.10}$$

Et par suite nous avons

$$\mu(i) = \beta\frac{1-\beta^{2i-2}}{1-\beta^{2i}}.$$

En effet de la relation

$$1+\beta^2+...+\beta^{2i-2} = \frac{1-\beta^{2i}}{1-\beta^2}.$$

Car

$(1-\beta^2)(1+\beta^2+...+\beta^{2i-2}) = 1+\beta^2+...+\beta^{2i-2}-\beta^2-\beta^4...-\beta^{2i} = 1-\beta^{2i}.$

On a

$$
\begin{aligned}
\mu(i) &= \beta - \frac{\beta^{2i}-1}{\frac{1-\beta^{2i}}{1-\beta^2}} \\
&= \frac{\beta(1-\beta^{2i}) - \beta^{2i-1}(1-\beta^2)}{1-\beta^{2i}} \\
&= \frac{\beta-\beta^{2i+1}-\beta^{2i-1}+\beta^{2i+1}}{1-\beta^{2i}} \\
&= \frac{\beta(1-\beta^{2i-2})}{1-\beta^{2i}} \\
&= \beta\frac{1-\beta^{2i-2}}{1-\beta^{2i}}.
\end{aligned}
$$

De l'expression (3.3.2) nous avons

$$\hat{\beta}_{RIVE} = \lim_{i\to+\infty}(\mu(i)+O_p(n^{-\frac{1}{2}})).$$

Nous remplaçons $\mu(i)$ par son expression (3.3.10) et nous avons

$$\hat{\beta}_{RIVE} = \lim_{i\to+\infty}(\beta - \frac{\beta^{2i-1}}{1+\beta^2+...+\beta^{2i-2}} + O_p(n^{-\frac{1}{2}})).$$

Donc

$$\lim_{n\to+\infty}\hat{\beta}_{RIVE} \to \beta,$$

en probabilité .

d'où (i).

ii) Le remplacement de $\mu(i)$ dans l'expression (3.3.8) donne

$$
\begin{aligned}
n^{-1}\sum_{t=1}^{n}(b_t^{(i-1)})^2 &= [1+\beta^2-\beta(\beta-(\frac{\beta^{2i-1}}{1+\beta^2+...+\beta^{2i-2}}))]\sigma^2+O_p(n^{-\frac{1}{2}}) \\
&= [1+\frac{\beta^{2i}}{1+\beta^2+...+\beta^{2i-2}}]\sigma^2+O_p(n^{-\frac{1}{2}}).
\end{aligned}
$$

Par conséquent

$$
\begin{aligned}
\underset{n\to+\infty}{lim}(n^{-1}\sum_{t=1}^{n}b_t^2) &= \underset{n\to+\infty}{lim}(n^{-1}\sum_{t=1}^{n}(\underset{i\to+\infty}{lim}b_t^{(i)})^2) \\
&= \underset{n\to+\infty}{lim}(\underset{i\to+\infty}{lim}n^{-1}\sum_{t=1}^{n}(b_t^{(i)})^2),
\end{aligned}
$$

en probabilité.

$$
\underset{i\to+\infty}{lim}(n^{-1}\sum_{t=1}^{n}(b_t^2)=\underset{i\to+\infty}{lim}[1+\frac{\beta^{2i}}{1+\beta^2+...+\beta^{2i-2}}]\sigma^2,
$$

avec $|\beta|<1$

Et par suite

$$
\underset{n\to+\infty}{lim}\hat{\sigma}^2_{RIVE}=\sigma^2, \tag{3.3.11}
$$

en probabilité.

D'où (ii).

iii) Maintenant nous devons distinguer les deux cas $\beta\in[-1,+1[$ et $\beta=\pm1$.

- Premier cas $\beta\in]-1,1[$

Si nous appliquons l'expression (3.3.10) nous avons

$$
\begin{aligned}
\mu(i)\mu(i-1) &= \beta\frac{1-\beta^{2i-2}}{1-\beta^{2i}}\beta\frac{1-\beta^{2(i-1)-2}}{1-\beta^{2(i-1)}} \\
&= \beta^2\frac{1-\beta^{2i-4}}{1-\beta^{2i}}.
\end{aligned}
$$

En suite

$$
\begin{aligned}
\mu(i)\mu(i-1)\mu(i-2) &= \beta^2\frac{1-\beta^{2i-4}}{1-\beta^{2i}}\beta\frac{1-\beta^{2i-6}}{1-\beta^{2i-4}} \\
&= \beta^3\frac{1-\beta^{2(i-3)}}{1-\beta^{2i}}.
\end{aligned}
$$

Nous avons donc

$$
\mu(i)\mu(i-1)...\mu(i-j) = \beta^{j+1}\frac{1-\beta^{2(i-j-1)}}{1-\beta^{2i}}. \tag{3.3.12}
$$

Si nous remplaçons ces expressions successivement dans (3.3.4) nous obtenons

$$
\begin{aligned}
b_t^{(i)} &= X_t + \beta\frac{1-\beta^{2i-2}}{1-\beta^{2i}}X_{t-1} + \beta^2\frac{1-\beta^{2i-4}}{1-\beta^{2i}}X_{t-2} + ... + \beta^{i-2}\frac{1-\beta^4}{1-\beta^{2i}}X_{t-(i-2)} + O_p(n^{-\frac{1}{2}}) \\
&= \frac{1}{1-\beta^{2i}}\sum_{j=0}^{i-2}\beta^i(1-\beta^{2(i-j)})X_{t-j} + O_p(n^{-\frac{1}{2}}) \\
&= \frac{1}{1-\beta^{2i}}\sum_{j=0}^{i-2}\beta^i X_{t-j} - \frac{\beta^i}{1-\beta^{2i}}\sum_{j=0}^{i-2}\beta^{i-j}X_{t-j} + O_p(n^{-\frac{1}{2}}).
\end{aligned}
$$

Ainsi

$$
\begin{aligned}
b_t &= \underset{i\to+\infty}{lim} b_t^{(i)} \\
&= \underset{i\to+\infty}{lim}[\frac{1}{1-\beta^{2i}}\sum_{j=0}^{i-2}\beta^i X_{t-j} - \frac{\beta^i}{1-\beta^{2i}}\sum_{j=0}^{i-2}\beta^{i-j}X_{t-j} + O_p(n^{-\frac{1}{2}})] \\
&= \underset{i\to+\infty}{lim}(\frac{1}{1-\beta^{2i}}\sum_{j=0}^{i-2}\beta^i X_{t-j}) - \underset{i\to+\infty}{lim}(\frac{\beta^i}{1-\beta^{2i}}\sum_{j=0}^{i-2}\beta^{i-j}X_{t-j}) + O_p(n^{-\frac{1}{2}}) \\
&= \sum_{j=0}^{+\infty}\beta^i X_{t-j} + O_p(n^{-\frac{1}{2}}) \\
&= \sum_{j=0}^{+\infty}\beta^i(\varepsilon_{t-j} - \beta^i\varepsilon_{t-j-1}) + O_p(n^{-\frac{1}{2}}) \\
&= \sum_{j=0}^{+\infty}\beta^i\varepsilon_{t-j} - \sum_{j=0}^{+\infty}\beta^{j+1}\varepsilon_{t-j-1} + O_p(n^{-\frac{1}{2}}).
\end{aligned}
$$

Nous avons donc

$$b_t = \varepsilon_t + O_p(n^{-\frac{1}{2}}).$$

-Deuxième cas $\beta = \pm 1$.

Nous avons

$$\mu(i) = \frac{\beta}{1+\beta^2-\beta\mu(i-1)}.$$

Si nous remplaçons $\beta = \pm 1$ dans l'expression précedente nous avons

$$\mu(i) = \frac{1}{2-\mu(i-1)}.$$

Donc

$$\mu(2) = \frac{1}{2-\mu(1)}.$$

Et comme $\mu(1) = 0$ nous obtenons

$$\mu(2) = \frac{1}{2}.$$

De meme nous avons

$$\mu(3) = \frac{1}{2 - \mu(2)},$$

$$\mu(3) = \frac{2}{3}.$$

Nous montrons facilement par réccurence que

$$\mu(i) = \frac{i-1}{i},$$

pour $i > 1$.

Calculons maintenant

$$\begin{aligned} \mu(i)\mu(i-1)...\mu(i-j) &= (\frac{i-1}{i})(\frac{i-2}{i-1})...(\frac{i-j}{i-j+1})(\frac{i-j-1}{i-j}) \\ &= \frac{i-j-1}{i} \\ &= 1 - \frac{j+1}{i}. \end{aligned}$$

Pour $\beta = -1$ nous obtenons le même résultat.

Enfin pour $\beta = \pm 1$ nous obtenons

$$\mu(i)\mu(i-1)...\mu(i-j) = \beta^{j+1}(1 - \frac{j+1}{i}).$$

Ainsi

$$
\begin{aligned}
b_t^{(i)} &= \sum_{j=0}^{i-2} \beta^j (1-\frac{j}{i}) X_{t-j} + O_p(n^{-\frac{1}{2}}) \\
&= \sum_{j=0}^{i-2} \beta^j (1-\frac{j}{i})(\varepsilon_{t-j} - \beta\varepsilon_{t-j-1}) + O_p(n^{-\frac{1}{2}}) \\
&= \sum_{j=0}^{i-2} \beta^j (1-\frac{j}{i})\varepsilon_{t-j} - \sum_{j=0}^{i-2} \beta^{j+1} (1-\frac{j}{i})\varepsilon_{t-j-1} + O_p(n^{-\frac{1}{2}}) \\
&= \varepsilon_t - \sum_{j=0}^{i-1} \frac{\beta^j}{i} \varepsilon_{t-j} + O_p(n^{-\frac{1}{2}}).
\end{aligned}
$$

Or

$$
\begin{aligned}
b_t &= \lim_{i\to+\infty} b_t^{(i)} \\
&= \lim_{i\to+\infty} (\varepsilon_t - \sum_{j=0}^{i-1} \frac{\beta^j}{i} \varepsilon_{t-j} + O_p(n^{-\frac{1}{2}})) \\
&= \varepsilon_t + O_p(n^{-\frac{1}{2}}),
\end{aligned}
$$

car

$$
\sum_{j=0}^{i-1} \frac{\beta^j}{i} \varepsilon_{t-j} = O_p(n^{-\frac{1}{2}}).
$$

Ainsi dans les deux cas nous avons obtenu

$$
b_t = \varepsilon_t + O_p(n^{-\frac{1}{2}}).
$$

Donc

$$
\lim_{n\to+\infty} b_t \to \varepsilon_t,
$$

en probabilité.

d'où (iii) .

iv) En remplaçant l'estimateur $\hat{\beta}_{RIVE}$ par son expression (3.2.2) dans la formule suivante

$$\lim_{n\to+\infty}(\sqrt{n}(\hat{\beta}_{RIVE}-\beta)) = \lim_{n\to+\infty}(-\frac{n^{-\frac{1}{2}}\sum\limits_{t=1}^{n}X_t b_{t-1}^{(i-1)}}{n^{-1}\sum\limits_{t=1}^{n}(b_{t-1}^{(i-1)})^2} - \beta\frac{n^{-\frac{1}{2}}\sum\limits_{t=1}^{n}\varepsilon_{t-1}^2}{n^{-1}\sum\limits_{t=1}^{n}\varepsilon_{t-1}^2}).$$

Par le résultat (3.3.11) on tire qu'asymptotiquement la variable aléatoire $\sqrt{n}(\hat{\beta}_{RIVE}$-$\beta)$ à la même distribution que la variable aléatoire

$$-n^{-\frac{1}{2}}\sum_{t=1}^{n}\frac{X_t b_{t-1}}{\sigma^2} - \beta n^{-\frac{1}{2}}\sum_{t=1}^{n}\frac{\varepsilon_{t-1}^2}{\sigma^2},$$

puisque $(\frac{\varepsilon_t\varepsilon_{t-1}}{\sigma^2})$ est une suite strictement stationnaire 1-dépendante de la variables aléatoires de moyenne égale à zéro et variance égale à 1, par une version du théorème centrale limite, la variable aléatoire $n^{\frac{1}{2}}(\hat{\beta}_{RIVE}$-$\beta)$ est asymptotiquement distribuée comme $N(0,1)$.

D'où (iv).

Chapitre 4

L'estimateur du RIV (Cas vectoriel)

4.1 Introduction

Dans ce chapitre, on aborde le cas vectoriel. La méthode est itérative, elle consiste à estimer le paramètre du processus $MM(1)$ multivarié.

On considère le modèle suivant

$$Z_t = (I - \Theta B)a_t \tag{4.1.1}$$

Avec Z_t =$[Z_{1t}, Z_{2t}, ..., Z_{mt}]'$ un vecteur centré.

a_t=$[a_{1t}, a_{2t}, ..., a_{mt}]'$ un bruit blanc avec une matrice de covariance définie positive Σ.

B est l'opérateur de retard : $BZ_t = Z_{t-1}$.

Le paramètre à estimer :$\Theta = [\Theta_{jk}]_{j,k=1,2,...,m}$,

et I est une matrice d'identité.

Ce processus est inversible si les valeurs propres de Θ sont dans le disque unité.

Soient

$Z_1, Z_2, ..., Z_n$, n observations générées par (4.1.1) définies

$$Z = [Z_1', Z_2', ..., Z_n']'$$

et a =$[a_1', a_2', ..., a_n']'$
et A_{-1}= BA.
donc le modèle devient

$$Z = A - A_{-1}\Theta'$$

Si A_{-1} était observable, alors Θ peut être estimée par la méthode des moindres carrés .

4.2 Principe de la méthode RIV

Soient
$Z_1, Z_2, ..., Z_n, n$ observations générées par (4.1.1) définies

$$Z = [Z_1', Z_2', ..., Z_n']'$$

et A =$[a_1', a_2', ..., a_n']'$
et A_{-1}= BA.
donc le modèle devient

$$Z = A - A_{-1}\Theta'.$$

On génère m variables aléatoires indépendantes par la série du bruit blanc $\{e_{0,1}^{(0)}, ..., e_{n-1,1}^{(0)}\}, ..., \{e_{0,m}^{(0)}, ..., e_{n-1,m}^{(0)}\}$.

Soit $E_{-1}^{(0)}$= $[e_{t,k}^{(0)}]_{t=0,...,n-1;k=1,2,...,m}$, avec $e_{t,k}^{(0)}$ sont tirées d'une distribution centrée, réduite .

Après avoir défini le nombre d'itération $i = 1$, les étapes suivantes sont répétées jusqu'à la convergence des éléments de $\hat{\Theta}^{(i)}$(ç-a-d $\underset{j,k}{max}|\hat{\Theta}_{jk}^{(i)} - \hat{\Theta}_{jk}^{(i-1)}| <$

ε avec $\varepsilon > 0$, fixé.

On a les observations Z et les $E_{-1}^{(0)}$, on veut avoir $\Theta'^{(1)}$,

$$Z = E^{(i-1)} - E_{-1}^{(i-1)}\Theta'^{(i)}.$$

Et puis pour $i = 1$

$$Z = E^{(0)} - E_{-1}^{(0)}\Theta'^{(1)},$$

Or on prend $E^{(0)} = 0$;
donc

$$Z = -E_{-1}^{(0)}\Theta'^{(1)},$$

$$Z' = -\Theta^{(1)}E_{-1}'^{(0)},$$

On multiplie par $E_{-1}^{(0)}$

$$-Z'E_{-1}^{(0)} = \Theta^{(1)}E_{-1}'^{(0)}E_{-1}^{(0)},$$

$$\hat{\Theta}^{(1)}(E_{-1}'^{(0)}E_{-1}^{(0)})(E_{-1}'^{(0)}E_{-1}^{(0)})^{-1} = -Z'E_{-1}^{(0)}(E_{-1}'^{(0)}E_{-1}^{(0)})^{-1},$$

et enfin

$$\hat{\Theta}^{(1)} = -Z'E_{-1}^{(0)}(E_{-1}'^{(0)}E_{-1}^{(0)})^{-1}.$$

Aprés avoir $\hat{\Theta}^{(1)}$, maintenant on veut avoir $E^{(1)}$

$$E^{(1)} = Z + E_{-1}^{(0)}\Theta'^{(1)}$$

$$E^{(2)} = Z + E_{-1}^{(1)}\Theta'^{(2)}$$

$$\vdots$$

$$E^{(i)} = Z + E_{-1}^{(i-1)}\Theta'^{(i)} \tag{4.2.1}$$

On multiplie (4.2.1) par $E_{-1}^{(i-1)}$;

$$E^{(i)}E_{-1}'^{(i-1)} = Z'E_{-1}^{(i-1)} + E_{-1}^{(i-1)}\Theta'^{(i)}E_{-1}'^{(i-1)};$$

On passe à l'espérance ;
on a $E(E^{(i)}E_{-1}'^{(i-1)}) = 0$.
Et puis

$$0 = \frac{1}{n}Z'E_{-1}^{(i-1)} + \frac{1}{n}E_{-1}^{(i-1)}\Theta'^{(i)}E_{-1}'^{(i-1)}.$$

Donc

$$\frac{1}{n}Z'E_{-1}^{(i-1)} = -\frac{1}{n}E_{-1}^{(i-1)}\Theta'^{(i)}E_{-1}'^{(i-1)}.$$

Et enfin,

$$\hat{\Theta}^{(i)\prime} = -\frac{1}{n}Z'E_{-1}^{(i-1)}(E_{-1}'^{(i-1)}E_{-1}^{(i-1)})^{-1}, \tag{4.2.2}$$

avec

$$E_{-1}^{(i-1)} = (e_{t,k}^{(i-1)})_{t=1,\dots,n;k=1,\dots,m}.$$

Et $E_{-1}'^{(i-1)}$ la transposé.

$$Z' = (Z_1', Z_2', \dots, Z_n'),$$

et $\hat{\Theta}^{(i)\prime}$ matrice .

On répète cette opération jusqu'à

$$\underset{j,k}{max}|\hat{\Theta}_{jk}^{(i)} - \hat{\Theta}_{jk}^{(i-1)}| < \varepsilon.$$

Le $\hat{\Theta}_{RIV}$ est donné par,

$$\hat{\Theta}_{RIV} = \underset{i\to\infty}{lim}\hat{\Theta}^{(i)},$$

et

$$\hat{\Sigma}^{(i)} = \frac{1}{n}E_{-1}^{\prime(i)}E_{-1}^{(i)},$$

et par conséquent Σ est estimé par ,

$$\hat{\Sigma} = \underset{i\to\infty}{lim}\hat{\Sigma}^{(i)}.$$

Dans cette section, nous étudions des propriétés asymptotiques de l'estimateur, les résultats sont exprimés par le théorème suivant :

4.3 Théorème

Si Z_t est un processus $MM(1)$ multivarié inversible alors

(i) $\lim \hat{\Theta} = \Theta$;

(ii) $\lim \hat{\Sigma} = \Sigma$.

Pour la démonstration de ce théorème, nous avons besoin des lemmes suivants.

Lemme 4.3.1. *On a pour $j < i$ et $i > 1$ l'éstimateur $\hat{\Theta}^{(j)}$ admet la décomposition suivante ,*

$$\hat{\Theta}^{(j)} = \nu(j) + O_p(n^{-\frac{1}{2}}).$$

Démonstration. Pour prouver cela il faut montrer que $E(\hat{\Theta}^{(j)}) = \nu(j)$ et que $\hat{\beta}^{(j)} \sim N(\nu(j), \frac{v_1}{n})$

On sait que

$$\hat{\Theta}^{(i)\prime} = -\frac{1}{n} Z' E_{-1}^{(i-1)} (E_{-1}^{\prime(i-1)} E_{-1}^{(i-1)})^{-1},$$

On pose

$$Y_t = Z' E_{-1}^{(i-1)}.$$

On passe à l'éspérance de Y_t

$$\begin{aligned} E(Y_t) &= E(Z' E_{-1}^{(i-1)}) \\ &= \nu(j) \end{aligned}$$

Y_t et Y_{t+1} ne sont pas indépendants.

Car

$$E(Y_t Y_{t+1}) \neq 0.$$

$$(Y_t = Z' E_{-1}^{(i-1)}, Y_{t+1} = Z' E_{-1}^{(i-1)})$$

$$Z_t = a_t - \Theta a_{t-1}, Z_t = (a_t, \Theta, a_{t-1}),$$

et

$$Z_{t+1} = a_{t+1} - \Theta a_t; Z_{t+1} = (a_{t+1}, \Theta, a_t),$$

donc Z_t et Z_{t+1} sont 1-dépendants.

Donc les Y_t sont 1-dépendants, et par le théorème T.C.L 2.1.2 du deuxième chapitre

$$\frac{1}{n} \sum_{t=1}^{n} Y_t \sim N(\nu(j), \frac{v_1}{n}).$$

On a

$$\begin{aligned}\hat{\Theta}^{(i)\prime}(E_{-1}^{\prime(i-1)}E_{-1}^{(i-1)}) &= -\frac{1}{n}Z'E_{-1}^{(i-1)}\\ &= \frac{1}{n}\sum_{t=1}^{n}Y_t,\end{aligned}$$

et $\frac{1}{n}\sum\limits_{t=1}^{n}Y_t \sim N(\nu(j), \frac{v_1}{n})$.

Donc $\hat{\Theta}^{(i)\prime}(E_{-1}^{\prime(i-1)}E_{-1}^{(i-1)}) \sim N(\nu(j), \frac{v_1}{n})$.

Par suite

$$\hat{\Theta}^{(j)} = \nu(j) + O_p(n^{-\frac{1}{2}}).$$

$\square$

Lemme 4.3.2. *On a*

$$\frac{1}{n}Z'E_{-2}^{(i-2)} = O_p(n^{-\frac{1}{2}}).$$

Démonstration. On a

$$E_{-2}^{(i-2)} = BE_{-1}^{(i-1)},$$

et

$$Z = A - A_{-1}\Theta'; Z' = A' - \Theta A'_{-1}.$$

$$\begin{aligned}\frac{1}{n}Z'E_{-2}^{(i-2)} &= \frac{1}{n}[Z'(Z_{-2} + E_{-2}^{(i-3)}\hat{\Theta}'^{(i-2)}]\\ &= \frac{1}{n}Z'Z_{-2} + \frac{1}{n}Z'E_{-2}^{(i-3)}\hat{\Theta}'^{(i-2)},\end{aligned}$$

avec $Z' = A' - \Theta A'_{-1}$ et $Z_{-2} = A_{-2} - A_{-3}\Theta'$

$$\begin{aligned}
\frac{1}{n}Z'Z_{-2} &= \tfrac{1}{n} \quad (A' - \Theta A'_{-1})(A_{-2} - A_{-3}\Theta') \\
&= \quad \frac{1}{n}A'A_{-2} - \frac{1}{n}A'A_{-3}\Theta' - \frac{1}{n}\Theta A'_{-1}A_{-2} + \frac{1}{n}\Theta A'_{-1}A_{-3}\Theta' \\
&= \quad \frac{1}{n}A'A_{-2} - \frac{1}{n}A'A_{-3}\Theta' - \frac{1}{n}\Theta A'_{-1}A_{-2} + \Theta A'_{-1}A_{-3}\Theta' \\
&= \quad 0,
\end{aligned}$$

et

$$\begin{aligned}
\frac{1}{n}Z'E_{-2}^{(i-3)}\hat{\Theta}'^{(i-2)} &= \frac{1}{n}Z'(Z_{-2} + E_{-2}^{(i-4)}\hat{\Theta}'^{(i-3)})\hat{\Theta}'^{(i-2)} \\
&= \frac{1}{n}Z'Z_{-2}\hat{\Theta}'^{(i-2)} + \frac{1}{n}Z'E_{-2}^{(i-4)}\hat{\Theta}'^{(i-3)}\hat{\Theta}'^{(i-2)},
\end{aligned}$$

or on a

$$\frac{1}{n}Z'Z_{-2}\hat{\Theta}'^{(i-2)} = 0.$$

Encore une fois

$$\begin{aligned}
\frac{1}{n}Z'E_{-2}^{(i-4)} &= \frac{1}{n}Z'(Z_{-2} + E_{-2}^{(i-5)}\hat{\Theta}'^{(i-4)}) \\
&= \frac{1}{n}Z'Z_{-2} + ...
\end{aligned}$$

ainsi de suite, donc on aura

$$\frac{1}{n}Z'E_{-2}^{(i-2)} = O_p(n^{-\frac{1}{2}}). \tag{4.3.1}$$

□

Affirmation 4.3.3. On a la formule suivante $\forall C, D$ deux matrices inversibles :

$$(C + D)^{-1} = C^{-1} - C^{-1}(C^{-1} + D^{-1})^{-1}C^{-1}.$$

Démonstration. On a

$$(C+D)^{-1} = C^{-1} - C^{-1}(C^{-1}+D^{-1})^{-1}C^{-1},$$

on multiplie des deux cotés par $(C+D)$

$$\begin{aligned}
(C+D)^{-1}(C+D) &= C^{-1} - C^{-1}(C^{-1}+D^{-1})^{-1}C^{-1}(C+D) \\
I &= C^{-1} - C^{-1}(C^{-1}+D^{-1})^{-1}C^{-1}(C+D)
\end{aligned}$$

et donc

$$\begin{aligned}
C^{-1} - C^{-1}(C^{-1}+D^{-1})^{-1}C^{-1}(C+D) &= I - C^{-1}(C^{-1}+D^{-1})^{-1} + C^{-1}D - C^{-1}(C^{-1}+D^{-1} \\
&= (I + C^{-1}D) - C^{-1}(C^{-1}+D^{-1})^{-1}[I + C^{-1}D] \\
&= (I + C^{-1}D)[I - ((C^{-1}+D^{-1})C)^{-1}] \\
&= (I + C^{-1}D)[I - ((D^{-1}CC^{-1}D + D^{-1}C)^{-1}] \\
&= (I + C^{-1}D)[I - (I + C^{-1}D)^{-1}C^{-1}D] \\
&= I
\end{aligned}$$

et enfin

$$C^{-1} - C^{-1}(C^{-1}+D^{-1})^{-1}C^{-1}(C+D) = I.$$

Inversement

$$\begin{aligned}
(C+D)(C^{-1}-C^{-1}(C^{-1}+D^{-1})^{-1}C^{-1}) &= CC^{-1}+DC^{-1}-(C^{-1}+D^{-1})^{-1}C^{-1}-DC^{-1}(C \\
&= (I+DC^{-1})-(I+DC^{-1})(C^{-1}+D^{-1})^{-1}C^{-1} \\
&= (I+DC^{-1})[I-(C(C^{-1}+D^{-1}))^{-1}] \\
&= (I+DC^{-1})[I-(CD^{-1}DC^{-1}+CD^{-1})^{-1}] \\
&= (I+DC^{-1})[I-(I+DC^{-1})^{-1}DC^{-1}] \\
&= I
\end{aligned}$$

$\square$

Lemme 4.3.4. *On a*

$$R^{(i-1)-1} = W_1 + W_2 + ... + W_{i-2},$$

avec W_1=$(\Theta')^{-1}\Sigma^{-1}\Theta^{-1}$ *;*

W_2=$(\Theta')^{-1}(\Theta')^{-1}\Sigma^{-1}\Theta^{-1}\Theta^{-1}$ *;*

W_3=$(\Theta')^{-1}(\Theta')^{-1}(\Theta')^{-1}\Sigma^{-1}\Theta^{-1}\Theta^{-1}\Theta^{-1}$;

Démonstration. On a

$$R^{(i)-1} = (\Theta')^{-1}(\Sigma^{-1} + R^{(i-1)-1})\Theta^{-1}.$$

Or $W_1 = R^{(2)-1}$

par reccurence, on suppose que cette formule est vraie a l'ordre i-1 et on la montre à l'ordre i

$$R^{(i-1)-1} = W_1 + W_2 + ... + W_{i-2},$$

et $R^{(i)-1} = W_1 + W_2 + ... + W_{i-1}$,
or

$$R^{(i)} = \Theta(\Sigma^{-1} + R^{(i-1)-1})^{-1}\Theta',$$

et

$$\begin{aligned} R^{(i)-1} &= (\Theta')^{-1}(\Sigma^{-1} + R^{(i-1)-1})\Theta^{-1} \\ &= (\Theta')^{-1}\Sigma^{-1}\Theta^{-1} + (\Theta')^{-1}R^{(i-1)-1}\Theta^{-1} \\ &= W_1 + (\Theta')^{-1}(W_1 + W_2 + ... + W_{i-2})\Theta^{-1} \\ &= W_1 + W_2 + ... + W_{i-1}. \end{aligned}$$

D'où le resultat. □

Lemme 4.3.5. *On a $R^{(i)}$ s'ecrit sous cette forme*

$$R^{(i)} = W_{i-1}^{-1} - W_{i-1}^{-1}\{I + (W_{i-1}^{-1}W_1 + W_{i-1}^{-1}W_2 + ... + W_{i-1}^{-1}W_{i-2})^{-1}\}^{-1}.$$

Proof. On a d'aprés le lemme (4.3.4)

$$R^{(i)-1} = R^{(i-1)-1} + W_{i-1},$$

donc

$$R^{(i)} = (R^{(i-1)-1} + W_{i-1})^{-1},$$

et puis on applique l'affirmation 4.3.3 on aura

$$(R^{(i-1)-1} + W_{i-1})^{-1} = W_{i-1}^{-1} - W_{i-1}^{-1}(W_{i-1}^{-1} + R^{(i-1)-1})^{-1}W_{i-1}^{-1}.$$

On remplace $R^{(i-1)}$ par

$$R^{(i-1)-1} = W_1 + W_2 + ... + W_{i-2},$$

daprés le lemme 4.3.4, donc

$$\begin{aligned}(R^{(i-1)-1} + W_{i-1})^{-1} &= W_{i-1}^{-1} - W_{i-1}^{-1}(W_{i-1}^{-1} + (W_1 + W_2 + ... + W_{i-2})^{-1})^{-1}W_{i-1}^{-1}. \\ &= W_{i-1}^{-1} - W_{i-1}^{-1}(I + (W_{i-1}^{-1}W_1 + W_{i-1}^{-1}W_2 + ... + W_{i-1}^{-1}W_{i-2})^{-1})^{-1}.\end{aligned}$$

□

Lemma 4.3.6. *On a*

$$R^{(i)} = W_{i-1}^{-1}[I - \{W_{i-1}^{-1} + (W_1 + W_2 + ... + W_{i-2})^{-1}\}^{-1}W_{i-1}^{-1}].$$

Proof. D'aprés le lemme 4.3.5 , on a

$$\begin{aligned}R^{(i)} &= W_{i-1}^{-1} - W_{i-1}^{-1}(I + (W_{i-1}^{-1}W_1 + W_{i-1}^{-1}W_2 + ... + W_{i-1}^{-1}W_{i-2})^{-1})^{-1} \\ &= W_{i-1}^{-1} - W_{i-1}^{-1}(I + W_{i-1}(W_1 + W_2 + ... + W_{i-2})^{-1})^{-1} \\ &= W_{i-1}^{-1} - W_{i-1}^{-1}(W_{i-1}^{-1}W_{i-1} + W_{i-1}(W_1 + W_2 + ... + W_{i-2})^{-1})^{-1} \\ &= W_{i-1}^{-1} - W_{i-1}^{-1}(W_{i-1}\{W_{i-1}^{-1} + (W_1 + W_2 + ... + W_{i-2})^{-1}\})^{-1} \\ &= W_{i-1}^{-1} - W_{i-1}^{-1}(\{W_{i-1}^{-1} + (W_1 + W_2 + ... + W_{i-2})^{-1}\}^{-1}W_{i-1}^{-1}) \\ &= W_{i-1}^{-1}[I - \{W_{i-1}^{-1} + (W_1 + W_2 + ... + W_{i-2})^{-1}\}^{-1}W_{i-1}^{-1}].\end{aligned}$$

□

4.3.1 Preuve du théorème

D'aprés le lemme 4.3.1 on a

Supposons qu'à l'étape $j < i$ et $(i > 1)$, l'éstimateur $\hat{\Theta}^{(j)}$ admet la décomposition suivante :

$$\hat{\Theta}^{(j)} = \mu(j) + O_p(n^{-\frac{1}{2}}), \tag{4.3.2}$$

pour $j < i$,

avec $\mu(1) = 0$ et $\mu(j) = O(1)$.

Soit

$$\hat{\Theta}^{(i)\prime} = -\frac{1}{n} Z' E_{-1}^{(i-1)} (E_{-1}^{\prime(i-1)} E_{-1}^{(i-1)})^{-1},$$

et $Z_{-1} = BZ$.

et $E_{-2}^{(i-2)} = BE_{-1}^{(i-1)}$.

Et

$$\begin{aligned} \frac{1}{n} Z' E_{-1}^{(i-1)} &= \frac{1}{n} Z'(Z_{-1} + E_{-2}^{(i-2)} \hat{\Theta}^{\prime(i-1)}) \\ &= \frac{1}{n} Z' Z_{-1} + \frac{1}{n} Z' E_{-2}^{(i-2)} \hat{\Theta}^{\prime(i-1)}, \end{aligned}$$

et d'aprés le lemme 4.3.2

$$\frac{1}{n} Z' E_{-2}^{(i-2)} = O_p(n^{-\frac{1}{2}}),$$

et

$$\frac{1}{n} Z' Z_{-1} = -\Theta\Sigma + O_p(n^{-\frac{1}{2}}),$$

car

$$
\begin{aligned}
\frac{1}{n}Z'Z_{-1} &= \frac{1}{n}(A' - \Theta A'_{-1})(A_{-1} - A_{-2}\Theta') \\
&= \frac{1}{n}A'A_{-1} - \frac{1}{n}A'A_{-2}\Theta' - \frac{1}{n}\Theta A'_{-1}A_{-1} + \frac{1}{n}\Theta A'_{-1}A_{-2}\Theta' \\
&= 0 - 0 - \Theta\Sigma + 0 \\
&= -\Theta\Sigma + O_p(n^{-\frac{1}{2}}).
\end{aligned}
$$

Donc

$$
\frac{1}{n}Z'E_{-1}^{(i-1)} = -\Theta\Sigma + O_p(n^{-\frac{1}{2}}). \tag{4.3.3}
$$

On a aussi

$$
\begin{aligned}
\frac{1}{n}E'^{(j)}_{-1}E^{(j)}_{-1} &= \frac{1}{n}(Z_{-1} + E^{(j-1)}_{-2}\hat{\Theta}^{(j)})'(Z_{-1} + E^{(j-1)}_{-2}\hat{\Theta}^{(j)}) \\
&= \frac{1}{n}Z'_{-1}Z_{-1} + \frac{1}{n}Z'_{-1}E^{(j-1)}_{-2}\hat{\Theta}^{(j)} + \frac{1}{n}\hat{\Theta}^{(j)}E'^{(j-1)}_{-2}Z_{-1} + \frac{1}{n}\hat{\Theta}^{(j)}E'^{(j-1)}_{-2}E^{(j-1)}_{-2}\hat{\Theta}^{(j)}) \\
&= \frac{1}{n}Z'_{-1}Z_{-1} + \frac{1}{n}\hat{\Theta}^{(j)}E'^{(j-1)}_{-2}Z_{-1} + O_p(n^{-\frac{1}{2}}),
\end{aligned}
$$

car

$$
\frac{1}{n}Z'_{-1}E^{(j-1)}_{-2} = O_p(n^{-\frac{1}{2}}),
$$

et $\frac{1}{n}\hat{\Theta}^{(j)}E'^{(j-1)}_{-2}E^{(j-1)}_{-2}\hat{\Theta}^{(j)} = O_p(n^{-\frac{1}{2}})$.
On a aussi

$$\begin{aligned}
\frac{1}{n}Z'_{-1}Z_{-1} &= \frac{1}{n}(A'_{-1} - \Theta A'_{-2})(A_{-1} - A_{-2}\Theta') \\
&= \tfrac{1}{n} \quad A'_{-1}A_{-1} - \frac{1}{n}A'_{-1}A_{-2}\Theta' - \frac{1}{n}\Theta A'_{-2}A_{-1} + \frac{1}{n}\Theta A'_{-2}A_{-2}\Theta' \\
&= \Sigma - 0 - 0 + \Theta\Sigma\Theta'. \\
&= \Sigma + \Theta\Sigma\Theta' + O_p(n^{-\frac{1}{2}}).
\end{aligned}$$

Et enfin

$$\frac{1}{n}E'^{(j)}_{-1}E^{(j)}_{-1} = \Sigma + \Theta\Sigma\Theta' + \hat{\Theta}^{(j)}\frac{1}{n}E'^{(j-1)}_{-2}Z_{-1} + O_p(n^{-1})$$

et

$$\frac{1}{n}E'^{(j-1)}_{-2}Z_{-1} = \Sigma\Theta' + O_p(n^{-\frac{1}{2}}),$$

car

$$\begin{aligned}
\frac{1}{n}E'^{(j-1)}_{-2}Z_{-1} &= \frac{1}{n}(Z_{-2} + E^{(j-2)}_{-3}\hat{\Theta}'^{(j-1)})'Z_{-1} \\
&= \frac{1}{n}Z'_{-2}Z_{-1} + \frac{1}{n}\hat{\Theta}^{(j-1)}E'^{(j-2)}_{-3}Z_{-1} \\
&= \Sigma\Theta' + O_p(n^{-\frac{1}{2}}).
\end{aligned}$$

Enfin

$$\frac{1}{n}E'^{(j)}_{-1}E^{(j)}_{-1} = \Sigma + \Theta\Sigma\Theta' + \hat{\Theta}^{(j)}\Sigma\Theta' + O_p(n^{-\frac{1}{2}}), \tag{4.3.4}$$

et

$$\frac{1}{n}E'^{(1)}_{-1}E^{(1)}_{-1} = \Sigma + \Theta\Sigma\Theta' + \hat{\Theta}^{(1)}\Sigma\Theta' + O_p(n^{-\frac{1}{2}}).$$

Or

$$\hat{\Theta}^{(1)} = \mu(1) + O_p(n^{-\frac{1}{2}}),$$

et $\mu(1) = 0$ donc $\hat{\Theta}^{(1)} = O_p(n^{-\frac{1}{2}})$
Et enfin

$$\frac{1}{n} E'^{(1)}_{-1} E^{(1)}_{-1} = \Sigma + \Theta\Sigma\Theta' + O_p(n^{-\frac{1}{2}}).$$

Et pour $j > 1$,
on a

$$\frac{1}{n} E'^{(j)}_{-1} E^{(j)}_{-1} = \Sigma + \Theta\Sigma\Theta' + \hat{\Theta}^{(j)}\Sigma\Theta' + O_p(n^{-\frac{1}{2}}).$$

Or

$$\hat{\Theta}^{(j)} = \mu(j) + O_p(n^{-\frac{1}{2}}).$$

On prend

$$\mu(j) = \Theta(I + R^{(j)}\Sigma^{-1})^{-1} = \Theta\Sigma(\Sigma + R^{(j)})^{-1}, \tag{4.3.5}$$

car

$$\begin{aligned}
\Theta\Sigma(\Sigma + R^{(j)})^{-1} &= \Theta\Sigma(\Sigma + R^{(j)}\Sigma^{-1}\Sigma)^{-1} \\
&= \Theta\Sigma(I + R^{(j)}\Sigma^{-1}(\Sigma))^{-1} \\
&= \Theta\Sigma\Sigma^{-1}(I + R^{(j)}\Sigma^{-1})^{-1} \\
&= \Theta(I + R^{(j)}\Sigma^{-1})^{-1}.
\end{aligned}$$

Alors pour $j > 1$, $R^{(2)} = \Theta\Sigma\Theta'$, et pour $j > 2$, on a

$$\begin{aligned} R^{(j)} &= \Theta(\Sigma^{-1} + R^{(j-1)-1})^{-1}\Theta' \qquad (4.3.6) \\ &= \Theta\{\Sigma - \Sigma(\Sigma + R^{(j-1)})^{-1}\Sigma\}\Theta' \qquad (4.3.7) \end{aligned}$$

On applique l'affirmation à (4.3.6)
avec $C = \Sigma^{-1}$ et $D = R^{(j-1)\text{-}1}$
donc

$$(\Sigma^{-1} + R^{(j-1)-1})^{-1} = \Sigma - \Sigma(\Sigma + R^{(j-1)})^{-1}\Sigma.$$

D'ou la formule (4.3.7).
On a toujours

$$\hat{\Theta}^{(j)} = \mu(j) + O_p(n^{-\frac{1}{2}}),$$

avec

$$\mu(j) = \Theta\Sigma(\Sigma + R^{(j)})^{-1},$$

donc$\hat{\Theta}^{(j)}$ devient

$$\hat{\Theta}^{(j)} = \Theta\Sigma(\Sigma + R^{(j)})^{-1} + O_p(n^{-\frac{1}{2}})$$

et puis on remplaçant $\hat{\Theta}^{(j)}$, (4.3.4) devient

$$\begin{aligned} \frac{1}{n}E'^{(j)}_{-1}E^{(j)}_{-1} &= \Sigma + \Theta\Sigma\Theta' - (\Theta\Sigma(\Sigma + R^{(j)})^{-1} + O_p(n^{-\frac{1}{2}}))\Sigma\Theta' + O_p(n^{-\frac{1}{2}}) \\ &= \Sigma + \Theta\Sigma\Theta' - \Theta\Sigma(\Sigma + R^{(j)})^{-1}\Sigma\Theta' + O_p(n^{-\frac{1}{2}}). \end{aligned}$$

Or

$$R^{(j+1)} = \Theta\Sigma\Theta' - \Theta\Sigma(\Sigma + R^{(j)})^{-1}\Sigma\Theta'.$$

D'où

$$\frac{1}{n}E'^{(j)}_{-1}E^{(j)}_{-1} = \Sigma + R^{(j+1)} + O_p(n^{-\frac{1}{2}}). \tag{4.3.8}$$

Donc on remplace enfin (4.3.3) et (4.3.8) dans $\hat{\Theta}^{(j)}$, on aura

$$\hat{\Theta}^{(j)} = \Theta\Sigma + O_p(n^{-\frac{1}{2}})(\Sigma + R^{(j+1)} + O_p(n^{-\frac{1}{2}}))^{-1}.$$

On a $\underset{i\to\infty}{lim} R^{(i)} \to 0$.
car

$$R^{(i)} = \Theta(\Sigma^{-1} + R^{(i-1)-1})^{-1}\Theta',$$

et puis

$$R^{(i)-1} = (\Theta')^{-1}(\Sigma^{-1} + R^{(i-1)-1})\Theta^{-1}.$$

Et les $R^{(i-1)-1}$ s'ecrivent en fonction de W_i d'aprés le lemme(4.3.6),
or

$$W_i = (\Theta')^{-1}\underset{i\,fois}{...}(\Theta')^{-1}\Sigma^{-1}\Theta^{-1}\underset{i\,fois}{...}\Theta^{-1},$$

et par conséquent

$$W_i^{-1} = \Theta\underset{i\,fois}{...}\Theta\Sigma\Theta'\underset{i\,fois}{...}\Theta',$$

donc $\underset{i\to\infty}{lim} W_i^{-1} = 0$
et comme le processus Z est inversible, les valeurs propres de Θ sont dans le disque unité.

Puis, comme $\underset{i\to\infty}{lim} W_i^{-1} = 0$, $R^{(i)} \to 0$ (le nombre d'ittération augmente).
Et quand $R^{(i)} \to 0$, on a

$$\underset{i\to\infty}{lim}\hat{\Theta}^{(i)} = \Theta.$$

Et on a aussi

$$\lim_{i\to\infty}\hat{\Sigma}^{(i)} = \lim_{i\to\infty}\frac{1}{n}E'^{(j)}_{-1}E^{(j)}_{-1},$$

et d'aprés (4.3.8) on a

$$\frac{1}{n}E'^{(j)}_{-1}E^{(j)}_{-1} = \Sigma + R^{(j+1)} + O_p(n^{-\frac{1}{2}}),$$

donc

$$\lim_{i\to\infty}\hat{\Sigma}^{(i)} = \lim_{i\to\infty}(\Sigma + R^{(j+1)} + O_p(n^{-\frac{1}{2}})),$$

et comme $R^{(i)}\to 0$ quand $i\to\infty$

Conclusion

Notre travail est une présentation des processus moyenne mobile en dimension finie, nous nous sommes intéressés aux principales méthodes d'estimation d'un moyenne mobile et en particulier à la méthode RIV proposée par Anna Clara Monti dans les cas scalaire et multivarié [1] et [2].

Cette méthode a été généralisé par Turbillon, C dans sa thèse de Doctorat [5]

Bibliographie

[1] Anna Clara Monti, "A new preliminary estimator for MA(1) models".In computational Statistics Data Analysis.21(1995) p.1-15.

[2] Anna Clara Monti, "A new preliminary estimator for $MA(1)$ models " in Journal of time series analysis, vol 19, No 2 (1998) p.210-219.

[3] Durbin, J.(1959)."Efficient Estimation of Parameters in Moving Average Models".Biometrika, 46 (1959) p. 306-313.

[4] Denise R.Osborn (1976).Maximum likehood estimation of moving average processes .Annals of Economic and Social Measurement 5/1,1976.

[5]Thèse de Doctorat Turbillon, C [2007], estimation et prévision des processus moyenne mobile fonctionnels .

[6] Brockwell, p, J and R.A.Davis."Time series: Théory and methods". (Springer, New York, 1987).

[7] Box, G.E.P.and G.M.Jenkins. "Time series analysis" : forecasting and control (Holden Day, San Franscisco.1970. revisededition. 1976).

[8] Fuller , W.A."Introduction to statistical time series " (Wiley, New York, 1976) .

[9] Magistère Mme SABRI Khadija

[10] Magistère Mr AZZOUZI Badreddin