Table des Matières

Remerciements

Nous avons préparé notre mémoire sous la direction de Monsieur le Professeur T.Mourid à qui vont mes remerciments les meilleurs et à qui nous exprimons notre profonde reconnaissance pour avoir diriger nos recherches.

Nous remercions très vivement Monsieur J.Benmansour qui a bien voulu examiner et accepter la présidence de notre jury.

Nos sincères remerciments vont également à Monsieur G.Senouci Berekci et Monsieur A.Labbas et Monsieur F. Boukhari qui ont bien voulu examiner notre travail.

Enfin, nous remercions toutes les personnes qui nous ont aidé à réaliser ce travail.

Introduction

Les processus autorégressifs ont beaucoup d'interêts dans plusieurs domaines: économie (marchés financiers), en biologie, en physique, en medecine...ect.

Les problemes statistiques concernant cette classe de processus sont essentiellement axés sur les problemes d'inférence sur les paramètres du modele. Parmi les méthodes d'estimation utilisées nous pouvons citer les "méthodes non-récursives" qui s'avèrent difficiles à mettre en euvre puisque dans ce modèle les variables aléatoires ne sont pas iid.

Les methodes récursives sont basées des équations récursives et utilisent des estimations initiales des paramètres inconnus pour obtenir de nouvelles estimations plus performantes a chaque fois qu'une nouvelle information est disponible.

Dans ce travail nous interessons aux propriétés des méthodes d'estimation récursive de paramètres dans les modèles autorégressifs où les observations ne sont pas iid et de leur mise en euvre numérique.

Dans ce mémoire nous développons les résultats sur les estimateurs récursifs obtenus dans les articles Teo Sharia suivants :

1. "**On Recursive Paremetric Estimation Theory**"

Technical report Royal Holloway, University of London (RHUL MA) -Jan. 2003

2."**Recursive parameter estimation convergence**"

Stat. infer. Stoch. Process **11** (2007),2,157-175.

3. "**Rate of convergence in recursive parameter estimation procedure**"

Georgian Mathematical Journal **14** (2007), 761-776.

Notre mémoire est constitué de deux parties. La première partie contient les résultats théoriques et est divisée en trois chapitres:

Dans le chapitre 1, nous considérons une classe d'estimateurs (la classe de M-estimateurs) pour des modèles statistiques généraux. Nous étudions leur convergence et leur approximation stochastique ainsi que l'utilisation de ces résultats dans des exemples.

Dans le chapitre 2, nous étudions la convergence d'estimateurs récursifs définis par

$$\hat{\theta}_t = \hat{\theta}_{t-1} + \Gamma_t^{-1}(\hat{\theta}_{t-1})\psi_t(\hat{\theta}_{t-1}), \qquad t \geq 1$$

où:

ψ_t est un processus vectoriel donné.

Γ_t est un processus matriciel normalisé.

$\hat{\theta}_0$ est une valeur initiale.

Nous donnons l'utilisation de ces resultats de convergence dans quelques exemples.

Dans le chapitre 3, nous présentons des resultats sur la vitesse de convergence de $\hat{\theta}_t \longrightarrow \theta$, et aussi des exemples montrant que sous certaines conditions de régularité les estimateurs récursifs sont asymptotiquement localement lineaires.

En Annexe, nous illustrons les résultats de la partie théorique par des simulations numériques à l'aide du logiciel R, et nous terminons par des resultats auxiliaires.

Chapitre 1

Estimation paramétrique récursive

1.1 Introduction

Nous considérons dans cette partie l'estimation paramétrique dans plusieurs types de modèles par des méthodes récursives.

Notre etude porte sur le développement des résultats de l'article de Sharia. T:

"**On recursive paremetric estimation theory**" technical Report RHUL MA (2003).

Soit $X_1, ..., X_n$ des variables aléatoires réelles indépendantes identiquement distribuiées (iid) de fonction de répartition F_θ dépendant d'un paramètre réel inconnu θ. Un estimateur de θ est défini par une statistique $\theta_n = \theta_n(X_1, ..., X_n)$ qui est en général une solution de l'équation suivante

$$\sum_{i=1}^{n} \psi\left(X_i, \theta\right) = 0 \tag{1.1.1}$$

où ψ est une une certaine fonction donnée.

Par exemple, si θ est un paramètre inconnu de la famille gaussienne $\mathcal{N}(\theta, \sigma^2)$ le choix de la fonction $\psi(x, \theta) = x - \theta$ donne l'estimateur de maximum du vraissemblance (emv).

En effet puisque $f(x, \theta) = \frac{1}{\sigma\sqrt{2\pi}} \exp(-\frac{1}{2\sigma^2}(x-\theta)^2)$ et en prenant la vraisemblance

$$L(X_1, ..., X_n, \theta) = \prod_{i=1}^{n} f(X_i, \theta) = \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} \exp(-\frac{1}{2\sigma^2}(X_i - \theta)^2)$$

la dérivée s'annule et donne l'équation suivante de type (1.1.1) : $\sum\limits_{i=1}^{n} \psi(X_i, \theta) = \sum\limits_{i=1}^{n} (X_i - \theta) = 0$

qui pour solution l'emv: $\theta_n(X_1, ..., X_n) = \frac{1}{n}\sum\limits_{i=1}^{n} X_i$ qui est la moyenne empirique. De même pour la même famille gaussienne $\mathcal{N}(\theta, \sigma^2)$ si on choisit $\psi(x, \theta) = sign(x - \theta) = 1$ si $x > \theta$ et -1 si $x < \theta$ alors l'estimateur de θ solution de $\sum\limits_{i=1}^{n} sign(X_i - \theta) = 0$ est la médiane empirique $\theta_n(X_1, ..., X_n) = med(X_1, ..., X_n)$.

Soit $f(x, \theta)$ la densité de X_i (par rapport à une mesure μ σ-finie) et supposons que cette densité est différentiable par rapport à θ. Si on choisit la fonction $\psi\,(x, \theta) = \acute{f}(x, \theta)/f(x, \theta)$ alors l'équation (1.1.1) devient

$$0 = \sum_{i=1}^{n} \psi(X_i, \theta) = \sum_{i=1}^{n} \frac{\acute{f}(X_i, \theta)}{f(X_i, \theta)} = \sum_{i=1}^{n} \frac{\partial}{\partial\theta}\left[\log f(X_i, \theta)\right] = \frac{\partial}{\partial\theta} L(X_1, ..., X_n, \theta)$$

où $L(X_1, ..., X_n, \theta)$ est la vraisemblance et par suit sa resolution donne l'estimateur de maximum du vraissemblance.

Supposons maintenant que les v.a $X_1, ..., X_n$ ne sont pas nécessairement indépendantes. Nous notons la densité de X_i conditionnelle à $X_{i-1}, ..., X_{i-k}$ où k est fixé, par

$$f_i(x, \theta) = f_i(x, \theta / X_{i-1}, ..., X_{i-k})$$

et on choisit $\psi_i(X_i, \theta) = f_i^{'}(X_i, \theta)/f_i(X_i, \theta)$. L'equation (1.1.1) s'écrit

$$0 = \sum_{i=1}^{n} \psi_i(X_i, \theta) = \sum_{i=1}^{n} \frac{\partial}{\partial\theta}\left[\log f_i(X_i, \theta)\right] = \frac{\partial}{\partial\theta} L(X_1, ..., X_n, \theta)$$

et sa résolution de cette équation donne l'estimateur de maximum du vraissemblance. Sous certaines conditions de régularité et d'ergodicité, il existe une suite de solutions possédant de bonnes propriétés asymptotiques (voir Ch.2).

Nous remarquons en général par la suite que si les fonctions ψ_i sont non linéaires, il est difficile de résoudre les équations donnant les estimateurs correspondants.

Nous considérons des procedures d'estimation récursives d'un paramètre $\theta \in \mathbb{R}^m$ où à chaque étape l'estimateur suivant est obtenu par un simple adjustement de l'estimateur précédent. En particulier, nous considérons une classe d'estimateur $(\hat{\theta}_n)$ défini récursivement par

$$\hat{\theta}_n = \hat{\theta}_{n-1} + \Gamma_n^{-1}(\hat{\theta}_{n-1})\psi_n(\hat{\theta}_{n-1}), \qquad n \geq 1 \tag{1.1.2}$$

où:

ψ_n est une fonction vectorielle quelconque.

Γ_n est un processs gaussien matriciel (peut être déterministe).

$\hat{\theta}_0 \in \mathbb{R}^m$ est un point initial.

En général la règle d'arret de la procedure d'estimation est basée sur le fait que la suite $(\hat{\theta}_n)$ reste constante.

Si on considère le modèle de v.a indépendantes identiquement distribuées (iid), des procedures d'estimation similaires à (1.1.2) utilisent des méthodes d'approximation stochastique. En particulier dans les articles de Campbell (1982, [2]) et dans Sharia (1998, [13]), on trouve le choix de $\psi_n(x,\theta) = \frac{\partial}{\partial\theta} \log f_n(x,\theta)$. Le but de notre étude est la convergence des estimateurs récursifs définis par (1.1.2) avec un point initial arbitraire $\hat{\theta}_0$.

1.2 Notations et préliminaires

Soit $X_t, t = 1, 2, ...$ des observations à valeurs dans un espace mesurable $(X, \mathcal{B}(X))$ muni d'une mesure μ σ-finie. Supposons que la loi des v.a X_t dépend d'un paramètre inconnu $\theta \in \Theta$, où $\Theta \subset \mathbb{R}^m$ et nous notons par $(P^\theta, \theta \in \Theta)$ la famille des lois correspondantes. L'espace $(\mathbb{R}^m, \mathcal{B}(\mathbb{R}^m))$ est muni de sa tribu Borelienne $\mathcal{B}(\mathbb{R}^m)$. Pour tout t, on suppose qu' il existe une densité conditionnelle régulière de X_t notées par

$$f_t\left(\theta, x_t/x_1^{t-1}\right) = f_t(\theta, x_t/x_{t-1}, ..., x_1)$$

où $f_1\left(\theta, x_1/x_1^0\right) = f_1(\theta, x_1/x_1) = f_1(\theta, x_1)$ est la densité de la variable aléatoire X_1.

Soit $\mathcal{F}_t = \sigma(X_1, ..., X_t)$ la tribu engendrée par les variables aléatoires $X_1, ..., X_t$. La transposée d'une matrice ou d'un vecteur est notée par T et $\langle u, v\rangle$ est le produit scalaire standard de $u, v \in \mathbb{R}^m, \langle u, v\rangle = u^T v$. Tous les vecteurs ci dessous sont notés par une colonne mais par notation ils sont écrits en ligne. Si h est une fonction réelle définie dans $\Theta \subset \mathbb{R}^m$, nous notons par $\dot{h}(\theta)$ le vecteur de dérivées partielles de $h(\theta)$ par rapport à θ qui est

$$\dot{h}(\theta) = \left(\frac{\partial}{\partial\theta^1} h(\theta), ..., \frac{\partial}{\partial\theta^m} h(\theta)\right)$$

Si pour tout $t = 1, 2, ...$, la dérivé $\dot{f}_t\left(\theta, x_t/x_1^{t-1}\right) = \left(\frac{\partial}{\partial\theta^1}\dot{f}_t, ..., \frac{\partial}{\partial\theta^m}\dot{f}_t\right)$ existe, alors nous définissons la fonction

$$l_t\left(\theta, x_t/x_1^{t-1}\right) = \frac{1}{f_t\left(\theta, x_t/x_1^{t-1}\right)}\dot{f}_t^T\left(\theta, x_t\ /x_1^{t-1}\right)$$

avec la convention $0/0 = 0$.

La matrice d'information conditionnelle de Fisher à l'étape $t = 1, 2, ...$ est définie par

$$\begin{aligned} i_t(\theta/x_1^{t-1}) &= E_\theta\left[\frac{\partial}{\partial\theta^i}\ln f_t(\theta, x_t/x_1^{t-1})\frac{\partial}{\partial\theta^j}\ln f_t(\theta, x_t/x_1^{t-1})\right] \\ &= E_\theta\left[\frac{\dot{f}_t^T\left(\theta, x_t/x_1^{t-1}\right)}{f_t\left(\theta, x_t/x_1^{t-1}\right)}\frac{\dot{f}_t\left(\theta, x_t/x_1^{t-1}\right)}{f_t\left(\theta, x_t/x_1^{t-1}\right)}\right] \\ &= E_\theta\left[l_t\left(\theta, x_t/x_1^{t-1}\right)l_t^T\left(\theta, x_t/x_1^{t-1}\right)\right] \\ &= \int l_t\left(\theta, z/x_1^{t-1}\right)l_t^T\left(\theta, z/x_1^{t-1}\right)f_t\left(\theta, z/x_1^{t-1}\right)\mu(dz) \end{aligned}$$

où $\dot{f}_t\left(\theta, x_t/x_1^{t-1}\right)$ est une matrice $1 \times m$ et $\dot{f}_t^T\left(\theta, x_t/x_1^{t-1}\right)$ est une matrice $m \times 1$.par suit $\dot{f}_t^T\left(\theta, x_t/x_1^{t-1}\right)\dot{f}_t\left(\theta, x_t/x_1^{t-1}\right)$ est une matrice $m \times m$. Nous notons

$$f_t(\theta) = f_t\left(\theta, X_t/X_1^{t-1}\right), l_t(\theta) = l_t\left(\theta, X_t/X_1^{t-1}\right), i_t(\theta) = i_t\left(\theta/X_1^{t-1}\right)$$

Nous disons que le processus $i_t(\theta)$ est *"prévisible"*, si la variable aléatoire $i_t(\theta)$ est $\mathcal{F}_{t-1}$-mesurable. Aussi par définition $i_t(\theta)$ est la version de l'espérance conditionnelle par rapport à $\mathcal{F}_{t-1}$, $i_t(\theta) = E_\theta\left(l_t(\theta)l_t^T(\theta)/\mathcal{F}_{t-1}\right)$.

L'information conditionnelle de Fisher au temps t est

$$I_t(\theta) = \sum_{s=1}^{t} i_s(\theta), \qquad t = 1, 2, ...$$

Si les X_t sont des variables aléatoires indépendentes, $I_t(\theta)$ se réduit à la matrice d'information de Fisher standard. Aussi $I_t(\theta)$ est dite espérance d'information de Fisher cumulée.

Pour la suite des fonctions d'estimation $\Psi = (\psi_t(\theta, x_t; x_{t-1}, ..., x_1))_{t\geq1}$ où $\psi_t \in \Psi, t \geq 1$ est une fonction mesurable.

Nous notons que la suite $(l_t\left(\theta, x_t/x_1^{t-1}\right)_{t\geq 1}) \in \Psi$, et dans ce cas la procedure récursive (de maximum du vraissemblance (mv)) est définie par

$$\hat{\theta}_t = \hat{\theta}_{t-1} + I_t^{-1}(\hat{\theta}_{t-1})l_t(\hat{\theta}_{t-1}), \qquad t \geq 1$$

Remarque 1 *Dans la procedure (1.1.2) nous avons $I_t = \Gamma_t$, tel que I_t est la matrice d'information de Fisher standard, l_t est un vecteur,et*

$$l_t(\hat{\theta}_{t-1}) = l_t\left(\hat{\theta}_{t-1}, X_t/X_1^{t-1}\right) = \frac{\dot{f}_t^T\left(\hat{\theta}_{t-1}, X_t/X_1^{t-1}\right)}{f_t\left(\hat{\theta}_{t-1}, X_t/X_1^{t-1}\right)}$$

Convention

Le paramètre $\theta \in \mathbb{R}^m$ est un vecteur fixé du paramètre. La convergence et toutes les relations entre les variables aléatoires sont définies sous une mesure P^θ (par exemple une suite de variables aléatoires $(\xi_t)_{t\geq 1}$ possède une propriété $P^\theta p.s$ s'il existe Ω^θ tel que $P^\theta(\Omega^\theta) = 1$ et pour tout $\omega \in \Omega^\theta : \xi_t(\omega)$ a cette propriété pour tout $t > t_0(\omega)$.

1.3 Convergence des estimateurs

Supposons que $\psi \in \Psi$ et $\Gamma_t(\theta)$ une matrice $m \times m$ prévisible, pour tout $\theta \in \mathbb{R}^m$, telle que $\det \Gamma_t(\theta) \neq 0, t \geq 1$. Considérons l'estimateur $\hat{\theta}_t$ défini par

$$\hat{\theta}_t = \hat{\theta}_{t-1} + \Gamma_t^{-1}(\hat{\theta}_{t-1})\psi_t(\hat{\theta}_{t-1}), \qquad t \geq 1 \tag{1.3.1}$$

où $\hat{\theta}_0 \in \mathbb{R}^m$ est un point initial arbitraire.

Soit $\theta \in \mathbb{R}^m$ une valeur fixée du paramètre, et pour tout $u \in \mathbb{R}^m$, nous définissons

$$b_t(\theta, u) = E_\theta\left[\psi_t(\theta + u)/\mathcal{F}_{t-1}\right]$$

Pour obtenir la convergence des estimateurs $\hat{\theta}_t$ nous avons besoin du lemme suivant.

Lemme 1 *Soit $(C_t(\theta))_{t\geq 1}$ un processus de matrices symétriques prévisibles $m \times m$ définies* *non-négatives. Notons par*

$$\Delta_t = \hat{\theta}_t - \theta, \quad et \quad V_t(u) = \langle C_t(\theta)u, u\rangle\,, \quad \Delta V_t(u) = V_t(u) - V_{t-1}(u)$$

et supposons que

$$\sum_{t=1}^{\infty} (1 + V_{t-1}(\Delta_{t-1}))^{-1} \left[\mathcal{K}_t(\theta)\right]^+ < \infty, \qquad P^\theta p.s \tag{1.3.2}$$

où

$$\begin{aligned}\mathcal{K}_t(\theta) \quad &= \quad \Delta V_t(\Delta_{t-1}) + 2\left\langle C_t(\theta)\Delta_{t-1}, \Gamma_t^{-1}(\Delta_{t-1}+\theta) b_t(\theta, \Delta_{t-1})\right\rangle \\ &\quad + E_\theta\left(\left[\Gamma_t^{-1}(\Delta_{t-1}+\theta)\psi_t(\Delta_{t-1}+\theta)\right]^T C_t(\theta)\Gamma_t^{-1}(\Delta_{t-1}+\theta)\psi_t(\Delta_{t-1}+\theta)/\mathcal{F}_{t-1}\right)\end{aligned} \tag{1.3.3}$$

Alors $V_t(\Delta_t)$ converge $P^\theta p.s$ vers une limite finie.

Preuve: Dans (1.3.1) nous ajoutons $-\theta$ et on a

$$\hat{\theta}_t - \theta = \hat{\theta}_{t-1} - \theta + \Gamma_t^{-1}(\Delta_{t-1}+\theta)\psi_t(\Delta_{t-1}+\theta), \qquad \forall t \geq 1$$

Or par definition $\hat{\theta}_{t-1} = \Delta_{t-1} + \theta$, donc

$$\Delta_t = \Delta_{t-1} + \Gamma_t^{-1}(\Delta_{t-1}+\theta)\psi_t(\Delta_{t-1}+\theta)$$

Un développement de Taylor à l'ordre 2 de $V_t(\Delta_t)$ au point Δ_{t-1}, donne

$$V_t(\Delta_t) = V_t(\Delta_{t-1}) + \dot{V}_t(\Delta_{t-1})^T(\Delta_t - \Delta_{t-1}) + \frac{1}{2}(\Delta_t - \Delta_{t-1})^T \ddot{V}_t(\tilde{\Delta}_{t-1})(\Delta_t - \Delta_{t-1})$$

où $\tilde{\Delta}_{t-1}$ est un point entre Δ_{t-1} et Δ_t. Posons

$$M := \Delta_t - \Delta_{t-1} = \Gamma_t^{-1}(\Delta_{t-1}+\theta)\psi_t(\Delta_{t-1}+\theta)$$

par suit

$$\langle M, M\rangle = (\Delta_t - \Delta_{t-1})^2 = M^T M$$

Donc

$$\begin{aligned} V_t(\Delta_t) &= V_t(\Delta_{t-1}) + \dot{V}_t(\Delta_{t-1})^T \Gamma_t^{-1}(\Delta_{t-1}+\theta)\psi_t(\Delta_{t-1}+\theta) \\ &+ \frac{1}{2}\left[\Gamma_t^{-1}(\Delta_{t-1}+\theta)\psi_t(\Delta_{t-1}+\theta)\right]^T \ddot{V}_t(\tilde{\Delta}_{t-1})\Gamma_t^{-1}(\Delta_{t-1}+\theta)\psi_t(\Delta_{t-1}+\theta) \end{aligned}$$

D'une part,

$$V_t(u) = \langle C_t(\theta)u, u\rangle = u^T C_t(\theta)u$$

et sa dérivée $\dot{V}_t(u) = 2C_t(\theta)u$ et $\ddot{V}_t(u) = 2C_t(\theta)$. Par conséquent

$$\begin{aligned} V_t(\Delta_t) &= V_t(\Delta_{t-1}) + 2\Delta_{t-1}^T C_t(\theta)\Gamma_t^{-1}(\Delta_{t-1}+\theta)\psi_t(\Delta_{t-1}+\theta) \\ &+ \left[\Gamma_t^{-1}(\Delta_{t-1}+\theta)\psi_t(\Delta_{t-1}+\theta)\right]^T C_t(\theta)\Gamma_t^{-1}(\Delta_{t-1}+\theta)\psi_t(\Delta_{t-1}+\theta) \end{aligned}$$

En utilisant $\langle u, v\rangle = u^T v$ on obtient

$$\begin{aligned} V_t(\Delta_t) &= V_t(\Delta_{t-1}) + 2\left\langle C_t(\theta)\Delta_{t-1}, \Gamma_t^{-1}(\Delta_{t-1}+\theta)\psi_t(\Delta_{t-1}+\theta)\right\rangle \\ &+ \left[\Gamma_t^{-1}(\Delta_{t-1}+\theta)\psi_t(\Delta_{t-1}+\theta)\right]^T C_t(\theta)\Gamma_t^{-1}(\Delta_{t-1}+\theta)\psi_t(\Delta_{t-1}+\theta) \end{aligned}$$

D'autre part de $\Delta V_t = V_t - V_{t-1}$ on a

$$V_t(\Delta_{t-1}) = V_{t-1}(\Delta_{t-1}) + \Delta V_t(\Delta_{t-1})$$

ce qui implique

$$\begin{aligned} V_t(\Delta_t) &= V_{t-1}(\Delta_{t-1}) + \Delta V_t(\Delta_{t-1}) + 2\left\langle C_t(\theta)\Delta_{t-1}, \Gamma_t^{-1}(\Delta_{t-1}+\theta)\psi_t(\Delta_{t-1}+\theta)\right\rangle \\ &+ \left[\Gamma_t^{-1}(\Delta_{t-1}+\theta)\psi_t(\Delta_{t-1}+\theta)\right]^T C_t(\theta)\Gamma_t^{-1}(\Delta_{t-1}+\theta)\psi_t(\Delta_{t-1}+\theta) \end{aligned}$$

Or par définition

$$b_t(\theta, \Delta_{t-1}) = E_\theta\left(\psi_t(\Delta_{t-1}+\theta)/\mathcal{F}_{t-1}\right)$$

et les v.a.$C_t, \Gamma_t, V_{t-1}, \Delta_{t-1}$ sont $\mathcal{F}_{t-1}$-mesurables, en prenant l'espérance conditionnelle à $\mathcal{F}_{t-1}$ on aboutit à

$$E_\theta\left(V_t(\Delta_t)/\mathcal{F}_{t-1}\right) = V_{t-1}(\Delta_{t-1}) + \mathcal{K}_t\left(\theta\right)$$

où

$$\begin{aligned}
\mathcal{K}_t(\theta) &= \Delta V_t(\Delta_{t-1}) + 2\left\langle C_t(\theta)\Delta_{t-1}, \Gamma_t^{-1}(\Delta_{t-1}+\theta)b_t(\theta,\Delta_{t-1})\right\rangle \\
&\quad + E_\theta\left(\left[\Gamma_t^{-1}(\Delta_{t-1}+\theta)\psi_t(\Delta_{t-1}+\theta)\right]^T C_t(\theta)\Gamma_t^{-1}(\Delta_{t-1}+\theta)\psi_t(\Delta_{t-1}+\theta)/\mathcal{F}_{t-1}\right)
\end{aligned}$$

Nous utilisons la décomposition de la v.a.$\mathcal{K}_t = [\mathcal{K}_t]^+ - [\mathcal{K}_t]^-$ on déduit que

$$\begin{aligned}
E_\theta\left(V_t(\Delta_t)/\mathcal{F}_{t-1}\right) &= V_{t-1}(\Delta_{t-1}) + \mathcal{K}_t = V_{t-1}(\Delta_{t-1}) + [\mathcal{K}_t]^+ - [\mathcal{K}_t]^- \\
&= V_{t-1}(\Delta_{t-1}) + (1+V_{t-1}(\Delta_{t-1}))(1+V_{t-1}(\Delta_{t-1}))^{-1}[\mathcal{K}_t]^+ - [\mathcal{K}_t]^- \\
&= V_{t-1}(\Delta_{t-1}) + (1+V_{t-1}(\Delta_{t-1}))B_t - [\mathcal{K}_t]^- \\
&= V_{t-1}(\Delta_{t-1})(1+B_t) + B_t - [\mathcal{K}_t]^-
\end{aligned}$$

où on a posé $B_t = (1+V_{t-1}(\Delta_{t-1}))^{-1}[\mathcal{K}_t]^+$. D'après (1.3.2), nous avons $\sum\limits_{t=1}^{\infty} B_t < \infty \; P^\theta p.s$ et nous appliquons le Lemme 2 de l'annexe 1, avec $X_{n-1} = V_{n-1}(\Delta_{n-1})$, $\beta_{n-1} = \xi_{n-1} = B_n$, $\zeta_{n-1} = [\mathcal{K}_t]^-$ qui sont des v.a.$\mathcal{F}_{n-1}$-mesurables. Donc

$$\left\{\sum_{i=1}^{\infty}\xi_{i-1} = \sum_{i=1}^{\infty}\beta_{i-1} = \sum_{i=1}^{\infty}B_i < \infty\right\} \subseteq \{V_t(\Delta_t) \longrightarrow V\} \cap \left\{\sum_{i=1}^{\infty}[\mathcal{K}_i]^- < \infty\right\}, P^\theta p.s$$

Par conséquent

$$\left\{\sum_{i=1}^{\infty}B_i < \infty\right\} \subseteq \{V_t(\Delta_t) \longrightarrow V\} \cap \left\{\sum_{i=1}^{\infty}[\mathcal{K}_i]^- < \infty\right\}, \quad P^\theta p.s$$

alors

$$1 = P^\theta\left(\sum_{i=1}^{\infty}B_i < \infty\right) \leq P^\theta\left(\{V_t(\Delta_t) \longrightarrow V\} \cap \left\{\sum_{i=1}^{\infty}[\mathcal{K}_i]^- < \infty\right\}\right)$$

Donc

$$P^\theta\left(\{V_t(\Delta_t) \longrightarrow V\} \cap \left\{\sum_{i=1}^{\infty}[\mathcal{K}_i]^- < \infty\right\}\right) = 1$$

c'est-à dire $P^\theta\left(\{V_t(\Delta_t)\longrightarrow V\}\right)=1$ ou encore $V_t(\Delta_t)\longrightarrow V, P^\theta p.s$ ce qui termine la preuve du Lemme. ■

Le résultat suivant donne la convergence des estimateurs avec une vitesse.

Corollaire [1.1] Soit $(a_t(\theta))_{t\geq 1}$ un processus scalaire prévisible non-décroissant tel que $a_t(\theta)\longrightarrow_{t\longrightarrow\infty}\infty$, et notons $\Delta a_t(\theta)=a_t(\theta)-a_{t-1}(\theta)$. Supposons que

$(R1)$

$$\lim_{t\longrightarrow\infty}\frac{\triangle a_t(\theta)}{a_{t-1}(\theta)}=0,\quad P^\theta p.s$$

$(R2)$ il existe une matrice C_θ symétrique définie positive et un processus scalaire $\mathcal{P}_t$ prévisible positif, tel que

$$2\left\langle C_\theta\Delta_{t-1},\Gamma_t^{-1}(\theta+\Delta_{t-1})b_t(\theta+\Delta_{t-1})\right\rangle+\mathcal{P}_t\leq-\lambda_t(\theta)\left\langle C_\theta\Delta_{t-1},\Delta_{t-1}\right\rangle \tag{1.3.4}$$

où $(\lambda_t(\theta))_{t\geq 1}$ est un processus scalaire prévisible satisfaisant

$$\sum_{i=1}^{\infty}\left[\frac{\triangle a_t(\theta)}{a_t(\theta)}-\lambda_t(\theta)\right]^+<\infty,\quad P^\theta p.s \tag{1.3.5}$$

$(R3)$ pour tout $0<\varepsilon<1$, et le processus $\mathcal{P}_t$ défini dans $(R2)$,

$$\sum_{i=1}^{\infty}a_t^\varepsilon(\theta)\left[E_\theta\left(\left\|\Gamma_t^{-1}(\theta+\Delta_{t-1})\psi_t(\theta+\Delta_{t-1})\right\|^2/\mathcal{F}_{t-1}\right)-\mathcal{P}_t\right]^+<\infty,\quad P^\theta p.s$$

Alors pour tout $\delta\in]0,1/2[$: $a_t^{2\delta}(\theta)(\hat\theta_t-\theta)^TC_\theta(\hat\theta_t-\theta)\longrightarrow 0$ $P^\theta p.s.$

Preuve: Vérifions les conditions du lemme 1 où $C_t(\theta)=a_t^{2\delta}(\theta)C_\theta$ et $\delta\in]0,1/2[$ et

$$V_t(\Delta_t)=a_t^{2\delta}\left\langle C_\theta\Delta_t,\Delta_t\right\rangle=a_t^{2\delta}\left\|\hat\theta_t-\theta\right\|^2$$

Seulement dans ce cas nous définissons la norme comme suite

$$\left\langle C_\theta x,x\right\rangle=\|x\|^2,\forall x\in\mathbb{R}^m$$

Nous notons

$$r_t = \left(\Delta a_t^{2\delta} - a_t^{2\delta}\lambda_t\right)/a_{t-1}^{2\delta} \text{ ,et } \tilde{\mathcal{P}}_t = a_t^{2\delta}(\mathcal{E}_t - \mathcal{P}_t)$$

où

$$\mathcal{E}_t = E_\theta\left(\left[\Gamma_t^{-1}(\theta+\Delta_{t-1})\psi_t(\theta+\Delta_{t-1})\right]^T C_\theta\left[\Gamma_t^{-1}(\theta+\Delta_{t-1})\psi_t(\theta+\Delta_{t-1})\right]/\mathcal{F}_{t-1}\right)$$

Le processus $\mathcal{K}_t$ est défini dans (1.3.3), nous avons donc

$$\begin{aligned}
\mathcal{K}_t &= \Delta a_t^{2\delta}\langle C_\theta\Delta_{t-1},\Delta_{t-1}\rangle + 2a_t^{2\delta}\left\langle C_\theta\Delta_{t-1},\Gamma_t^{-1}(\theta+\Delta_{t-1})b_t(\theta,\Delta_{t-1})\right\rangle + a_t^{2\delta}\mathcal{E}_t \\
&= \Delta a_t^{2\delta}\langle C_\theta\Delta_{t-1},\Delta_{t-1}\rangle + 2a_t^{2\delta}\left\langle C_\theta\Delta_{t-1},\Gamma_t^{-1}(\theta+\Delta_{t-1})b_t(\theta,\Delta_{t-1})\right\rangle + a_t^{2\delta}\mathcal{P}_t + \tilde{\mathcal{P}}
\end{aligned}$$

Par $(R2),(1.3.4)$ nous avons

$$\begin{aligned}
\mathcal{K}_t &\leq \Delta a_t^{2\delta}\langle C_\theta\Delta_{t-1},\Delta_{t-1}\rangle - a_t^{2\delta}\lambda_t\langle C_\theta\Delta_{t-1},\Delta_{t-1}\rangle + \tilde{\mathcal{P}}_t \\
&\leq \left(\Delta a_t^{2\delta} - a_t^{2\delta}\lambda_t\right)\langle C_\theta\Delta_{t-1},\Delta_{t-1}\rangle + \tilde{\mathcal{P}}_t \\
&\leq \left(\Delta a_t^{2\delta} - a_t^{2\delta}\lambda_t\right)\frac{a_{t-1}^{2\delta}}{a_{t-1}^{2\delta}}\langle C_\theta\Delta_{t-1},\Delta_{t-1}\rangle + \tilde{\mathcal{P}}_t \\
&\leq r_t a_{t-1}^{2\delta}\langle C_\theta\Delta_{t-1},\Delta_{t-1}\rangle + \tilde{\mathcal{P}}_t
\end{aligned}$$

Comme C_θ est définie non-négative, nous avons

$$[\mathcal{K}_t]^+ \leq [r_t]^+ a_{t-1}^{2\delta}\langle C_\theta\Delta_{t-1},\Delta_{t-1}\rangle + \left[\tilde{\mathcal{P}}_t\right]^+$$

Ce qui implique que

$$\begin{aligned}
(1+V_{t-1}(\Delta_{t-1}))^{-1}[\mathcal{K}_t]^+ &= \left(1+a_{t-1}^{2\delta}\langle C_\theta\Delta_{t-1},\Delta_{t-1}\rangle\right)^{-1}[\mathcal{K}_t]^+ \\
&\leq \left(1+a_{t-1}^{2\delta}\langle C_\theta\Delta_{t-1},\Delta_{t-1}\rangle\right)^{-1}[r_t]^+ a_{t-1}^{2\delta}\langle C_\theta\Delta_{t-1},\Delta_{t-1}\rangle \\
&\quad + \left(1+a_{t-1}^{2\delta}\langle C_\theta\Delta_{t-1},\Delta_{t-1}\rangle\right)^{-1}\left[\tilde{\mathcal{P}}_t\right]^+ \\
&\leq [r_t]^+ + \left[\tilde{\mathcal{P}}_t\right]^+
\end{aligned}$$

car $\left(1+a_{t-1}^{2\delta}\left\langle C_\theta\Delta_{t-1},\Delta_{t-1}\right\rangle\right)^{-1} a_{t-1}^{2\delta}\left\langle C_\theta\Delta_{t-1},\Delta_{t-1}\right\rangle \leq 1$, et $\left(1+a_{t-1}^{2\delta}\left\langle C_\theta\Delta_{t-1},\Delta_{t-1}\right\rangle\right)^{-1} \leq 1$. Nous avons

$$\begin{aligned}\left[\tilde{\mathcal{P}}_t\right]^+ &= \left[a_t^{2\delta}(\mathcal{E}_t-\mathcal{P}_t)\right]^+ = a_t^{2\delta}\left[\mathcal{E}_t-\mathcal{P}_t\right]^+ \\ &= a_t^{2\delta}(\theta)\left[E_\theta\left(\left\|\Gamma_t^{-1}(\theta+\Delta_{t-1})b_t(\theta+\Delta_{t-1})\right\|^2/\mathcal{F}_{t-1}\right)-\mathcal{P}_t\right]^+\end{aligned}$$

Pour $\varepsilon = 2\delta$, et $\delta \in]0,1/2[$, par $(R3)$ $(\varepsilon \in]0,1[)$ nous avons

$$\sum_{t=1}^{\infty}\left[\tilde{\mathcal{P}}_t\right]^+ = \sum_{t=1}^{\infty} a_t^{2\delta}(\theta)\left[E_\theta\left(\left\|\Gamma_t^{-1}(\theta+\Delta_{t-1})\psi_t(\theta+\Delta_{t-1})\right\|^2/\mathcal{F}_{t-1}\right)-\mathcal{P}_t\right]^+ < \infty$$

Or

$$\sum_{t=1}^{\infty}(1+V_{t-1}(\Delta_{t-1}))^{-1}\left[\mathcal{K}_t\right]^+ \leq \sum_{t=1}^{\infty}\left[r_t\right]^+ + \sum_{t=1}^{\infty}\left[\tilde{\mathcal{P}}_t\right]^+$$

Reste à montrer que $\sum_{t=1}^{\infty}[r_t]^+ < \infty$. En effet comme $\Delta a_t^{2\delta} = a_t^{2\delta} - a_{t-1}^{2\delta}$ nous reécrivons r_t comme suit

$$r_t = (a_t^{2\delta} - a_{t-1}^{2\delta} - a_t^{2\delta}\lambda_t)/a_{t-1}^{2\delta} = \left(a_t^{2\delta}(1-\lambda_t) - a_{t-1}^{2\delta}\right)a_{t-1}^{-2\delta} = \left(a_t a_{t-1}^{-1}\right)^{2\delta}(1-\lambda_t) - 1$$

Par le développement limité $(1+x)^{2\delta} = 1 + 2\delta x + O(x^2)$ nous avons

$$\left(a_t a_{t-1}^{-1}\right)^{2\delta} = \left(\frac{a_{t-1}+\Delta a_t}{a_{t-1}}\right)^{2\delta} = \left(1+\frac{\Delta a_t}{a_{t-1}}\right)^{2\delta} = 1 + 2\delta\frac{\Delta a_t}{a_{t-1}} + O\left(\frac{\Delta a_t}{a_{t-1}}\right)^2$$

Par $(R1)$, $O\left(\Delta a_t/a_{t-1}\right)^2 \longrightarrow_{t\longrightarrow\infty} 0$. En posant

$$\eta_t = \frac{\Delta a_t}{a_t} - \lambda_t$$

r_t se réduit à

$$r_t = \left(a_t a_{t-1}^{-1}\right)^{2\delta}(1-\lambda_t) - 1 = \left(a_t a_{t-1}^{-1}\right)^{2\delta}\left(1+\eta_t-\frac{\Delta a_t}{a_t}\right) - 1$$

Mais $\eta_t = [\eta_t]^+ - [\eta_t]^-$, et $[\eta_t]^+ \geq \eta_t$ par suit en posant $\delta_t^{(1)} = O\left(\Delta a_t / a_{t-1}\right)^2$ on aboutit à

$$\begin{aligned}
r_t &\leq \left(a_t a_{t-1}^{-1}\right)^{2\delta}\left(1+\eta_t^+ - \frac{\Delta a_t}{a_t}\right) - 1 \\
&\leq \left(1 + 2\delta\frac{\Delta a_t}{a_{t-1}} + \delta_t^{(1)}\right)\left(1+\eta_t^+ - \frac{\Delta a_t}{a_t}\right) - 1 \\
&= 1 + \eta_t^+ - \frac{\Delta a_t}{a_t} - 1 + 2\delta\frac{\Delta a_t}{a_{t-1}} + 2\delta\eta_t^+\frac{\Delta a_t}{a_{t-1}} - 2\delta\frac{\Delta a_t}{a_{t-1}}\frac{\Delta a_t}{a_t} \\
&\quad + \delta_t^{(1)} + \delta_t^{(1)}\eta_t^+ - \delta_t^{(1)}\frac{\Delta a_t}{a_t} \\
&= -\frac{\Delta a_t}{a_{t-1}} + \frac{\Delta a_t}{a_t}\frac{\Delta a_t}{a_{t-1}} + 2\delta\frac{\Delta a_t}{a_{t-1}} - 2\delta\frac{\Delta a_t}{a_{t-1}}\frac{\Delta a_t}{a_t} + \delta_t^{(1)} - \frac{\Delta a_t}{a_t}\delta_t^{(1)} \\
&\quad + \eta_t^+ + 2\delta\eta_t^+\frac{\Delta a_t}{a_{t-1}} + \eta_t^+\delta_t^{(1)}
\end{aligned}$$

D'autre part

$$-\frac{\Delta a_t}{a_t} = \frac{\Delta a_t}{a_{t-1}}\left(\frac{-a_{t-1}}{a_t}\right) = \frac{\Delta a_t}{a_{t-1}}\left(-1 + \frac{\Delta a_t}{a_t}\right) = -\frac{\Delta a_t}{a_{t-1}} + \frac{\Delta a_t}{a_t}\frac{\Delta a_t}{a_{t-1}}$$

et donc

$$\begin{aligned}
r_t &\leq -\left(1-2\delta\right)\frac{\Delta a_t}{a_{t-1}} + \delta_t^{(1)} - \frac{\Delta a_t}{a_t}\delta_t^{(1)} + \left(1-2\ \delta\right)\frac{\Delta a_t}{a_t}\frac{\Delta a_t}{a_{t-1}} \\
&\quad + \eta_t^+ + 2\delta\eta_t^+\frac{\Delta a_t}{a_{t-1}} + \eta_t^+\delta_t^{(1)} \\
&= \frac{\Delta a_t}{a_{t-1}}\left(-(1-2\delta) + \left(\frac{\Delta a_t}{a_{t-1}}\right)^{-1}\delta_t^{(1)}\left(1 - \frac{\Delta a_t}{a_t}\right) + (1-2\delta)\frac{\Delta a_t}{a_t}\right) \\
&\quad + \eta_t^+ + 2\delta\eta_t^+\frac{\Delta a_t}{a_{t-1}} + \eta_t^+\delta_t^{(1)} \\
&= \frac{\Delta a_t}{a_{t-1}}\left(-(1-2\delta) + \delta_t^{(2)}\right) + \delta_t^{(3)}
\end{aligned}$$

où

$$\delta_t^{(2)} = \left(\frac{\Delta a_t}{a_{t-1}}\right)^{-1}\delta_t^{(1)}\left(1 - \frac{\Delta a_t}{a_t}\right) + (1-2\delta)\frac{\Delta a_t}{a_t},$$

$$\delta_t^{(3)} = \eta_t^+ + 2\delta\eta_t^+\frac{\Delta a_t}{a_{t-1}} + \eta_t^+\delta_t^{(1)} = \eta_t^+(1 + 2\delta\frac{\Delta a_t}{a_{t-1}} + \delta_t^{(1)})$$

D'aprés $(R1)$, et $\delta_t^{(1)} \longrightarrow 0$ et par la convention $0/0 = 0$, on a $(\Delta a_t / a_{t-1})^{-1} \delta_t^{(1)} \longrightarrow_{t \longrightarrow \infty} 0$, et donc $\delta_t^{(2)} \longrightarrow_{t \longrightarrow \infty} 0$, et pour $t > t_0$, nous avons

$$\left|\delta_t^{(3)}\right| = \eta_t^+ \left|1 + 2\delta \frac{\Delta a_t}{a_{t-1}} + \delta_t^{(1)}\right| < \eta_t^+ (1 + \varepsilon)$$

qui implique d'aprés (1.3.5) que

$$\sum_{t>t_0}^{\infty} \left|\delta_t^{(3)}\right| \leq (1 + \varepsilon) \sum_{t>t_0}^{\infty} \eta_t^+ < \infty \text{ et donc } \sum_{t=1}^{\infty} \left|\delta_t^{(3)}\right| < \infty$$

Comme $\delta \in]0, 1/2[$ on a $1 - 2\delta > 0$, et de la formule $r_t = \frac{\Delta a_t}{a_{t-1}} \left(-(1 - 2\delta) + \delta_t^{(2)}\right) + \delta_t^{(3)}$ nous avons pour t assez grand $[r_t]^+ \leq \left|\delta_t^{(3)}\right|$. Par conséquent

$$\sum_{t=1}^{\infty} [r_t]^+ \leq \sum_{t=1}^{\infty} \left|\delta_t^{(3)}\right| < \infty$$

Donc les conditions du lemme 1 sont satisfaites et la suite des v.a.$a_t^{2\delta} \left\|\hat{\theta}_t - \theta\right\|^2 \longrightarrow V$ *p.s* où V est une limite finie. Or pour tout $\delta \in]0, 1/2[$: $\lim_{t \longrightarrow \infty} a_t^{2\delta} = \infty$ par suit $V = 0$ *p.s.*(avec la convention $0/0 = 0$). ■

Remarque 2 *Remarquons que dans* (1.3.4) *le membre de gauche est négatif donc le premièr terme à gauche de* (1.3.4) *est aussi négatif puisque* $\mathcal{P}_t > 0$. *En supposant que* $\mathcal{P}_t = 0$, *les parties positives dans* (1.3.5) *sont nulles en général (ou assez petits) dans beaucoup d'examples. D'autre part, le choix* $\mathcal{P}_t = 0$ *signifie que la condition* $(R3)$ *devient plus restrictive et impose des modèles probabilistes plus forts.le choix* $\mathcal{P}_t = 0$ *est naturel dans le cas iid sachant que touts les conditions sont satisfaites (i.e Remarque 3). Maintenant, si le premier terme de gauche de* (1.3.4) *est négatif avec une grande valeur absolue alors il est possible d'introduire* $\mathcal{P}_t > 0$ *sans compromettre* (1.3.5) *une possibilité doit être*

$$\mathcal{P}_t = \left\|\Gamma_t^{-1}(\theta + \Delta_{t-1}) b_t(\theta + \Delta_{t-1})\right\|^2$$

Dans ce cas comme $b_t(\theta, u) = E_\theta \left(\psi_t(\theta + u)/\mathcal{F}_{t-1}\right)$ *et* $\Gamma_t^{-1}(\theta + u)$ *sont des processus prévisibles,* *alors nous peuvons écrit la condition dans* $(R3)$ *comme suit*

$$\sum_{t=1}^{\infty} a_t^{\varepsilon}(\theta) E_\theta \left(\left\| \Gamma_t^{-1}(\theta+\Delta_{t-1}) \left(\psi_t(\theta+\Delta_{t-1}) - b_t(\theta,\Delta_{t-1}) \right) \right\|^2 / \mathcal{F}_{t-1} \right) < \infty$$

Remarque 3 *Considérons le cas iid avec* $f_t(\theta, z/x_1^{t-1}) = f(\theta, z), \psi_t(\theta) = \psi_t(\theta, z)|_{z=X_t}$ *et* $\int \psi(\theta,z) f(\theta,z)\mu(dz) = 0$, *et* $\gamma(\theta)$ *une matrice non-aléatoire inversible et* $\Gamma_t(\theta) = t\gamma(\theta)$. *Alors*

$$b_t(\theta,u) = b(\theta,u) = \int \psi(\theta+u,z) f(\theta,z)\mu(dz)$$

implique que

$$b_t(\theta,0) = \int \psi(\theta,z) f(\theta,z)\mu(dz) = 0$$

Notons $\Delta_t = \hat{\theta}_t - \theta$ *et nous reécrivons* (1.3.1) *sous la forme*

$$\hat{\theta}_t = \hat{\theta}_{t-1} + \Gamma_t^{-1}(\hat{\theta}_{t-1})\psi_t(\hat{\theta}_{t-1})$$

par suit

$$\hat{\theta}_t - \theta = \hat{\theta}_{t-1} - \theta + \frac{1}{t}\left[\gamma^{-1}(\theta+\Delta_{t-1})\psi(\theta+\Delta_{t-1},z)|_{z=X_t} \right]$$

qui est

$$\begin{aligned} \Delta_t &= \Delta_{t-1} + \frac{1}{t}\left[\gamma^{-1}(\theta+\Delta_{t-1})\psi(\theta+\Delta_{t-1},X_t) \right] \\ &= \Delta_{t-1} + \frac{1}{t}\left[\gamma^{-1}(\theta+\Delta_{t-1}) b_t(\theta,\Delta_{t-1}) + \gamma^{-1}(\theta+\Delta_{t-1}) \left(\psi(\theta+\Delta_{t-1},X_t) - b_t(\theta,\Delta_{t-1}) \right) \right] \end{aligned}$$

Donc

$$\Delta_t = \Delta_{t-1} + \frac{1}{t}\left(\gamma^{-1}(\theta+\Delta_{t-1}) b_t(\theta,\Delta_{t-1}) + \varepsilon_t^{\theta} \right) \tag{1.3.6}$$

où

$$\varepsilon_t^{\theta} = \gamma^{-1}(\theta+\Delta_{t-1}) \left(\psi(\theta+\Delta_{t-1},X_t) - b_t(\theta,\Delta_{t-1}) \right)$$

L'équation (1.3.6) *définit l'approximation stochastique de Robbins Monro (cf.[7],[8]) qui con-*

verge vers la solution de l'équation

$$R^{\theta}(u) := \gamma^{-1}(\theta+u)b(\theta,u) = 0$$

quand les valeurs de la fonction $R^{\theta}(u)$ *peuvent observées avec une erreur* ε_t^{θ} *d'espérance nulle. En général, la récurence* (1.3.1) *ne peut pas être considérée dans le contexte de l'approximation stochastique classique (cf.[8]). Pour le cas iid, les conditions de corollaire [1.1] peuvent etre écrites comme* $(B1)$ *et* $(B2)$ *dans la proposition 4 (e.g Remarque 5) qui sont des conditions standards dans les procedures d'approximation stochastique de type* (1.3.6) *(cf.[5],[11]).*

1.4 Exemples

1. Cas iid

Nous considérons des observations $X_1, X_2, ...$ iid de densité commune $f(x,\theta), \theta \in \mathbb{R}^m$. Soit $\psi(\theta,z)$ est une fonction d'estimation verifiant

$$\int \psi(\theta,z)f(\theta,z)\mu(dz) = 0$$

Nous définissons un estimateur récursif $\hat{\theta}_t$ par

$$\hat{\theta}_t = \hat{\theta}_{t-1} + \frac{1}{t}\gamma^{-1}(\hat{\theta}_{t-1})\psi(\hat{\theta}_{t-1}, X_t), \qquad t \geq 1 \tag{1.4.1}$$

où $\gamma(\theta)$ est une matrice non-aléatoire telle que $\gamma^{-1}(\theta)$ existe pour tout $\theta \in \mathbb{R}^m, \hat{\theta}_0 \in \mathbb{R}^m$ est une valeur initiale.

Nous avons le résultat suivant.

Proposition 4 *Supposons que* $\hat{\theta}_t \longrightarrow \theta$ P^{θ}*p.s, et*

$(B1)$ *il existe une matrice* C_{θ} *symétrique et définie non-négative, telle que pour tout u assez petit*

$$\left\langle C_{\theta}u, \gamma^{-1}(\theta+u)E^{\theta}\psi(\theta+u, X_1)\right\rangle \leq -\frac{1}{2}\left\langle C_{\theta}u, u\right\rangle$$

$(B2)$

$$E_{\theta}\left\|\gamma^{-1}(\theta+u)\psi(\theta+u)\right\|^2 = O(1) \quad quand \quad u \longrightarrow 0.$$

Alors pour tout $\delta \in]0,1/2[\,, t^{2\delta}(\hat{\theta}_t-\theta)^T C_\theta(\hat{\theta}_t-\theta) \longrightarrow 0 \;\; P^\theta p.s.$

Preuve: Le résultat est une conséquence du corollaire [1.1] si on choisit $a_t(\theta)=t; \mathcal{P}_t=0$; $\lambda_t(\theta)=1/t$ et on a

$$\lim_{t\longrightarrow\infty} a_t(\theta)=t=\infty, \text{ et } \quad \lim_{t\longrightarrow\infty}\frac{\Delta a_t(\theta)}{a_{t-1}(\theta)}=\frac{t-(t-1)}{t-1}=\frac{1}{t-1}=0 \;\; P^\theta p.s$$

Donc $(R1)$ est satisfaite et $\frac{\Delta a_t(\theta)}{a_t(\theta)}-\lambda_t(\theta)=\frac{1}{t}-\frac{1}{t}=0$ implique que

$$\sum_{t=1}^{\infty}\left[\frac{\Delta a_t(\theta)}{a_t(\theta)}-\lambda_t(\theta)\right]^+=0 \;\; P^\theta p.s$$

D'ou la convergence définie par (1.3.5). En prenant $\Delta_{t-1}=u=\hat{\theta}-\theta$ assez petit puisque $\hat{\theta}_t \longrightarrow \theta \; P^\theta p.s$, et de $(B1)$, nous avons

$$\begin{aligned}\left\langle C_\theta u, \gamma^{-1}(\theta+u)E^\theta\psi(\theta+u,X_1)\right\rangle &= \left\langle C_\theta\Delta_{t-1}, \gamma^{-1}(\theta+\Delta_{t-1})b_t(\theta+\Delta_{t-1})\right\rangle \\ &\le -\frac{1}{2}\left\langle C_\theta u, u\right\rangle\end{aligned}$$

Alors

$$\begin{aligned}2\left\langle C_\theta\Delta_{t-1}, \frac{1}{t}\gamma^{-1}(\theta+\Delta_{t-1})b_t(\theta+\Delta_{t-1})\right\rangle &= \left\langle C_\theta\Delta_{t-1}, \Gamma_t^{-1}(\theta+\Delta_{t-1})b_t(\theta+\Delta_{t-1})\right\rangle \\ &\le -\frac{1}{t}\left\langle C_\theta\Delta_{t-1}, \Delta_{t-1}\right\rangle\end{aligned}$$

Ainsi la condition $(R2)$ est satisfaite.

Nous avons d'après $(B2)$ pour $\varepsilon \in]0,1[$

$$\begin{aligned}& a_t^\varepsilon(\theta)\left[E_\theta\left(\left\|\Gamma_t^{-1}(\theta+\Delta_{t-1})\,\psi_t(\theta+\Delta_{t-1})\right\|^2/\mathcal{F}_{t-1}\right)-\mathcal{P}_t\right]^+ \\ =\;& t^\varepsilon\left[E_\theta\left(\left\|\frac{1}{t}\gamma^{-1}(\theta+u)\psi(\theta+u)\right\|^2/\mathcal{F}_{t-1}\right)\right]^+ \\ =\;& t^\varepsilon E_\theta\left(\left\|\frac{1}{t}\gamma^{-1}(\theta+u)\psi(\theta+u)\right\|^2/\mathcal{F}_{t-1}\right) \\ =\;& t^{\varepsilon-2}E_\theta\left\|\gamma^{-1}(\theta+u)\psi(\theta+u)\right\|^2\end{aligned}$$

Par suit de $(B2)$:$E_\theta \left\|\gamma^{-1}(\theta+u)\psi(\theta+u)\right\|^2 = O(1)$ on deduit du fait que $2-\varepsilon > 1$

$$\sum_{t=1}^{\infty} \frac{1}{t^{2-\varepsilon}} E_\theta \left\|\gamma^{-1}(\theta+u)\psi(\theta+u)\right\|^2 < \infty$$

Donc $(R3)$ et les conditions de corollaire [1.1] sont satisfaites. On tire que pour tout $\delta \in]0, 1/2[$: $t^{2\delta}(\hat{\theta}-\theta)^T C_\theta(\hat{\theta}-\theta) \longrightarrow 0 \quad P^\theta p.s.$ ∎

Remarque 5 *Dans Remarque 3 cas iid on a vu que les procedures récursives sont utilisées dans les algorithmes stochastiques où les conditions qui donnent une bonne vitesse de convergence sont exprimeés en terme de stabilité de matrices. Une matrice A est dite* stable *si les parties réelles de ses valeurs propres sont négatives. Dans les algorithmes stochastique on exige l'existence de la représentation (voir Remarque 2)*

$$R^\theta(u) = B^\theta u + o(\|u\|) \quad \textit{quand} \quad u \longrightarrow 0 \tag{1.4.2}$$

où $S^\theta = B^\theta + \frac{1}{2}\mathbf{1}$ est stable. Donc le maximum des parties réelles des valeurs propres de B^θ est infereur à $-1/2$. Par suit (cf.[6],Ch.6,§3,Corollaire 3.1) il existe une matrice C_θ symétrique définie positive telle que

$$\left\langle C_\theta u, B^\theta u \right\rangle < -\frac{1}{2}\left\langle C_\theta u, u\right\rangle$$

implique que $B^\theta u < -\frac{1}{2}u$ or $R^\theta(u) = \gamma^{-1}(\theta+u)b(\theta,u) = \gamma^{-1}(\theta+u)E_\theta\psi(\theta+u, X_1)$ donc

$$\begin{aligned} \left\langle C_\theta u, \gamma^{-1}(\theta+u)E^\theta\psi(\theta+u,X_1)\right\rangle &= \left\langle C_\theta u, R^\theta(u)\right\rangle = \left\langle C_\theta u, B^\theta u\right\rangle \\ &\leq -\frac{1}{2}\left\langle C_\theta u, u\right\rangle \ \ \textit{(pour u assez petit)} \end{aligned}$$

Donc (1.4.2) *implique* $(B1)$.

Comme exemple considérons

$$f(\theta, x) = \frac{1}{\pi\left(1+(x-\theta)^2\right)}$$

la densité de Cauchy de paramètre $\theta \in \mathbb{R}$. Nous avons

$$\frac{\dot{f}}{f}(\theta, x) = \frac{2(x-\theta)}{1+(x-\theta)^2} \quad \text{et} \quad \frac{\partial^2}{\partial \theta^2} \log f(\theta, x) = \frac{2(x-\theta)^2 - 2}{\left(1+(x-\theta)^2\right)^2}$$

L'information de Fisher est

$$\begin{aligned} i(\theta) &= -\int \frac{\partial^2}{\partial \theta^2} \log f(\theta, x) f(\theta, x)\, dx \\ &= -\int_{-\infty}^{+\infty} \frac{2(x-\theta)^2 - 2}{\pi\left(1+(x-\theta)^2\right)^3} dx \end{aligned}$$

Le théorème des résidus donne

$$i(\theta) = \int_{-\infty}^{+\infty} g(z)dz = 2\pi i \sum_i res\left(g(z), z_i\right) = \int_{\Gamma} g(z)dz + \int_{-R}^{+R} g(z)dz$$

où Γ est le demi cercle (positif) de centre 0 et de diamètre $2R$ (si $z \in [-R, +R]$ on pose $z = x$, et si $z \in \Gamma : z = Re^{i\theta}$). Par conséquent

$$\begin{aligned} i(\theta) &= \int_{\Gamma} -\frac{2(z-\theta)^2 - 2}{\pi\left(1+(z-\theta)^2\right)^3} dz + \int_{-R}^{+R} -\frac{2(x-\theta)^2 - 2}{\pi\left(1+(x-\theta)^2\right)^3} dx \\ &= 2\pi i \sum_i res\left(g(z), z_i\right) \end{aligned}$$

Calculons les points de singularité

$$\left(1+(z-\theta)^2\right)^3 = 0 \Longleftrightarrow \left(\begin{array}{ll} z = i + \theta & \text{pôle triple} \\ z = -i + \theta \notin \Gamma & \end{array} \right)$$

Donc

$$res\left(f(z), i+\theta\right) = \frac{1}{2!} \frac{d^2}{dz^2} (z - i - \theta)^3 f(z) = \frac{1}{4\pi i}$$

Nous avons ($z = Re^{i\theta}$)

$$\lim_{R\to\infty} -\frac{2(z-\theta)^2-2}{\pi\left(1+(z-\theta)^2\right)^3} = \lim_{R\to\infty} -\frac{R^2}{R^6} = 0$$

Donc

$$i(\theta) = \int_{-\infty}^{+\infty} -\frac{2(x-\theta)^2-2}{\pi\left(1+(x-\theta)^2\right)^3}dx = 2\pi i\sum_i res\left(f(z), z_i\right) = 2\pi i\left(\frac{1}{4\pi i}\right) = \frac{1}{2}$$

Par suit

$$I_t(\theta) = \sum_{s=1}^{t} i_s(\theta) = \sum_{s=1}^{t}\frac{1}{2} = \frac{t}{2} \quad \text{et} \quad l_t(\theta) = \frac{\dot{f}}{f}(\theta, x) = \frac{2(x-\theta)}{1+(x-\theta)^2}$$

Ainsi l'estimateur de maximum du vraissemblance est donné par

$$\hat{\theta}_t = \hat{\theta}_{t-1} + I_t^{-1}\left(\hat{\theta}_{t-1}\right) l_t\left(\hat{\theta}_{t-1}\right) = \hat{\theta}_{t-1} - \frac{2}{t}\frac{2\left(X_t-\hat{\theta}_{t-1}\right)}{1+\left(X_t-\hat{\theta}_{t-1}\right)^2}$$

Vérifions les conditions du proposition 4. On a

$$\begin{aligned} b_t(\theta, u) &= b(\theta, u) = \int \psi(\theta+u, x) f(\theta, x)\mu(dx) \\ &= \int \frac{2(x-u)}{1+(x-u)^2}\frac{1}{\pi(1+x^2)}dx \end{aligned}$$

Par le théorème des résidus

$$\begin{aligned} b(\theta, u) &= \frac{2}{\pi}\int_{\Gamma} \frac{z-u}{1+(z-u)^2}\frac{dz}{1+z^2} + \frac{2}{\pi}\int_{-R}^{+R}\frac{x-u}{1+(x-u)^2}\frac{dx}{1+x^2} \\ &= 2\pi i\sum_i res\left(h(z), z_i\right) \end{aligned}$$

Les points de singularité sont $\left(1+(z-u)^2\right)\left(1+z^2\right) = 0$: $z = i+u, z = -i+u \notin \Gamma, z = i$, $z = -i \notin \Gamma$ et on a

$$res(h, i+u) = \lim_{z\to i+u}(z-i-u)\,h(z) = \frac{1}{2u(u+2i)}$$

$$res\,(h,i) = \lim_{z\to i}(z-i)\,h(z) = \frac{i-u}{-2iu\,(2i-u)}$$

Mais $z\in\Gamma$ $(z=Re^{i\theta})$ donc

$$\lim_{R\to\infty}\frac{z-u}{\left(1+(z-u)^2\right)(1+z^2)} = \lim_{R\to\infty}\frac{z}{z^4} = 0$$

Donc

$$\begin{aligned} b\,(\theta,u) &= 2\pi i\sum_i res\,(h(z),z_i) \\ &= 2\pi i\left[\frac{1}{2u\,(u+2i)}+\frac{i-u}{-2iu\,(2i-u)}\right] = -\frac{2u}{4+u^2} \end{aligned}$$

D'autre part, de la même facon on a par le théorème des residus:

$$\begin{aligned} J &: = \int\left(\frac{\dot f}{f}(\theta+u,x)\right)^2 f\,(\theta,x)\,dx = \int\left(\frac{2\,(x-u)}{1+(x-u)^2}\right)^2\frac{dx}{\pi\,(1+x^2)} \\ &= \frac{4}{\pi}\int_\Gamma\left(\frac{z-u}{1+(z-u)^2}\right)^2\frac{dz}{1+z^2}+\frac{4}{\pi}\int_{-R}^{+R}\left(\frac{x-u}{1+(x-u)^2}\right)^2\frac{dx}{1+x^2} \\ &= 2\pi i\sum_i res\,(k(z),z_i) \end{aligned}$$

Les pôles sont: $z=i, z=-i\notin\Gamma, z=i+u$ pôle double, $z=-i+u\notin\Gamma$ et

$$res\,(k(z),i) = \lim_{z\to i}(z-i)\,k(z) = \frac{(i-u)^2}{2iu^2\,(2i-u)^2}$$

$$res\,(k(z),i+u) = \frac{1}{1!}\frac{d}{dz}\,(z-i-u)^2\,k(z) = \frac{u\,(2i+u)-2i\,(i+u)}{4iu^2\,(2i+u)^2}$$

On a

$$\lim_{R\to\infty}\left(\frac{z-u}{1+(z-u)^2}\right)^2\frac{1}{1+z^2} = \lim_{R\to\infty}\frac{z^2}{z^6} = 0$$

Par conséquent par un calcul

$$J = \frac{4}{\pi}\int\left(\frac{x-u}{1+(x-u)^2}\right)^2\frac{dx}{1+x^2} = 2\pi i\sum_i res\,(k(z),z_i) = 2\frac{4+3u^2}{(4+u^2)^2}$$

Maintenant verifions les conditions $(B1)$,$(B2)$ du proposition 4 pour $\gamma(\theta) = I_t(\theta)/t$ et $l_t(\theta) = \psi(\theta, X_t)$. Pour $0 < \varepsilon < 1/2$ on a

$$\begin{aligned}
\frac{i^{-1}(\theta)\, b(\theta,u)}{u} &= \gamma^{-1}(\theta+u) E^\theta \psi(\theta+u, X_1)/u = \frac{2}{u}\left(-\frac{2u}{4+u^2}\right) = -\frac{4}{4+u^2} \\
&= -1 + \frac{u^2}{4+u^2} \leq -1+\varepsilon \leq -\frac{1}{2} \quad \text{(pour } u \text{ assez petit)}
\end{aligned}$$

donc si on choisit $C_\theta = 1$:

$$\left\langle u, \gamma^{-1}(\theta+u) E^\theta \psi(\theta+u, X_1)\right\rangle = \left\langle u, 2\left(-\frac{2u}{4+u^2}\right)\right\rangle \leq -\frac{1}{2}\langle u, u\rangle$$

Donc $(B1)$ est satisfaite, et pour $(B2)$ on a

$$\begin{aligned}
E_\theta \left\|\gamma^{-1}(\theta+u)\,\psi(\theta+u)\right\|^2 &= E_\theta \left\| 2\left(\frac{2(X_1-u)}{1+(X_1-u)^2}\right)\right\|^2 \\
&= 16 \int \left(\frac{x-u}{1+(x-u)^2}\right)^2 \frac{1}{1+x^2} dx \\
&= 8\pi \frac{4+3u^2}{(4+u^2)^2} \leq 8\pi(1+2\varepsilon) \quad \text{quand } u \longrightarrow 0
\end{aligned}$$

donc

$$E_\theta \left\|\gamma^{-1}(\theta+u)\psi(\theta+u)\right\|^2 = O(1) \quad \text{quand} \quad u \longrightarrow 0$$

et on a $\hat{\theta}_t \longrightarrow \theta$ (cf.[13].Corollaire 4.1). Alors les conditions du proposition 4 sont satisfaites, donc $\forall \delta \in]0, 1/2[: t^\delta \left(\hat{\theta}_t - \theta\right) \longrightarrow 0 \ P^\theta p.s.$

2. Processus de Markov

Nous considèrons la famille exponentielle de processus de Markov au sens de Feigin (1981,[4]). C'est une chaine de Markov homogène de densité de transition d'une etape donnée par

$$f(y; \theta, x) = h(x,y) \exp\left(\theta^T m(y,x) - \beta(\theta; x)\right)$$

où $m(y,x) \in R^m, \beta(\theta; x) \in R^m$. Avec nos notations $f_t(\theta) = f(X_t; \theta; X_{t-1})$ et nous avons

$$l_t(\theta) = \frac{d}{d\theta} \log f_t(\theta) = m(X_t, X_{t-1}) - \dot{\beta}^T(\theta; X_{t-1})$$

Dans ce cas $l_t(\theta)$ est une difference de martingale :$E_\theta\left(l_t(\theta)/\mathcal{F}_{t-1}\right) = 0$ (cf.[4]) donc

$$E_\theta\left(m(X_t, X_{t-1})/\mathcal{F}_{t-1}\right) = \dot{\beta}^T(\theta; X_{t-1})$$

Or

$$i_t(\theta) = -E_\theta\left(\frac{d^2}{d\theta^2}\log f_t(\theta)/\mathcal{F}_{t-1}\right) = -E_\theta\left(-\ddot{\beta}(\theta; X_{t-1})/\mathcal{F}_{t-1}\right) = \ddot{\beta}(\theta; X_{t-1})$$

et l'information de Fisher

$$I_t(\theta) = \sum_{s=1}^{t} i_s(\theta) = \sum_{s=1}^{t} \ddot{\beta}(\theta; X_{t-1})$$

Pour $\psi_t := l_t, \Gamma_t := I_t$, l'estimateur MV verifie

$$\begin{aligned}
\hat{\theta}_t &= \hat{\theta}_{t-1} + I_t^{-1}(\hat{\theta}_{t-1}) l_t(\hat{\theta}_{t-1}) \\
&= \hat{\theta}_{t-1} + \left(\sum_{s=1}^{t} \ddot{\beta}(\hat{\theta}_{t-1}; X_{s-1})\right)^{-1} \left(m(X_t, X_{t-1}) - \dot{\beta}^T(\hat{\theta}_{t-1}; X_{t-1})\right), \quad t \geq 1
\end{aligned}$$

Mais de $E_\theta\left(m(X_t, X_{t-1})/\mathcal{F}_{t-1}\right) = \dot{\beta}^T(\theta; X_{t-1})$ on a

$$\begin{aligned}
i_t(\theta) &= \ddot{\beta}(\theta; X_{t-1}) = E_\theta\left(l_t(\theta) l_t^T(\theta)/\mathcal{F}_{t-1}\right) \\
&= E_\theta\left[\left(m(X_t, X_{t-1}) - \dot{\beta}^T(\theta, X_{t-1})\right)\left(m^T(X_t, X_{t-1}) - \dot{\beta}(\theta, X_{t-1})\right)/\mathcal{F}_{t-1}\right] \\
&= E_\theta\left(m(X_t, X_{t-1}) m^T(X_t, X_{t-1})/\mathcal{F}_{t-1}\right) - E_\theta\left(m(X_t, X_{t-1})/\mathcal{F}_{t-1}\right)\dot{\beta}(\theta, X_{t-1}) \\
&\quad -\dot{\beta}^T(\theta, X_{t-1}) E_\theta\left(m^T(X_t, X_{t-1})/\mathcal{F}_{t-1}\right) + \dot{\beta}^T(\theta, X_{t-1})\dot{\beta}(\theta, X_{t-1}) \\
&= E_\theta\left(m(X_t, X_{t-1}) m^T(X_t, X_{t-1})/\mathcal{F}_{t-1}\right) - \dot{\beta}^T(\theta, X_{t-1})\dot{\beta}(\theta, X_{t-1}) \\
&\quad -\dot{\beta}^T(\theta, X_{t-1})\dot{\beta}(\theta, X_{t-1}) + \dot{\beta}^T(\theta, X_{t-1})\dot{\beta}(\theta, X_{t-1}) \\
&= E_\theta\left(m(X_t, X_{t-1}) m^T(X_t, X_{t-1})/\mathcal{F}_{t-1}\right) - \dot{\beta}^T(\theta, X_{t-1})\dot{\beta}(\theta, X_{t-1}) = \ddot{\beta}(\theta; X_{t-1})
\end{aligned}$$

Ce qui implique que

$$E_\theta\left(m(X_t, X_{t-1}) m^T(X_t, X_{t-1})/\mathcal{F}_{t-1}\right) = \ddot{\beta}(\theta; X_{t-1}) + \dot{\beta}^T(\theta, X_{t-1})\dot{\beta}(\theta, X_{t-1}) \qquad (1.4.3)$$

Comme $\dot{\beta}^T(\theta+u,X_{t-1})$ est $\mathcal{F}_{t-1}$-mesurable alors

$$\begin{aligned} b_t(\theta,u) &= E_\theta\left(l_t(\theta+u)/\mathcal{F}_{t-1}\right) \\ &= E_\theta\left(m(X_t,X_{t-1})-\dot{\beta}^T(\theta+u,X_{t-1})/\mathcal{F}_{t-1}\right) \\ &= E_\theta\left(m(X_t,X_{t-1})/\mathcal{F}_{t-1}\right)-\dot{\beta}^T(\theta+u,X_{t-1}) \\ &= \dot{\beta}^T(\theta,X_{t-1})-\dot{\beta}^T(\theta+u,X_{t-1}) \end{aligned} \tag{1.4.4}$$

Nous avons: $(u^Tu=\|u\|^2=tr\left(uu^T\right)$, et $tr\,(A+B)=trA+trB)$, donc

$$\begin{aligned} \|b_t(\theta,u)\|^2 &= \left\|\dot{\beta}^T(\theta,X_{t-1})-\dot{\beta}^T(\theta+u,X_{t-1})\right\|^2 \\ &= tr\left[\left(\dot{\beta}^T(\theta,X_{t-1})-\dot{\beta}^T(\theta+u,X_{t-1})\right)\left(\dot{\beta}^T(\theta,X_{t-1})-\dot{\beta}^T(\theta+u,X_{t-1})\right)^T\right] \\ &= \left(\dot{\beta}^T(\theta,X_{t-1})-\dot{\beta}^T(\theta+u,X_{t-1})\right)^T\left(\dot{\beta}^T(\theta,X_{t-1})-\dot{\beta}^T(\theta+u,X_{t-1})\right) \\ &= \left(\dot{\beta}(\theta,X_{t-1})-\dot{\beta}(\theta+u,X_{t-1})\right)\left(\dot{\beta}^T(\theta,X_{t-1})-\dot{\beta}^T(\theta+u,X_{t-1})\right) \\ &= \dot{\beta}(\theta,X_{t-1})\dot{\beta}^T(\theta,X_{t-1})-\dot{\beta}(\theta,X_{t-1})\dot{\beta}^T(\theta+u,X_{t-1}) \\ &\quad -\dot{\beta}(\theta+u,X_{t-1})\dot{\beta}^T(\theta,X_{t-1})+\dot{\beta}(\theta+u,X_{t-1})\dot{\beta}^T(\theta+u,X_{t-1}) \end{aligned}$$

et

$$\begin{aligned} \|l_t(\theta+u)\|^2 &= \left\|m(X_t,X_{t-1})-\dot{\beta}^T(\theta+u,X_{t-1})\right\|^2 \\ &= \left(m(X_t,X_{t-1})-\dot{\beta}^T(\theta+u,X_{t-1})\right)^T\left(m(X_t,X_{t-1})-\dot{\beta}^T(\theta+u,X_{t-1})\right) \\ &= tr\left[\left(m(X_t,X_{t-1})-\dot{\beta}^T(\theta+u,X_{t-1})\right)\left(m^T(X_t,X_{t-1})-\dot{\beta}(\theta+u,X_{t-1})\right)\right] \\ &= tr\left(m(X_t,X_{t-1})m^T(X_t,X_{t-1})\right)-tr\left(m(X_t,X_{t-1})\dot{\beta}(\theta+u,X_{t-1})\right) \\ &\quad -tr\left(\dot{\beta}^T(\theta+u,X_{t-1})m^T(X_t,X_{t-1})\right)+tr\left(\dot{\beta}^T(\theta+u,X_{t-1})\dot{\beta}(\theta+u,X_{t-1})\right) \end{aligned}$$

En utilisant (1.4.3), (1.4.4) nous avons

$$
\begin{aligned}
E_\theta\left(\left\|l_t(\theta+u)\right\|^2/\mathcal{F}_{t-1}\right) &= tr\ddot{\beta}(\theta;X_{t-1}) + \left\|\dot{\beta}^T(\theta,X_{t-1}) - \dot{\beta}^T(\theta+u,X_{t-1})\right\|^2 \\
&= tr\ddot{\beta}(\theta;X_{t-1}) + \left\|b_t(\theta,u)\right\|^2 \qquad (1.4.5)
\end{aligned}
$$

Ces expressions sont utilisees pour vérifier les conditions du Corollaire [1.1] pour des choix des fonctions m,β.

Supposons que $\theta \in \mathbb{R}$ et que

$$
f(y;\theta,x) = h(x,y)\exp\left(\theta m(y,x) - \beta(\theta;x)\right)
$$

où $\beta(\theta;x) = \gamma(\theta)h(x), h(.) \geq 0$ et $\ddot{\gamma}(.) \geq 0$ (cf.[4]). Alors

$$
l_t(\theta) = \frac{d}{d\theta}\log f(X_t;\theta,X_{t-1}) = m(X_t,X_{t-1}) - \dot{\gamma}(\theta)h(X_{t-1})
$$

et

$$
\frac{d^2}{d\theta^2}\log f(X_t;\theta,X_{t-1}) = -\ddot{\gamma}(\theta)h(X_{t-1})
$$

Mais $\ddot{\beta}(\theta;X_{t-1}) = \ddot{\gamma}(\theta)h(X_{t-1})$ est $\mathcal{F}_{t-1}$-mesurable, donc

$$
i_t(\theta) = -E_\theta\left(\frac{d^2}{d\theta^2}\log f(X_t;\theta,X_{t-1})/\mathcal{F}_{t-1}\right) = \ddot{\gamma}(\theta)h(X_{t-1})
$$

et

$$
\begin{aligned}
I_t(\theta) &= \sum_{s=1}^{t} i_s(\theta) = \sum_{s=1}^{t}\ddot{\gamma}(\theta)h(X_{s-1}) \\
&= \ddot{\gamma}(\theta)\sum_{s=1}^{t}h(X_{s-1}) = \ddot{\gamma}(\theta)H_t
\end{aligned}
$$

Si $\ddot{\gamma}(\theta) \neq 0$, l'estimateur MV verifie

$$\begin{aligned}
\hat{\theta}_t &= \hat{\theta}_{t-1} + I_t^{-1}(\hat{\theta}_{t-1}) l_t(\hat{\theta}_{t-1}) \\
&= \hat{\theta}_{t-1} + \frac{1}{\ddot{\gamma}(\theta_{t-1}) H_t} \Big(m(X_t, X_{t-1}) - \dot{\gamma}(\hat{\theta}_{t-1}) h(X_{t-1}) \Big)
\end{aligned} \tag{1.4.6}$$

Le résultat suivant donne les conditions suffissantes pour la convergence de (1.4.6).

Proposition 6 *Supposons que $H_t \longrightarrow \infty$ $P^\theta p.s$, et que $\dot{\gamma}$ est lineaire ou les conditions suivantes sont satisfaites*

$(M1)$

$$\frac{h(X_{t-1})}{H_t} \longrightarrow 0 \ P^\theta p.s$$

$(M2)$ *pour tout a,b finie*

$$0 < \inf_{u \in [a,b]} \ddot{\gamma}(u) \leq \sup_{u \in [a,b]} \ddot{\gamma}(u) < \infty$$

$(M3)$ *il existe une constante B, telle que pour tout $u \in \mathbb{R}$*

$$\frac{1 + \dot{\gamma}^2(u)}{\ddot{\gamma}^2(u)} \leq B(1 + u^2)$$

Alors $\hat{\theta}_t$ défini par (1.4.6) *est convergent:$\hat{\theta}_t \longrightarrow \theta$ $P^\theta p.s$ pour toute valeur initiale $\hat{\theta}_0$.*

Preuve: Nous utilisons Théorème 7 (cf Ch 2). Vérifions les conditions $(G1), (G2)(G3)$ du Théorème 7 avec $\psi_t(\theta) = l_t(\theta) = m(X_t, X_{t-1}) - \dot{\gamma}(\theta) h(X_{t-1})$, $\Gamma_t(\theta) = I_t(\theta) = H_t \ddot{\gamma}(\theta)$, et $V_t = u^2$ ($V_t(0) = 0$, $\inf_u V_\theta(u) = 0$, et $\inf_{\|u\| \geq \varepsilon} V_\theta(u) > 0$), donc $(G1)$ est satisfaite.

Pour $(G3)$

$$\begin{aligned}
\mathcal{N}_t(u) &= \dot{V}_\theta(u) \Gamma_t^{-1}(\theta + u) E_\theta(\psi_t(\theta + u) / \mathcal{F}_{t-1}) \\
&\quad + \frac{1}{2} \sup_\nu \left\| \ddot{V}_\theta(\nu) \right\| E_\theta\left(\left\| \Gamma_t^{-1}(\theta + u) \psi_t(\theta + u) \right\|^2 / \mathcal{F}_{t-1} \right)
\end{aligned}$$

Mais $\dot{V}_\theta(u) = 2u$ et $\ddot{V}_\theta(u) = 2$, $\sup_\nu \left\| \ddot{V}_\theta(\nu) \right\| = 2$, $\Gamma_t(\theta + u) = H_t \ddot{\gamma}(\theta + u)$;

$$\begin{aligned} E_\theta\left(\psi_t(\theta + u) / \mathcal{F}_{t-1}\right) &= b_t(\theta, u) = \dot{\beta}^T(\theta, X_{t-1}) - \dot{\beta}^T(\theta + u, X_{t-1}) \\ &= h(X_{t-1})\left(\dot{\gamma}(\theta) - \dot{\gamma}(\theta + u)\right) \ \text{(d'aprés (1.4.4))} \end{aligned} \tag{1.4.7}$$

et de (1.4.5)

$$E_\theta\left(l_t^2(\theta + u) / \mathcal{F}_{t-1}\right) = \ddot{\gamma}(\theta) h(X_{t-1}) + b_t^2(\theta, u) \tag{1.4.8}$$

et

$$\begin{aligned} \left\|\Gamma_t^{-1}(\theta + u) \psi_t(\theta + u)\right\|^2 &= \left\| \frac{m(X_t, X_{t-1}) - \dot{\gamma}(\theta + u) h(X_{t-1})}{H_t \ddot{\gamma}(\theta + u)} \right\|^2 \\ &= \frac{\left(m(X_t, X_{t-1}) - \dot{\gamma}(\theta + u) h(X_{t-1})\right)^2}{H_t^2 \ddot{\gamma}^2(\theta + u)} \end{aligned}$$

Par suit

$$\begin{aligned} \mathcal{N}_t(u) &= 2u \frac{1}{H_t \ddot{\gamma}(\theta + u)} b_t(\theta, u) + \frac{1}{H_t^2 \ddot{\gamma}^2(\theta + u)} E_\theta\left(l_t^2(\theta + u) / \mathcal{F}_{t-1}\right) \\ &= 2u \frac{1}{H_t \ddot{\gamma}(\theta + u)} \left[h(X_{t-1})\left(\dot{\gamma}(\theta) - \dot{\gamma}(\theta + u)\right)\right] \\ &\quad + \frac{1}{H_t^2 \ddot{\gamma}^2(\theta + u)} \left(\ddot{\gamma}(\theta) h(X_{t-1}) + b_t^2(\theta, u)\right) \\ &= 2u \frac{h(X_{t-1})}{H_t} \frac{\dot{\gamma}(\theta) - \dot{\gamma}(\theta + u)}{\ddot{\gamma}(\theta + u)} + \frac{h(X_{t-1})}{H_t^2} \frac{\ddot{\gamma}(\theta)}{\ddot{\gamma}^2(\theta + u)} \\ &\quad + \frac{h^2(X_{t-1})}{H_t^2} \frac{\left(\dot{\gamma}(\theta) - \dot{\gamma}(\theta + u)\right)^2}{\ddot{\gamma}^2(\theta + u)} \\ &= \frac{h(X_{t-1})}{H_t} \frac{\dot{\gamma}(\theta) - \dot{\gamma}(\theta + u)}{\ddot{\gamma}(\theta + u)} u \left(2 + \frac{h(X_{t-1})}{H_t} \frac{\dot{\gamma}(\theta) - \dot{\gamma}(\theta + u)}{u \ddot{\gamma}(\theta + u)}\right) \\ &\quad + \frac{h(X_{t-1})}{H_t^2} \frac{\ddot{\gamma}(\theta)}{\ddot{\gamma}^2(\theta + u)} \\ &= : \mathcal{N}_{1t}(u) + \mathcal{N}_{2t}(u) \end{aligned}$$

Montrons que pour tout t assez grand on a

$$2 + \frac{h(X_{t-1})}{H_t} \frac{\dot{\gamma}(\theta) - \dot{\gamma}(\theta + u)}{u \ddot{\gamma}(\theta + u)} \geq 1$$

Si $\dot{\gamma}$ est linéaire on a $\frac{\dot{\gamma}(\theta)-\dot{\gamma}(\theta+u)}{u\ddot{\gamma}(\theta+u)} = \frac{-u\ddot{\gamma}(\xi)}{u\ddot{\gamma}(\theta+u)} = -1$ et $\frac{h(X_{t-1})}{H_t} = \frac{\Delta H_t}{H_t} \leq 1$ ($h(.) \geq 0$). Donc l'inégalité ci-dessus est vérifiée. Pour le cas non-linéaire, nous avons (avec $u \neq 0$)

$$\left|\frac{\dot{\gamma}(\theta)-\dot{\gamma}(\theta+u)}{u\ddot{\gamma}(\theta+u)}\right| = \frac{\ddot{\gamma}(\theta+\tilde{u})}{\ddot{\gamma}(\theta+u)}, \text{ où } |\tilde{u}| \leq |u| \tag{1.4.9}$$

Pour $|u| \leq M, 0 < M < \infty$, alors il s'ensuit de $(M2)$ que le membre de gauche de (1.4.9) est borné. Pour $|u| > M$, nous utilisons $(M3)$ et l'inégalité $(a-b)^2 \leq 2a^2 + 2b^2$, pour obtenir

$$\begin{aligned}
\frac{(\dot{\gamma}(\theta)-\dot{\gamma}(\theta+u))^2}{\ddot{\gamma}^2(\theta+u)} &\leq \frac{2\dot{\gamma}^2(\theta)+2\dot{\gamma}^2(\theta+u)}{\ddot{\gamma}^2(\theta+u)} = \frac{2\dot{\gamma}^2(\theta)-2\dot{\gamma}^2(\theta+u)+4\dot{\gamma}^2(\theta+u)}{\ddot{\gamma}^2(\theta+u)} \\
&\leq 2\frac{\dot{\gamma}^2(\theta)-\dot{\gamma}^2(\theta+u)}{\ddot{\gamma}^2(\theta+u)} + 4\frac{\dot{\gamma}^2(\theta+u)}{\ddot{\gamma}^2(\theta+u)} \\
&\leq 4\frac{\dot{\gamma}^2(\theta+u)}{\ddot{\gamma}^2(\theta+u)} \quad (\text{car } \dot{\gamma} \text{ est croissante par } (M2)) \\
&\leq 4\frac{1+\dot{\gamma}^2(\theta+u)}{\ddot{\gamma}^2(\theta+u)} - \frac{4}{\ddot{\gamma}^2(\theta+u)} \\
&\leq 4\frac{1+\dot{\gamma}^2(\theta+u)}{\ddot{\gamma}^2(\theta+u)} \leq \tilde{B}\left(1+u^2\right) \quad (\text{par } (M3))
\end{aligned}$$

Par suit $\left|\frac{\dot{\gamma}(\theta)-\dot{\gamma}(\theta+u)}{u\ddot{\gamma}(\theta+u)}\right|$ est bornée par $\sqrt{\tilde{B}\left(1+u^2\right)/u^2} = \sqrt{\tilde{B}\left(1/u^2+1\right)} \leq \sqrt{\tilde{B}\left(1/M^2+1\right)}$. Donc $\forall u$ le membre de gauche de (1.4.9) est borné. Comme $\frac{h(X_{t-1})}{H_t} \to 0$ l'inégalité $2 + \frac{h(X_{t-1})}{H_t}\frac{\dot{\gamma}(\theta)-\dot{\gamma}(\theta+u)}{u\ddot{\gamma}(\theta+u)} \geq 1$ est vérifiée pour t assez grand et $\ddot{\gamma}(.) \geq 0$ ceci impliquent que

$$\begin{aligned}
\mathcal{N}_{1t}(u) &= \frac{h(X_{t-1})}{H_t}\frac{\dot{\gamma}(\theta)-\dot{\gamma}(\theta+u)}{u\ddot{\gamma}(\theta+u)}u^2\left(2+\frac{h(X_{t-1})}{H_t}\frac{\dot{\gamma}(\theta)-\dot{\gamma}(\theta+u)}{u\ddot{\gamma}(\theta+u)}\right) \\
&= \frac{h(X_{t-1})}{H_t}\frac{-u\ddot{\gamma}(\xi)}{u\ddot{\gamma}(\theta+u)}u^2\left(2+\frac{h(X_{t-1})}{H_t}\frac{\dot{\gamma}(\theta)-\dot{\gamma}(\theta+u)}{u\ddot{\gamma}(\theta+u)}\right) \leq 0
\end{aligned}$$

donc pour tout t assez grand de $\mathcal{N}_t(u) = \mathcal{N}_{1t}(u) + \mathcal{N}_{2t}(u) \leq \mathcal{N}_{2t}(u)$ par conséquent $[\mathcal{N}_t(u)]^+ \leq [\mathcal{N}_{2t}(u)]^+ = \mathcal{N}_{2t}(u)$. En utilisant $(M3)$ et de facon similaire que ci dessus pour montrer que $\frac{\ddot{\gamma}(\theta)}{\ddot{\gamma}^2(\theta+u)} \leq B_1(1+u^2)$ on obtient

$$\frac{1}{1+u^2}[\mathcal{N}_t(u)]^+ \leq \frac{1}{1+u^2}[\mathcal{N}_{2t}(u)]^+ = \frac{1}{1+u^2}\frac{h(X_{t-1})}{H_t^2}\frac{\ddot{\gamma}(\theta)}{\ddot{\gamma}^2(\theta+u)} \leq \frac{h(X_{t-1})}{H_t^2}B_1$$

où B_1 peut dépendre de θ. Du Lemme 3 dans l'annexe 1, avec $H_t = \sum_{s=1}^{t} h(X_{s-1}), h(.) \geq 0$ et

$\varepsilon = 1$ alors

$$\begin{aligned}
\sum_{t=1}^{\infty} \left(1 + V_{\theta}\left(\Delta_{t-1}\right)\right)^{-1} \left[\mathcal{N}_{t}\left(\Delta_{t-1}\right)\right]^{+} &= \sum_{t=1}^{\infty} \frac{1}{1+u^{2}} \left[\mathcal{N}_{t}\left(u\right)\right]^{+} \\
&\leq B_{1} \sum_{t=1}^{\infty} \frac{h\left(X_{t-1}\right)}{H_{t}^{2}} = B_{1} \sum_{t=1}^{\infty} \frac{\Delta H_{t}}{H_{t}^{2}} < \infty,\ P^{\theta} p.s
\end{aligned}$$

Donc la condition $(G3)$ de Théorème 7 (Ch.2) est satisfaite.

Vérifiant la condition $(G2)$. Par $[x]^{-} \geq -x$, nous obtenons pour t assez grand

$$\begin{aligned}
\left[\mathcal{N}_{t}\left(u\right)\right]^{-} &\geq -\mathcal{N}_{t}\left(u\right) = -\mathcal{N}_{1t}\left(u\right) - \mathcal{N}_{2t}\left(u\right) \\
&= -\frac{h\left(X_{t-1}\right)}{H_{t}} \frac{\dot{\gamma}\left(\theta\right) - \dot{\gamma}\left(\theta+u\right)}{\ddot{\gamma}\left(\theta+u\right)} u \left(2 + \frac{h\left(X_{t-1}\right)}{H_{t}} \frac{\dot{\gamma}\left(\theta\right) - \dot{\gamma}\left(\theta+u\right)}{u\ddot{\gamma}\left(\theta+u\right)}\right) \\
&\quad -\frac{h\left(X_{t-1}\right)}{H_{t}^{2}} \frac{\ddot{\gamma}\left(\theta\right)}{\ddot{\gamma}^{2}\left(\theta+u\right)} \\
&\geq -\frac{h\left(X_{t-1}\right)}{H_{t}} \frac{\dot{\gamma}\left(\theta\right) - \dot{\gamma}\left(\theta+u\right)}{\ddot{\gamma}\left(\theta+u\right)} u - \frac{h\left(X_{t-1}\right)}{H_{t}^{2}} \frac{\ddot{\gamma}\left(\theta\right)}{\ddot{\gamma}^{2}\left(\theta+u\right)} \\
&= \frac{h\left(X_{t-1}\right)}{H_{t}} \frac{-\left(\dot{\gamma}\left(\theta\right) - \dot{\gamma}\left(\theta+u\right)\right)}{u\ddot{\gamma}\left(\theta+u\right)} u^{2} - \frac{h\left(X_{t-1}\right)}{H_{t}^{2}} \frac{\ddot{\gamma}\left(\theta\right)}{\ddot{\gamma}^{2}\left(\theta+u\right)} \\
&= \frac{h\left(X_{t-1}\right)}{H_{t}} \frac{\ddot{\gamma}\left(\theta+\tilde{u}\right)}{\ddot{\gamma}\left(\theta+u\right)} u^{2} - \frac{h\left(X_{t-1}\right)}{H_{t}^{2}} \frac{\ddot{\gamma}\left(\theta\right)}{\ddot{\gamma}^{2}\left(\theta+u\right)}
\end{aligned}$$

où $|\tilde{u}| \leq |u|$. Alors il s'ensuit de $(M2)$ que

$$\sup_{\varepsilon \leq |u| \leq \frac{1}{\varepsilon}} \frac{\ddot{\gamma}\left(\theta\right)}{\ddot{\gamma}^{2}\left(\theta+u\right)} < R \quad \text{et} \quad \inf_{\varepsilon \leq |u| \leq \frac{1}{\varepsilon}} \frac{\ddot{\gamma}\left(\theta+\tilde{u}\right)}{\ddot{\gamma}\left(\theta+u\right)} u^{2} > r > 0$$

où les constantes positives R,r qui peuvent dépendre de θ. Notons aussi que dans le cas $\dot{\gamma}$ linéaire on a

$$\frac{\dot{\gamma}\left(\theta\right) - \dot{\gamma}\left(\theta+u\right)}{u\ddot{\gamma}\left(\theta+u\right)} = \frac{-\dot{\gamma}\left(u\right)}{u\ddot{\gamma}\left(\theta+u\right)} = -1$$

Ce qui implique que

$$\left[\mathcal{N}_{t}\left(u\right)\right]^{-} \geq \frac{h\left(X_{t-1}\right)}{H_{t}} u^{2} - \frac{h\left(X_{t-1}\right)}{H_{t}^{2}} \frac{\ddot{\gamma}\left(\theta\right)}{\ddot{\gamma}^{2}\left(\theta+u\right)}$$

Dans cas général ($\dot{\gamma}$ non linéaire) nous avons de la minoration de $\left[\mathcal{N}_{t}\left(u\right)\right]^{-}$ precedente nous avons

$$\begin{aligned}\inf_{\varepsilon\le|u|\le\frac{1}{\varepsilon}}\left[\mathcal{N}_t\left(u\right)\right]^- &\ge \inf_{\varepsilon\le|u|\le\frac{1}{\varepsilon}}\frac{\ddot{\gamma}\left(\theta+\tilde{u}\right)}{\ddot{\gamma}\left(\theta+u\right)}u^2\frac{h\left(X_{t-1}\right)}{H_t}+\inf_{\varepsilon\le|u|\le\frac{1}{\varepsilon}}-\frac{\ddot{\gamma}\left(\theta\right)}{\ddot{\gamma}^2\left(\theta+u\right)}\frac{h\left(X_{t-1}\right)}{H_t^2}\\ &\ge \frac{h\left(X_{t-1}\right)}{H_t}r-R\frac{h\left(X_{t-1}\right)}{H_t^2}\end{aligned}$$

Par le lemme 3 dans l'annexe 1 avec $H_t=\sum_{s=1}^{t}h\left(X_{s-1}\right)\longrightarrow\infty$ et $\varepsilon=1$ on a:

$$\sum_{t=1}^{\infty}\frac{h\left(X_{t-1}\right)}{H_t}=\sum_{t=1}^{\infty}\frac{\Delta H_t}{H_t}=\infty \quad\text{et}\quad \sum_{t=1}^{\infty}\frac{h\left(X_{t-1}\right)}{H_t^2}=\sum_{t=1}^{\infty}\frac{\Delta H_t}{H_t^2}<\infty$$

Donc

$$\sum_{t=1}^{\infty}\inf_{\varepsilon\le|u|\le\frac{1}{\varepsilon}}\left[\mathcal{N}_t\left(u\right)\right]^-\ge\sum_{t=1}^{\infty}\frac{h\left(X_{t-1}\right)}{H_t}r-\sum_{t=1}^{\infty}R\frac{h\left(X_{t-1}\right)}{H_t^2}=\infty$$

Par suit la condition $(G2)$ est satisfaite. Par le Théorème 7 (Ch.2) nous déduisons que $\hat{\theta}_t\longrightarrow\theta$ $P^\theta p.s$ pour tout valeur initiale θ_0. ■

Nous étudions la vitesse de convergence dans le schema récursif precedent.

Corollaire [1.2] Supposons que $\hat{\theta}_t$ est défini par (1.4.6),$\hat{\theta}_t\longrightarrow\theta$ $P^\theta p.s$, et

1) $H_t\longrightarrow\infty$ $P^\theta p.s$

2) $\frac{h(X_t)}{H_t}\longrightarrow 0$ $P^\theta p.s$

3) $\ddot{\gamma}\left(.\right)$ est une fonction positive continue.

Alors $H_t^\delta\left(\hat{\theta}_t-\theta\right)\longrightarrow 0$ $P^\theta p.s$ pour tout $\delta\in\left]0,1/2\right[$.

Preuve: Vérifions les conditions du Corollaire [1.1] avec $\psi_t\left(\theta\right)=l_t\left(\theta\right)=m(X_t,X_{t-1})-\dot{\gamma}\left(\theta\right)h(X_{t-1})$, $\Gamma_t\left(\theta\right)=I_t\left(\theta\right)=H_t\ddot{\gamma}\left(\theta\right)$, $a_t\left(\theta\right)=H_t$, $C_\theta=1$, et

$$\mathcal{P}_t=\Gamma_t^{-2}\left(\theta+\Delta_{t-1}\right)b_t^2\left(\theta,\Delta_{t-1}\right)=H_t^{-2}\ddot{\gamma}^{-2}\left(\theta+\Delta_{t-1}\right)b_t^2\left(\theta,\Delta_{t-1}\right)$$

Comme $\Delta H_t=H_t-H_{t-1}=h(X_{t-1})$, $(R1)$ est traduite dans *2)* du corollaire.

Pour $(R2)$ comme $\dot{\gamma}\left(\theta\right)-\dot{\gamma}\left(\theta+u\right)=-u\ddot{\gamma}\left(\theta+\tilde{u}\right)$ où $|\tilde{u}|\le|u|$, et par (1.4.7) la partie gauche de (1.3.4) devient

$$\begin{aligned}
& 2\left\langle C_\theta \Delta_{t-1}, \Gamma_t^{-1}\left(\theta+\Delta_{t-1}\right) b_t\left(\theta, \Delta_{t-1}\right)\right\rangle+\mathcal{P}_t \\
= & 2 C_\theta \Delta_{t-1}^T \Gamma_t^{-1}\left(\theta+\Delta_{t-1}\right) b_t\left(\theta, \Delta_{t-1}\right)+\mathcal{P}_t \\
= & 2 \Delta_{t-1}^T H_t^{-1} \ddot{\gamma}^{-1}\left(\theta+\Delta_{t-1}\right) h(X_{t-1})(\dot{\gamma}(\theta)-\dot{\gamma}(\theta+u)) \\
& +H_t^{-2} \ddot{\gamma}^{-2}\left(\theta+\Delta_{t-1}\right)\left[h(X_{t-1})\left(-\ddot{\gamma}\left(\theta+\tilde{\Delta}_{t-1}\right) \Delta_{t-1}\right)\right]^2 \\
= & -2 \frac{h(X_{t-1})}{H_t} \frac{\ddot{\gamma}\left(\theta+\tilde{\Delta}_{t-1}\right)}{\ddot{\gamma}\left(\theta+\Delta_{t-1}\right)} \Delta_{t-1}^2+\frac{h^2(X_{t-1})}{H_t^2}\left(\frac{\ddot{\gamma}\left(\theta+\tilde{\Delta}_{t-1}\right)}{\ddot{\gamma}\left(\theta+\Delta_{t-1}\right)}\right)^2 \Delta_{t-1}^2
\end{aligned}$$

Comme $\ddot{\gamma}$ est continue et $\Delta_{t-1}=\hat{\theta}_{t-1}-\theta \longrightarrow 0$, pour tout petit $\tilde{\varepsilon}>0$ (qui peut être dépend de θ), nous avons $\left|\tilde{\Delta}_{t-1}\right| \leq\left|\Delta_{t-1}\right|$ et $1-\tilde{\varepsilon}<\frac{\ddot{\gamma}(\theta+\tilde{\Delta}_{t-1})}{\ddot{\gamma}(\theta+\Delta_{t-1})}<1+\tilde{\varepsilon}$ pour tout t assez grand. On a

$$-2 \frac{h(X_{t-1})}{H_t} \frac{\ddot{\gamma}\left(\theta+\tilde{\Delta}_{t-1}\right)}{\ddot{\gamma}\left(\theta+\Delta_{t-1}\right)} \leq-2(1-\tilde{\varepsilon}) \frac{h(X_{t-1})}{H_t}$$

et

$$0 \leq \frac{h^2(X_{t-1})}{H_t^2}\left(\frac{\ddot{\gamma}\left(\theta+\tilde{\Delta}_{t-1}\right)}{\ddot{\gamma}\left(\theta+\Delta_{t-1}\right)}\right)^2 \leq \frac{h^2(X_{t-1})}{H_t^2}(1+\tilde{\varepsilon})^2$$

donc

$$\begin{aligned}
& 2\left\langle C_\theta \Delta_{t-1}, \Gamma_t^{-1}\left(\theta+\Delta_{t-1}\right) b_t\left(\theta, \Delta_{t-1}\right)\right\rangle+\mathcal{P}_t \\
\leq & \left(-2(1-\tilde{\varepsilon}) \frac{h(X_{t-1})}{H_t}+\frac{h^2(X_{t-1})}{H_t^2}(1+\tilde{\varepsilon})^2\right) \Delta_{t-1}^2 \\
= & -\lambda_t(\theta)\left\langle C_\theta \Delta_{t-1}, \Delta_{t-1}\right\rangle
\end{aligned}$$

où

$$\lambda_t(\theta)=2(1-\tilde{\varepsilon}) \frac{h(X_{t-1})}{H_t}-(1+\tilde{\varepsilon})^2 \frac{h^2(X_{t-1})}{H_t^2} \quad \text{et} \quad C_\theta=1$$

Donc (1.3.4) est verifiée. Pour verifier la condition (1.3.5) on a

$$\begin{aligned}
\frac{h(X_{t-1})}{H_t}-\lambda_t(\theta) & =\frac{h(X_{t-1})}{H_t}\left(1-(2-2 \tilde{\varepsilon})+(1-\tilde{\varepsilon})^2 \frac{h(X_{t-1})}{H_t}\right) \\
& =\frac{h(X_{t-1})}{H_t}\left(-1+2 \tilde{\varepsilon}+(1-\tilde{\varepsilon})^2 \frac{h(X_{t-1})}{H_t}\right) \qquad (1.4.10)
\end{aligned}$$

Puisque $\tilde{\varepsilon}$ est arbitraire, on prend $-1+2\tilde{\varepsilon}<0$, $\frac{h(X_{t-1})}{H_t} \longrightarrow 0$, $h \geq 0$; $H_t \geq 0$ donc (1.4.10) est négative pour tout $t > T_0$ assez grand. Par suit la condition (1.3.5) devient

$$\sum_{s=1}^{\infty}\left[\frac{h(X_{t-1})}{H_t}-\lambda_t(\theta)\right]^+ = \sum_{s=1}^{T_0}\left[\frac{h(X_{t-1})}{H_t}-\lambda_t(\theta)\right]^+ < \infty \ P^\theta p.s$$

Donc $(R2)$ est satisfaite. Pour vérifier $(R3)$, et par (1.4.7),(1.4.8) on a

$$\begin{aligned}
& a_t^{\varepsilon}(\theta)\left[E_\theta\left(\Gamma_t^{-2}(\theta+\Delta_{t-1})\,\psi_t^2(\theta+\Delta_{t-1})/\mathcal{F}_{t-1}\right)-\mathcal{P}_t\right] \\
= \; & H_t^{\varepsilon}\left[E_\theta\left(H_t^{-2}\ddot{\gamma}^{-2}(\theta+\Delta_{t-1})\,l_t^2(\theta+\Delta_{t-1})/\mathcal{F}_{t-1}\right)-\mathcal{P}_t\right] \\
= \; & H_t^{\varepsilon}\left[H_t^{-2}\ddot{\gamma}^{-2}(\theta+\Delta_{t-1})\,E_\theta\left(l_t^2(\theta+\Delta_{t-1})/\mathcal{F}_{t-1}\right)-H_t^{-2}\ddot{\gamma}^{-2}(\theta+\Delta_{t-1})\,b_t^2(\theta,\Delta_{t-1})\right] \\
= \; & H_t^{\varepsilon-2}\ddot{\gamma}^{-2}(\theta+\Delta_{t-1})\left(\ddot{\gamma}(\theta)\,h(X_{t-1})+b_t^2(\theta,\Delta_{t-1})\right) \\
& -H_t^{\varepsilon-2}\ddot{\gamma}^{-2}(\theta+\Delta_{t-1})\,b_t^2(\theta,\Delta_{t-1}) \\
= \; & H_t^{\varepsilon-2}\ddot{\gamma}^{-2}(\theta+\Delta_{t-1})\,\ddot{\gamma}(\theta)\,h(X_{t-1}) = \frac{h(X_{t-1})}{H_t^{2-\varepsilon}}\frac{\ddot{\gamma}(\theta)}{\ddot{\gamma}^2(\theta+\Delta_{t-1})} \geq 0
\end{aligned}$$

D'aprés le Lemme 3 dans l'annexe 1, avec $d_t = H_t$;$\Delta_{t-1} \longrightarrow 0$, et $0<\varepsilon<1$, et $1+\acute{\varepsilon} = 2-\varepsilon \Longrightarrow \acute{\varepsilon} = 1-\varepsilon > 0$, on a $\frac{\ddot{\gamma}(\theta)}{\ddot{\gamma}^2(\theta+\Delta_{t-1})} \longrightarrow_{t\longrightarrow\infty} \frac{\ddot{\gamma}(\theta)}{\ddot{\gamma}^2(\theta)} = \ddot{\gamma}^{-1}(\theta) \geq 0$, qui ne dépend pas de t. Donc

$$\begin{aligned}
& \sum_{t=1}^{\infty} H_t^{\varepsilon}\left[E_\theta\left(\Gamma_t^{-2}(\theta+\Delta_{t-1})\,l_t^2(\theta+\Delta_{t-1})/\mathcal{F}_{t-1}\right)-\mathcal{P}_t\right] \\
= \; & \frac{\ddot{\gamma}(\theta)}{\ddot{\gamma}^2(\theta+\Delta_{t-1})}\sum_{t=1}^{\infty}\frac{h(X_{t-1})}{H_t^{2-\varepsilon}} < \infty \ P^\theta p.s
\end{aligned}$$

par conséquent $(R3)$ est satisfaite. Finalement par le Corollaire [1.1], nous avons $H_t^{\delta}\left(\hat{\theta}_t-\theta\right) \longrightarrow 0 \ P^\theta p.s$ pour tout $\delta \in \left]0,1/2\right[$. ■

Nous étudions un exemple de processus de markov qui est le processus autorégressif gaussien défini par

$$X_t = \theta X_{t-1} + Z_t, \qquad t = 1,2,...$$

où: $\theta \in \mathbb{R}, X_0 = 0$, et Z_t sont des va iid de densité gaussienne $\mathcal{N}(0,1)$. Dans ce modèle, la densité des transitions est donnée par

$$f(y;\theta,x) = h(x,y)\exp\left\{\theta xy - \frac{1}{2}x^2\theta^2\right\}$$

Si on pose $\gamma(\theta) = \theta^2/2$ et $h(x) = x^2$, on a

$$l_t(\theta) = \frac{d}{d\theta}\log f(X_t;\theta,X_{t-1}) = X_t X_{t-1} - X_{t-1}^2\theta$$

et

$$I_t(\theta) = \ddot{\gamma}(\theta) H_t = H_t = \sum_{s=1}^{t} h(X_{s-1}) = \sum_{s=1}^{t} X_{s-1}^2$$

Par suit

$$\hat{\theta}_t = \hat{\theta}_{t-1} + \frac{1}{I_t}\left(X_t X_{t-1} - X_{t-1}^2\hat{\theta}_{t-1}\right) \text{ où } I_t = I_{t-1} + X_{t-1}^2 \tag{1.4.11}$$

Nous notons que dans ce cas la vitesse dans l'information de Fisher conditionnelle I_t varie avec les valeurs de θ. On pose

$$\mathcal{K}_t(\theta) = \left\{\begin{array}{ll} t\left(1-\theta^2\right)^{-1} & \text{si } |\theta| < 1 \\ \frac{1}{2}t^2 & \text{si } |\theta| = 1 \\ \theta^{2t}\left(\theta^2-1\right)^{-2} & \text{si } |\theta| > 1 \end{array}\right\} \tag{1.4.12}$$

Pour $|\theta| < 1$, le processus $(X_t)_{t\geq 1}$ est strictement stationnaire et la loi forte des grands nombres donne (cf.[1],[18]):

$$I_t(\theta)/\mathcal{K}_t(\theta) = \frac{1}{\left(1-\theta^2\right)^{-1}}\frac{1}{t}\sum_{s=1}^{t} X_{s-1}^2 \longrightarrow_{t\longrightarrow\infty} \frac{1}{\left(1-\theta^2\right)^{-1}}E\left(X_1^2\right) = \frac{\left(1-\theta^2\right)^{-1}}{\left(1-\theta^2\right)^{-1}} = 1$$

Pour $|\theta| > 1$, nous avons (cf.[1],[18]) $I_t(\theta)/\mathcal{K}_t(\theta) \longrightarrow_{t\longrightarrow\infty} \chi^2(1)$ $p.s.$ Pour $|\theta| = 1$ la v.a.$I_t(\theta)/\mathcal{K}_t(\theta)$ converge en loi mais non en probabilité (cf.[1],[18]). On aussi $I_t(\theta) \longrightarrow_{t\longrightarrow\infty} \infty$ $p.s$ pour tout θ (cf.[17],Ch.VII,5.5). D'une part $\dot{\gamma}(\theta) = \theta$ est linéaire, et $H_t = I_t$ alors les conditions de Proposition 6 sont satisfaites. Par conséquent, l'estimateur récursif $\hat{\theta}_t$ est convergent, $\forall\theta \in \mathbb{R}$ et $\forall\theta_0$.

Pour établir la vitesse de convergence, on se place dans le cas $|\theta| < 1$, et vérifions les conditions du Corollaire [1.2]. D'une part comme $H_t = I_t$ et la loi forte des grands nombres:

$$\lim_{t\longrightarrow\infty}\frac{1}{t}I_t(\theta)=\lim_{t\longrightarrow\infty}\frac{1}{t}\sum_{s=1}^{t}X_{s-1}^2=E\left(X_1^2\right)=\frac{1}{1-\theta^2}<\infty \;\; P^\theta p.s \tag{1.4.13}$$

implique $H_t=I_t\to\infty$ $p.s.$(cf.[17]). En posant $d_t:=\frac{t}{t+1}\frac{I_{t+1}}{I_t}=\frac{I_{t+1}/(t+1)}{I_t/t}\longrightarrow_{t\longrightarrow\infty}1$ $p.s$ par (1.4.13) alors

$$\frac{\Delta I_{t+1}}{I_t}:=\frac{I_{t+1}-I_t}{I_t}=\frac{I_{t+1}}{I_t}-1=\frac{t+1}{t}d_t-1\longrightarrow 0$$

Par suit $\frac{h(X_t)}{I_t}=\frac{X_t^2}{I_t}=\frac{\Delta I_{t+1}}{I_t}\longrightarrow_{t\longrightarrow\infty}0$ $P^\theta p.s.$ En fin $\ddot{\gamma}(\theta)=1$ et donc les conditions du Corollaire [1.2] sont satisfaites et on a la vitesse de convergence dans le cas $|\theta|<1$ et pour tout $\delta\in]0,1/2[$: $I_t^\delta(\hat{\theta}_t-\theta)\to 0$. Notons que pour le cas $|\theta|>1$, la condition *2)* du Corollaire [1.2] n'est pas toujours satisfaite. En effet on a $d_t:=\frac{I_{t+1}/\mathcal{K}_{t+1}}{I_t/\mathcal{K}_t}$ et $\frac{\Delta I_{t+1}}{I_t}=\frac{\mathcal{K}_t}{\mathcal{K}_{t+1}}d_t-1\to\theta^{-2}-1$ en loi et aussi en proba, par suit $\frac{h(X_t)}{I_t}=\frac{X_t^2}{I_t}=\frac{\Delta I_{t+1}}{I_t}\to\theta^{-2}-1\neq 0$. Pour le cas $|\theta|=1$ on a $d_t:=\frac{I_{t+1}/\mathcal{K}_{t+1}}{I_t/\mathcal{K}_t}$ et $\frac{\Delta I_{t+1}}{I_t}=\frac{\mathcal{K}_t}{\mathcal{K}_{t+1}}d_t-1\to 0$ en loi.

Chapitre 2

Convergence

2.1 Introduction

Soit $X_1, ..., X_n$ des observations iid de loi dépendant d'un paramètre réel inconnu θ. Rappellons du Ch 1 qu'un M-estimateur $\hat{\theta}_n = \hat{\theta}_n (X_1, ..., X_n)$ de θ est défini comme solution de l'équation

$$\sum_{i=1}^{n} \psi (X_i, \theta) = 0 \tag{2.1.1}$$

Supposons maintenant que les v.a ne sont pas iid, le M-estimateur de θ est défini comme solution de l'équation

$$\sum_{i=1}^{n} \psi_i (\theta) = 0 \tag{2.1.2}$$

où $\psi_i (\theta) = \psi_i \left(X_{i-k}^{i}; \theta \right)$ avec $X_{i-k}^{i} = (X_{i-k}, ..., X_i)$.

Si les fonctions ψ_i sont linéaires et pour la moyenne empirique on a

$$\begin{aligned} \hat{\theta}_n &= \bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i = \frac{1}{n} \left(\frac{n-1}{n-1} \sum_{i=1}^{n-1} X_i + X_n \right) \\ &= \frac{n-1}{n} \left(\frac{1}{n-1} \sum_{i=1}^{n-1} X_i \right) + \frac{1}{n} X_n = \frac{n-1}{n} \hat{\theta}_{n-1} + \frac{1}{n} X_n \end{aligned}$$

En générale, Si on pose $\sum_{i=1}^{n} \psi_i (\theta) = M_n (\theta)$, on a $M_n(\hat{\theta}_n) = 0$ et $M_{n-1}(\hat{\theta}_{n-1}) = 0$ et un développement de Taylor d'ordre 1 de $M_n(\hat{\theta}_n)$ au point $\hat{\theta}_{n-1}$ donne

$$\begin{aligned} 0 &= M_n(\hat{\theta}_n) - M_{n-1}(\hat{\theta}_{n-1}) \\ &= M_n(\hat{\theta}_{n-1}) + \acute{M}_n(\hat{\theta}_{n-1})(\hat{\theta}_n - \hat{\theta}_{n-1}) + o(\hat{\theta}_{n-1}) - M_{n-1}(\hat{\theta}_{n-1}) \end{aligned}$$

et on obtient sous certaines conditions

$$\hat{\theta}_n \approx \hat{\theta}_{n-1} - \frac{\psi_n(\hat{\theta}_{n-1})}{\acute{M}_n(\hat{\theta}_{n-1})}$$

Dans le cas iid avec $\psi(x,\theta) = \frac{\acute{f}(x,\theta)}{f(x,\theta)}$ (emv) et par la loi forte des grands nombres on a

$$\begin{aligned} \frac{\acute{M}(\theta)}{n} &= \frac{1}{n}\sum_{i=1}^{n} \acute{\psi}_i(\theta) = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{\acute{f}(X_1,\theta)}{f(X_1,\theta)}\right)' \approx E\left(\acute{\psi}_1(\theta)\right) \\ &= E\left(\frac{\acute{f}(X_1,\theta)}{f(X_1,\theta)}\right)' = -i(\theta) \end{aligned}$$

Donc

$$\hat{\theta}_n = \hat{\theta}_{n-1} + \frac{1}{ni(\hat{\theta}_{n-1})}\frac{\acute{f}(X_n,\hat{\theta}_{n-1})}{f(X_n,\hat{\theta}_{n-1})}, \quad n \geq 1 \tag{2.1.3}$$

Considérons la classe d'estimateurs suivante

$$\hat{\theta}_n = \hat{\theta}_{n-1} + \Gamma_n^{-1}(\hat{\theta}_{n-1})\psi_n(\hat{\theta}_{n-1}), \quad n \geq 1 \tag{2.1.4}$$

où ψ_n est une fonction mesurable et $\Gamma_n(\theta)$ une matrice $m \times m$ prévisible,pour tout $\theta \in \mathbb{R}^m$, telle que $\det \Gamma_n(\theta) \neq 0, n \geq 1$ et $\hat{\theta}_0 \in \mathbb{R}^m$ est un point initial arbitraire.

Dans cette partie nous donnons des conditions suffisantes (différentes de celles du Ch.1) de convergence de cette classe d'estimateurs.

2.2 Convergence des estimateurs

Nous avons le théorème suivant

Théorème 7 *Soit $\theta \in \mathbb{R}^m$ et $V_\theta(u) : \mathbb{R}^m \longrightarrow \mathbb{R}$ une fonction réelle non-négative continue de dérivées partielles d'ordre 2 bornées telle que*

$(G1)$ $V_\theta\left(0\right)=0$, *et pour tout* $\varepsilon\in\left]0,1\right[$: $\inf\limits_{\|u\|\geq\varepsilon}V_\theta\left(u\right)>0$

$(G2)$ *il existe un ensemble* $A\in\mathcal{F}$,*tel que* $P^\theta(A)>0$, *et pour tout* $\varepsilon\in\left]0,1\right[$,

$$\sum_{t=1}^{\infty}\inf_{\varepsilon\leq V_\theta(u)\leq 1/\varepsilon}\left[\mathcal{N}_t\left(u\right)\right]^-=\infty$$

dans A, *où*

$$\begin{aligned}\mathcal{N}_t(u) &= \dot{V}_\theta(u)\Gamma_t^{-1}(\theta+u)b_t(\theta+u)\\ &+\frac{1}{2}\sup_{v}\left\|\ddot{V}_\theta(v)\right\|E_\theta\left(\left\|\Gamma_t^{-1}(\theta+u)\psi_t(\theta+u)\right\|^2/\mathcal{F}_{t-1}\right)\end{aligned}$$

$(G3)$ *pour* $\Delta_t=\hat{\theta}_t-\theta$,

$$\sum_{t=1}^{\infty}\left(1+V_\theta(\Delta_{t-1})\right)^{-1}\left[\mathcal{N}_t(\Delta_{t-1})\right]^+<\infty,\quad P^\theta p.s$$

Alors $\hat{\theta}_t\longrightarrow\theta$ P^θ*p.s.pour tout* $\hat{\theta}_0$ *valeur initiale.*

Preuve: Comme $\hat{\theta}_t$ est defini par (2.1.4) et $\Delta_t=\hat{\theta}_t-\theta$ alors

$$\begin{aligned}\Delta_t &= \hat{\theta}_t-\theta=\hat{\theta}_{t-1}+\Gamma_t^{-1}\left(\theta+\Delta_{t-1}\right)\psi_t\left(\theta+\Delta_{t-1}\right)-\theta\\ &= \Delta_{t-1}+\Gamma_t^{-1}\left(\theta+\Delta_{t-1}\right)\psi_t\left(\theta+\Delta_{t-1}\right)\end{aligned}$$

Un développement de Taylor de $V_\theta(\Delta_t)$ à l'ordre 2 au point Δ_{t-1} donne

$$V_\theta(\Delta_t)=V_\theta(\Delta_{t-1})+\dot{V}_\theta(\Delta_{t-1})(\Delta_t-\Delta_{t-1})+\frac{1}{2}\ddot{V}_\theta(\tilde{\Delta}_t)(\Delta_t-\Delta_{t-1})^2$$

où $\tilde{\Delta}_t\in\mathbb{R}^m$. Donc en posant $M_t=\Delta_t-\Delta_{t-1}$ et $(\Delta_t-\Delta_{t-1})^2=M^TM$ on obtient

$$\begin{aligned}V_\theta(\Delta_t) &= V_\theta(\Delta_{t-1})+\dot{V}_\theta(\Delta_{t-1})\Gamma_t^{-1}\left(\theta+\Delta_{t-1}\right)\psi_t\left(\theta+\Delta_{t-1}\right)\\ &+\frac{1}{2}\left[\Gamma_t^{-1}\left(\theta+\Delta_{t-1}\right)\psi_t\left(\theta+\Delta_{t-1}\right)\right]^T\ddot{V}_\theta(\tilde{\Delta}_t)\Gamma_t^{-1}\left(\theta+\Delta_{t-1}\right)\psi_t\left(\theta+\Delta_{t-1}\right)\end{aligned}$$

L'espérance conditionnelle par rapport à $\mathcal{F}_{t-1}$ donne

$$\begin{aligned}
E_\theta\left(V_\theta(\Delta_t)/\mathcal{F}_{t-1}\right) &= V_\theta(\Delta_{t-1}) + \dot{V}_\theta(\Delta_{t-1})\Gamma_t^{-1}\left(\theta+\Delta_{t-1}\right) E_\theta\left(\psi_t\left(\theta+\Delta_{t-1}\right)/\mathcal{F}_{t-1}\right) \\
&\quad + \frac{1}{2}\ddot{V}_\theta(\tilde{\Delta}_t) E_\theta\left(\left\|\Gamma_t^{-1}\left(\theta+\Delta_{t-1}\right)\psi_t\left(\theta+\Delta_{t-1}\right)\right\|^2/\mathcal{F}_{t-1}\right) \\
&\leq V_\theta(\Delta_{t-1}) + \dot{V}_\theta(\Delta_{t-1})\Gamma_t^{-1}(\theta+\Delta_{t-1}) b_t(\theta+\Delta_{t-1}) \\
&\quad + \frac{1}{2}\sup_{v}\left\|\ddot{V}_\theta(v)\right\| E_\theta\left(\Gamma_t^{-1}(\theta+\Delta_{t-1})\psi_t(\theta+\Delta_{t-1})^2/\mathcal{F}_{t-1}\right) \\
&=: V_\theta(\Delta_{t-1}) + \mathcal{N}_t(\Delta_{t-1})
\end{aligned}$$

où $\sup\limits_{v}\left\|\ddot{V}_\theta(v)\right\|$ existe et $V_\theta(\Delta_{t-1})$,$\Gamma_t^{-1}(\theta+\Delta_{t-1})$ sont $\mathcal{F}_{t-1}$-mesurables. De la décomposition $\mathcal{N}_t(\Delta_{t-1}) = [\mathcal{N}_t(\Delta_{t-1})]^+ - [\mathcal{N}_t(\Delta_{t-1})]^-$ et l'inégalité précédente on deduit

$$\begin{aligned}
E_\theta\left(V_\theta(\Delta_t)/\mathcal{F}_{t-1}\right) &\leq V_\theta(\Delta_{t-1}) + \mathcal{N}_t(\Delta_{t-1}) \\
&\leq V_\theta(\Delta_{t-1}) + [\mathcal{N}_t(\Delta_{t-1})]^+ - [\mathcal{N}_t(\Delta_{t-1})]^- \\
&\leq V_\theta(\Delta_{t-1}) + \left(1+V_\theta(\Delta_{t-1})\right)\left(1+V_\theta(\Delta_{t-1})\right)^{-1}[\mathcal{N}_t(\Delta_{t-1})]^+ \\
&\quad - [\mathcal{N}_t(\Delta_{t-1})]^- \\
&\leq V_\theta(\Delta_{t-1}) + \left(1+V_\theta(\Delta_{t-1})\right) B_t - [\mathcal{N}_t(\Delta_{t-1})]^- \\
&= V_\theta(\Delta_{t-1})(1+B_t) + B_t - [\mathcal{N}_t(\Delta_{t-1})]^- \qquad (2.2.1)
\end{aligned}$$

où $B_t := \left(1+V_\theta(\Delta_{t-1})\right)^{-1}[\mathcal{N}_t(\Delta_{t-1})]^+$. Par la condition $(G3)$,$\sum\limits_{t=1}^{\infty} B_t < \infty$. Nous appliquons le Lemme 2 dans l'annexe 1 avec $X_n = V_\theta(\Delta_n), \beta_{n-1} = B_n = \xi_{n-1}, \zeta_{n-1} = [\mathcal{N}_n(\Delta_{n-1})]^-$, et $V_\theta(\Delta_n), B_{n+1}, [\mathcal{N}_n(\Delta_n)]^- \in \mathcal{F}_n$,$\forall n \geq 0$, donc

$$\left\{\sum_{s=1}^{\infty}\xi_{s-1} = \sum_{s=1}^{\infty}\beta_{s-1} = \sum_{s=1}^{\infty} B_s < \infty\right\} \subseteq \{V_\theta(\Delta_t) \longrightarrow V\} \cap \left\{\sum_{s=1}^{\infty}[\mathcal{N}_s(\Delta_{s-1})]^-\right\} < \infty$$

c'est à dire

$$\left\{\sum_{s=1}^{\infty} B_s < \infty\right\} \subseteq \{V_\theta(\Delta_t) \longrightarrow V\} \cap \left\{\sum_{s=1}^{\infty}[\mathcal{N}_s(\Delta_{s-1})]^- < \infty\right\}$$

et d'aprés $(G3)$, nous avons

$$P^\theta\left(\left\{\sum_{s=1}^{\infty} B_s < \infty\right\}\right) = 1 \leq P^\theta\left(\{V_\theta(\Delta_t) \longrightarrow V\} \cap \left\{\sum_{s=1}^{\infty}[\mathcal{N}_s(\Delta_{s-1})]^- < \infty\right\}\right) = 1$$

ou encore

$$P^{\theta}\left(\{V_{\theta}(\Delta_t)\longrightarrow V\}\right)=1 \text{ et } P^{\theta}\left(\left\{\sum_{s=1}^{\infty}[N_s(\Delta_{s-1})]^{-}<\infty\right\}\right)=1$$

par suit les processus $V_{\theta}(\Delta_t)$ et $Y_t=\sum_{s=1}^{t}[\mathcal{N}_s(\Delta_{s-1})]^{-}$ convergent $p.s$ vers des limites finies Comme $V_{\theta}(\Delta_t)\geq 0$ on a $V\geq 0$. Supposons que $V>0$ $p.s$, alors il existe un $\varepsilon>0$ et t_0 tel que tout $t>t_0$: $\varepsilon\leq V_{\theta}(\Delta_t)\leq 1/\varepsilon$. Donc $s>t_0$: $[\mathcal{N}_s(\Delta_{s-1})]^{-}\geq \inf\limits_{\varepsilon\leq V_{\theta}(u)\leq 1/\varepsilon}[\mathcal{N}_s(u)]^{-}$. Par $(G2)$ on obtient

$$\sum_{s=t_0}^{\infty}[\mathcal{N}_s(\Delta_{s-1})]^{-}\geq\sum_{s=t_0}^{\infty}\inf_{\varepsilon\leq V_{\theta}(u)\leq 1/\varepsilon}[\mathcal{N}_s(u)]^{-}=\infty$$

dans l'ensemble A avec $P(A)>0$. Ce qui contredit la convergence presque sure de $Y_t=\sum_{s=1}^{\infty}[N_s(\Delta_{s-1})]^{-}$. D'où $V=0$ et $\lim\limits_{t\longrightarrow\infty}V_{\theta}(\Delta_t)=0$ $p.s$. Comme V_{θ} est continue et par $(G1)$ on a $V(0)=0=\lim\limits_{t\longrightarrow\infty}V_{\theta}(\Delta_t)=V_{\theta}\left(\lim\limits_{t\longrightarrow\infty}\Delta_t\right)$ ce qui donne $\lim\limits_{t\longrightarrow\infty}\Delta_t=0$ $P^{\theta}p.s$ où $\lim\limits_{t\longrightarrow\infty}\hat{\theta}_t=\theta$ $P^{\theta}p.s.$ ■

Corollaire [2.1] Supposons que

$(C1)$

$$u^{T}\Gamma_t^{-1}(\theta+u)\,b_t(\theta,u)<0 \text{ si } u\neq 0 \quad P^{\theta}p.s$$

$(C2)$ pour tout $\varepsilon\in\,]0,1[$,

$$\sum_{t=1}^{\infty}\inf_{\varepsilon\leq\|u\|\leq\frac{1}{\varepsilon}}\left|u^{T}\Gamma_t^{-1}(\theta+u)\,b_t(\theta,u)\right|=\infty$$

$(C3)$ $\forall u\in\mathbb{R}^m$, il existe un processus scalaire prévisible $(B_t^{\theta})_{t\geq 1}$, tel que

$$E_{\theta}\left(\left\|\Gamma_t^{-1}(\theta+u)\psi_t(\theta+u)\right\|^2/\mathcal{F}_{t-1}\right)\leq B_t^{\theta}\left(1+\|u\|^2\right)$$

et

$$\sum_{t=1}^{\infty}B_t^{\theta}<\infty,\quad P^{\theta}p.s$$

Alors $\hat{\theta}_t\longrightarrow\theta$ $P^{\theta}p.s$ pour tout $\hat{\theta}_0$ valeur initiale.

Preuve: Nous allons vérifier les conditions de Théorème 7. On prend pour $u \in \mathbb{R}^m, \theta \in \mathbb{R}^m$ $V_\theta(u) = \langle u, u \rangle = u^T u = \|u\|^2$ qui est non-négative continue et 2 fois dérivable avec $\dot{V}_\theta(u) = 2u^T$ et $\ddot{V}_\theta(u) = 2 \times \mathbf{1}$, et $\sup_u \left\| \ddot{V}_\theta\left(u\right) \right\| = 2, V_\theta(0) = 0, \forall \varepsilon \in]0,1[: \inf_{\|u\| \geq \varepsilon} V_\theta(u) = \inf_{\|u\| \geq \varepsilon} \|u\|^2 = \varepsilon^2 > 0$. Donc $(G1)$ est satisfaite.

Pour vérifier $(G3)$, nous avons

$$\begin{aligned}
\mathcal{N}_t(u) &= \dot{V}_\theta(u)\Gamma_t^{-1}(\theta+u)b_t(\theta,u) \\
&\quad + \frac{1}{2}\sup_v \left\| \ddot{V}_\theta\left(v\right) \right\| E_\theta\left(\left\| \Gamma_t^{-1}(\theta+u)\psi_t(\theta+u) \right\|^2 / \mathcal{F}_{t-1} \right) \\
&= 2u^T\Gamma_t^{-1}(\theta+u)b_t(\theta,u) + \frac{1}{2} 2E_\theta\left(\left\| \Gamma_t^{-1}(\theta+u)\psi_t(\theta+u) \right\|^2 / \mathcal{F}_{t-1} \right) \\
&= 2u^T\Gamma_t^{-1}(\theta+u)b_t(\theta,u) + E_\theta\left(\left\| \Gamma_t^{-1}(\theta+u)\psi_t(\theta+u) \right\|^2 / \mathcal{F}_{t-1} \right) \qquad (2.2.2)
\end{aligned}$$

Par $(C1)$ on a

$$\begin{aligned}
[\mathcal{N}_t(u)]^+ &= \left[2u^T\Gamma_t^{-1}(\theta+u)b_t(\theta,u) + E_\theta\left(\left\| \Gamma_t^{-1}(\theta+u)\psi_t(\theta+u) \right\|^2 / \mathcal{F}_{t-1} \right) \right]^+ \\
&\leq E_\theta\left(\left\| \Gamma_t^{-1}(\theta+u)\psi_t(\theta+u) \right\|^2 / \mathcal{F}_{t-1} \right)
\end{aligned}$$

par $(C3)$ on a

$$\begin{aligned}
\sum_{t=1}^{\infty} (1+V_\theta(\Delta_{t-1}))^{-1} [\mathcal{N}_t(\Delta_{t-1})]^+ &= \sum_{t=1}^{\infty} \left(1+\|\Delta_{t-1}\|^2\right)^{-1} [\mathcal{N}_t(\Delta_{t-1})]^+ \\
&\leq \sum_{t=1}^{\infty} (1+\|\Delta_{t-1}\|^2)^{-1} B_t (1+\|\Delta_{t-1}\|^2) \\
&= \sum_{t=1}^{\infty} B_t < \infty \qquad (2.2.3)
\end{aligned}$$

Donc $(G3)$ est satisfaite.

Pour verifier $(G2)$, de $[a]^- \geq -a$ et $(C1)$, nous avons

$$\begin{aligned}
\inf [\mathcal{N}_t(u)]^- &\geq \inf \left[-2u^T\Gamma_t^{-1}(\theta+u)b_t(\theta,u) - E_\theta\left(\left\| \Gamma_t^{-1}(\theta+u)\psi_t(\theta+u) \right\|^2 / \mathcal{F}_{t-1} \right) \right] \\
&\geq \inf \left| 2u^T\Gamma_t^{-1}(\theta+u)b_t(\theta,u) \right| - \sup E_\theta\left(\left\| \Gamma_t^{-1}(\theta+u)\psi_t(\theta+u) \right\|^2 / \mathcal{F}_{t-1} \right)
\end{aligned}$$

où inf et sup sont trouvée sur $\left\{u : \varepsilon \leq \|u\|^2 \leq 1/\varepsilon\right\}$. Par $(C3)$ nous avons

$$\sup\left[E_\theta\left(\left\|\Gamma_t^{-1}(\theta+u)\psi_t(\theta+u)\right\|^2/\mathcal{F}_{t-1}\right)\right] \leq B_t^\theta(1+\|u\|^2) \leq B_t^\theta(1+1/\varepsilon^2)$$

où $\sum\limits_{t=1}^{\infty} B_t < \infty$. Maintenant nous utilisons $(C2)$ pour obtenir

$$\sum_{t=1}^{\infty} \inf\left[\mathcal{N}_t(u)\right]^- \geq \sum_{t=1}^{\infty} \inf\left|2u^T\Gamma_t^{-1}(\theta+u)b_t(\theta,u)\right| - (1+1/\varepsilon^2)\sum_{t=1}^{\infty} B_t = \infty$$

D'où $(G2)$. Finalement les conditions de Théorème 7 sont satisfaites et par suit $\hat{\theta}_t \longrightarrow \theta \;\; P^\theta p.s$ pour tout $\hat{\theta}_0$. ■

Remarque 8 *De la preuve de Théorème 7 on déduit que si les conditions* $(G1)$ *et* $(G3)$ *sont satisfaites, alors* $(\hat{\theta}_t - \theta)^2$ *converge* $P^\theta p.s$ *vers une limite finie pour tout valeur initiale* $\hat{\theta}_0$ *(cf.Ch 1 lemme 1). En particulièr, pour garantir cet convergence il est suffissant de nécessiter les conditions* $(C1)$ *et* $(C3)$ *de corollaire [2.1] (si nous prenons* $V_\theta(u) = \langle u, u\rangle = u^T u = \|u\|^2$ *et* (2.2.3)*).*

2.3 Exemples

1. Cas iid

Considérons le cas des observations $X_1, X_2, ...$ iid de densité de probabilité $f(x,\theta)$, $\theta \in \mathbb{R}^m$. Supposons que $\psi(\theta, z)$ est une fonction d'estimation vérifiant $\int \psi(\theta,z) f(z,\theta)\mu(dz) = 0$ et nous définissons l'estimateur récursif $\hat{\theta}_t$ par

$$\hat{\theta}_t = \hat{\theta}_{t-1} + \frac{1}{t}\gamma^{-1}(\hat{\theta}_{t-1})\psi(\hat{\theta}_{t-1}, X_t), \qquad t \geq 1 \tag{2.3.1}$$

où $\gamma(\theta)$ est une matrice non-aléatoire, telle que $\gamma^{-1}(\theta)$ existe pour tout $\theta \in \mathbb{R}^m$, et $\hat{\theta}_0 \in \mathbb{R}^m$ est une valeur initiale.

Nous avons le résultat suivant

Corollaire [2.2] Supposons que pour $\theta \in \mathbb{R}^m$, les conditions suivantes sont satisfaites:

(I) pour tout $0 < \varepsilon < 1$,

$$\sup_{\varepsilon \leq \|u\| \leq 1/\varepsilon} u^T \gamma^{-1}(\theta + u) \int \psi(\theta + u, x) f(\theta, x) \mu(dx) < 0$$

(II) pour tout $u \in \mathbb{R}^m, \exists K_\theta > 0$,

$$\int \left\| \gamma^{-1}(\theta + u) \psi(\theta + u, x) \right\|^2 f(\theta, x) \mu(dx) \leq K_\theta (1 + \|u\|^2)$$

Alors l'estimateur $\hat{\theta}_t$ converge $p.s$ pour tout valeur initiale $\hat{\theta}_0$.

Preuve: On pose $\Gamma_t(\theta) = t\gamma(\theta)$, et $b_t(\theta, u) = b(\theta, u) = \int \psi(\theta + u, x) f(\theta, x) \mu(dx)$. Nous avons *(I)* et *(II)* impliquent $(C1), (C2), (C3)$ de corollaire [2.1]. En effet, la condition (I) implique pour tout $t \geq 1$,

$$\begin{aligned}
& \sup_{\varepsilon \leq \|u\| \leq 1/\varepsilon} u^T \gamma^{-1}(\theta + u) \int \psi(\theta + u, x) f(\theta, x) \mu(dx) \\
= & \left(\sup_{\varepsilon \leq \|u\| \leq 1/\varepsilon} u^T \gamma^{-1}(\theta + u) b_t(\theta, u) \right) \times \frac{1}{t} \\
= & \sup_{\varepsilon \leq \|u\| \leq 1/\varepsilon} u^T \frac{1}{t} \gamma^{-1}(\theta + u) b_t(\theta, u) \\
\leq & \sup_{t \geq 1} \sup_{\varepsilon \leq \|u\| \leq 1/\varepsilon} u^T \Gamma_t^{-1}(\theta + u) b_t(\theta, u) < 0
\end{aligned}$$

Comme ε est quelconque dans $]0, 1[$ ceci implique que $\sup_{t \geq 1} u^T \Gamma_t^{-1}(\theta + u) b_t(\theta, u) < 0$ si $u \neq 0$ qui est $(C1)$. Pour $(C2)$ la condition (I) implique (voir ci dessus) que pour tout $0 < \varepsilon < 1$ on a $\inf_{\varepsilon \leq \|u\| \leq \ /\varepsilon} \left| u^T \Gamma_t^{-1}(\theta + u) b_t(\theta, u) \right| > \delta > 0$, par suit

$$\sum_{t=1}^{\infty} \inf_{\varepsilon \leq \|u\| \leq 1/\varepsilon} \left| u^T \Gamma_t^{-1}(\theta + u) b_t(\theta, u) \right| > \delta \sum_{t=1}^{\infty} 1 = \infty. \quad P^\theta p.s$$

qui est $(C2)$. Pour $(C3)$ la condition (II) implique

$$\begin{aligned}
& \frac{1}{t^2}\int \left\|\gamma^{-1}(\theta+u)\,\psi(\theta+u,x)\right\|^2 f(\theta,x)\mu(dx) \\
= & \int \left\|\Gamma_t^{-1}(\theta+u)\,\psi(\theta+u,x)\right\|^2 f(\theta,x)\mu(dx) \\
\leq & \frac{1}{t^2}K_\theta(1+\|u\|^2)
\end{aligned}$$

ou encore

$$E_\theta\left(\left\|\Gamma_t^{-1}(\theta+u)\,\psi(\theta+u,x)\right\|^2/\mathcal{F}_{t-1}\right)\leq B_t^\theta(1+\|u\|^2),\quad \forall u\in\mathbb{R}^m$$

où $\sum\limits_{t=1}^{\infty} B_t^\theta = \sum\limits_{t=1}^{\infty}\frac{K_\theta}{t^2}<\infty$. D'où la condition $(C3)$. Finalement du corollaire [2.1] on deduit $\hat{\theta}_t \longrightarrow \theta\ P^\theta p.s.$ ■

Il existent des résultats similaires (cas iid) dans Khas'minskii et Nevelson (1972,[6],Ch 8,§4) et Fabian (1978,[3]) avec d'autres conditions alternatives. Nous notons que dans (2.3.1), la suite normalisée est $\Gamma_t(\theta)=t\gamma(\theta)$, mais les corollaires [2.1] et [2.2] permettent de considérer des procedures avec un processus arbitraire prévisible $\Gamma_t(\theta)$.

2. Procedures linéaires

Considérons les estimateurs récursifs definis par

$$\hat{\theta}_t=\hat{\theta}_{t-1}+\Gamma_t^{-1}\left(h_t-\gamma_t\hat{\theta}_{t-1}\right),\qquad t\geq 1 \tag{2.3.2}$$

où Γ_t et γ_t sont des processus prévisibles, h_t est un processus adapté à $\mathcal{F}_t,\forall t\geq 1$ et ne dépendant pas de θ. Les résultats suivants donnent des conditions suffissantes de convergence de (2.3.2) dans le cas linéaire où $\psi_t(\theta)=h_t-\gamma_t\theta$ est une différence de martingales où $\theta\in\mathbb{R}$.

Corollaire [2.3] Supposons que pour tout $\theta\in\mathbb{R}$,

(a) $E_\theta(h_t/\mathcal{F}_{t-1})=\gamma_t\theta$ pour $t\geq 1$, $P^\theta p.s.$

(b) $0\leq\gamma_t/\Gamma_t\leq 2-\delta$ pour quelque $\delta>0$, et $\sum\limits_{t=1}^{\infty}\gamma_t/\Gamma_t=\infty$ dans un ensemble A de probabilité P^θ positive,

(c)

$$\sum_{t=1}^{\infty}\frac{E_\theta\left((h_t-\theta\gamma_t)^2/\mathcal{F}_{t-1}\right)}{\Gamma_t^2}<\infty,\qquad P^\theta p.s$$

Alors $\hat{\theta}_t \longrightarrow \theta \;\; P^\theta p.s$ pour tout valeur $\hat{\theta}_0$.

Preuve: Nous vérifions les conditions de Théorème 7 pour $V_\theta(u) = u^2$. La condition (a) implique

$$\begin{aligned} b_t(\theta,u) \;\; &: = E_\theta\left(\psi_t\left(\theta+u\right)/\mathcal{F}_{t-1}\right) = E_\theta\left(\left(h_t - \left(\theta+u\right)\gamma_t\right)/\mathcal{F}_{t-1}\right) \\ &= E_\theta\left(h_t/\mathcal{F}_{t-1}\right) - \theta\gamma_t - u\gamma_t = -u\gamma_t \end{aligned}$$

Par suit

$$\begin{aligned} E_\theta\left(\psi_t^2(\theta+u)/\mathcal{F}_{t-1}\right) &= E_\theta\left(\left(h_t-(\theta+u)\gamma_t\right)^2/\mathcal{F}_{t-1}\right) \\ &= E_\theta\left((h_t-\theta\gamma_t)^2 + u^2\gamma_t^2 - 2u\gamma_t(h_t-\theta\gamma_t)/\mathcal{F}_{t-1}\right) \\ &= E_\theta\left((h_t-\theta\gamma_t)^2/\mathcal{F}_{t-1}\right) + u^2\gamma_t^2 \;\; (\gamma_t \text{ est } \mathcal{F}_{t-1}\text{-mes.}) \\ &= :\mathcal{P}_t^\theta + u^2\gamma_t^2 \end{aligned}$$

Donc de (2.2.2) nous avons

$$\begin{aligned} \mathcal{N}_t(u) \;\; &: = 2u^T\Gamma_t^{-1}(\theta+u)b_t(\theta,u) + E_\theta\left(\Gamma_t^{-2}(\theta+u)\psi_t^2(\theta+u)/\mathcal{F}_{t-1}\right) \\ &= 2u^T\Gamma_t^{-1}(\theta+u)(-u\gamma_t) + \Gamma_t^{-2}(\theta+u)\left(\mathcal{P}_t^\theta + u^2\gamma_t^2\right) \;\; (\Gamma_t^{-1} \text{ est } \mathcal{F}_{t-1}\text{-mes}) \\ &= -2u^2\gamma_t\Gamma_t^{-1}(\theta+u) + \Gamma_t^{-2}(\theta+u)\mathcal{P}_t^\theta + u^2\gamma_t^2\Gamma_t^{-2}(\theta+u) \\ &= -\delta u^2\gamma_t\Gamma_t^{-1}(\theta+u) + \delta u^2\gamma_t\Gamma_t^{-1}(\theta+u) - 2u^2\gamma_t\Gamma_t^{-1}(\theta+u) \\ &\quad +\Gamma_t^{-2}(\theta+u)\mathcal{P}_t^\theta + u^2\gamma_t^2\Gamma_t^{-2}(\theta+u) \\ &= -\delta u^2\gamma_t\Gamma_t^{-1}(\theta+u) - (2-\delta)u^2\gamma_t\Gamma_t^{-1}(\theta+u) + \Gamma_t^{-2}\mathcal{P}_t^\theta + u^2\gamma_t^2\Gamma_t^{-2}(\theta+u) \\ &= -\delta u^2\gamma_t\Gamma_t^{-1}(\theta+u) - u^2\gamma_t\Gamma_t^{-1}(\theta+u)\left((2-\delta) - \gamma_t\Gamma_t^{-1}(\theta+u)\right) \\ &\quad +\Gamma_t^{-2}(\theta+u)\mathcal{P}_t^\theta \end{aligned}$$

Par la condition (b) et $[x]^- \geq -x$, nous avons

$$\begin{aligned}
[\mathcal{N}_t(u)]^- &\geq \delta u^2 \gamma_t \Gamma_t^{-1}(\theta+u) + u^2 \gamma_t \Gamma_t^{-1}(\theta+u) \left((2-\delta) - \gamma_t \Gamma_t^{-1}(\theta+u)\right) \\
&\quad - \Gamma_t^{-2}(\theta+u) \mathcal{P}_t^\theta \\
&\geq \delta u^2 \gamma_t \Gamma_t^{-1}(\theta+u) - \Gamma_t^{-2}(\theta+u) \mathcal{P}_t^\theta
\end{aligned}$$

et donc

$$\begin{aligned}
\inf_{\varepsilon \leq u^2 \leq 1/\varepsilon} [\mathcal{N}_t(u)]^- &\geq \inf_{\varepsilon \leq u^2 \leq 1/\varepsilon} \left[\delta u^2 \gamma_t \Gamma_t^{-1}(\theta+u) - \Gamma_t^{-2}(\theta+u) \mathcal{P}_t^\theta\right] \\
&= \inf_{\varepsilon \leq u^2 \leq 1/\varepsilon} \delta u^2 \gamma_t \Gamma_t^{-1}(\theta+u) - \Gamma_t^{-2}(\theta+u) \mathcal{P}_t^\theta \\
&\geq \delta \varepsilon \gamma_t \Gamma_t^{-1}(\theta+u) - \Gamma_t^{-2}(\theta+u) \mathcal{P}_t^\theta
\end{aligned}$$

Donc dans A, et par (b), (c) nous avons

$$\sum_{t=1}^{\infty} \inf_{\varepsilon \leq u^2 \leq 1/\varepsilon} [\mathcal{N}_t(u)]^- \geq \delta \varepsilon \sum_{t=1}^{\infty} \gamma_t \Gamma_t^{-1}(\theta+u) - \sum_{t=1}^{\infty} \Gamma_t^{-2}(\theta+u) \mathcal{P}_t^\theta = \infty$$

et donc $(G2)$ est satisfaite. Pour la condition $(G3)$, de la relation

$$\begin{aligned}
\mathcal{N}_t(u) &= -\delta u^2 \gamma_t \Gamma_t^{-1}(\theta+u) - u^2 \gamma_t \Gamma_t^{-1}(\theta+u) \left((2-\delta) - \gamma_t \Gamma_t^{-1}(\theta+u)\right) \\
&\quad + \Gamma_t^{-2}(\theta+u) \mathcal{P}_t^\theta
\end{aligned}$$

où les deux premièrs termes sont négatifs on a $[\mathcal{N}_t(u)]^+ \leq \Gamma_t^{-2}(\theta+u)\mathcal{P}_t^\theta$, et par (c), nous avons

$$\sum_{t=1}^{\infty} (1+\Delta_{t-1}^2)^{-1} [\mathcal{N}_t(u)]^+ \leq \sum_{t=1}^{\infty} [\mathcal{N}_t(u)]^+ \leq \sum_{t=1}^{\infty} \Gamma_t^{-2}(\theta+u) \mathcal{P}_t^\theta < \infty, \qquad P^\theta p.s$$

D'où $(G3)$. Les conditions de Théorème 7 sont satisfaites et donc $\hat{\theta}_t \longrightarrow \theta$ $P^\theta p.s.$ ∎

Remarque 9 *Si on prend* $\Delta\Gamma_t = \Gamma_t - \Gamma_{t-1} = \gamma_t$ *alors*

$$\hat{\theta}_t = \Gamma_t^{-1} \left(\hat{\theta}_0 + \sum_{s=1}^{t} h_s(X_s) \right) \tag{2.3.3}$$

On peut vérifier que la différence $\hat{\theta}_t - \hat{\theta}_{t-1}$ *de* (2.3.3) *permet d'obtenir* (2.3.2) :

$$\begin{aligned}
\hat{\theta}_t - \hat{\theta}_{t-1} &= \Gamma_t^{-1}\left(\hat{\theta}_0 + \sum_{s=1}^{t} h_s\right) - \Gamma_t^{-1}\Gamma_t\Gamma_{t-1}^{-1}\left(\hat{\theta}_0 + \sum_{s=1}^{t-1} h_s\right) \\
&= \Gamma_t^{-1}\left[\left(\hat{\theta}_0 + \sum_{s=1}^{t-1} h_s\right)\left(1 - \Gamma_t\Gamma_{t-1}^{-1}\right) + h_t\right] \\
&= \Gamma_t^{-1}\left[h_t - \left(\Gamma_t\Gamma_{t-1}^{-1} - 1\right)\left(\hat{\theta}_0 + \sum_{s=1}^{t-1} h_s\right)\right] \\
&= \Gamma_t^{-1}\left[h_t - \left(\Gamma_t - \Gamma_{t-1}\right)\Gamma_{t-1}^{-1}\left(\hat{\theta}_0 + \sum_{s=1}^{t-1} h_s\right)\right] \\
&= \Gamma_t^{-1}\left(h_t - \gamma_t\hat{\theta}_{t-1}\right)
\end{aligned}$$

Il est aussi intéressant d'observer que dans ce cas puisque $\Delta\Gamma_t = \gamma_t \Longleftrightarrow \Gamma_t = \sum\limits_{s=1}^{t} \gamma_s$, *et*

$$\begin{aligned}
\hat{\theta}_t &= \Gamma_t^{-1}\left(\hat{\theta}_0 + \sum_{s=1}^{t} h_s\left(X_s\right)\right) \\
&= \Gamma_t^{-1}\hat{\theta}_0 + \Gamma_t^{-1}\sum_{s=1}^{t} h_s\left(X_s\right) - \theta\left(\Gamma_t^{-1}\sum_{s=1}^{t}\gamma_s\right) + \theta \\
&= \Gamma_t^{-1}\hat{\theta}_0 + \Gamma_t^{-1}\sum_{s=1}^{t}\left(h_s\left(X_s\right) - \gamma_s\theta\right) + \theta \\
&= :\Gamma_t^{-1}\hat{\theta}_0 + \Gamma_t^{-1}M_t^{\theta} + \theta
\end{aligned}$$

où $M_t^{\theta} = \sum\limits_{s=1}^{t}\left(h_s\left(X_s\right) - \gamma_s\theta\right)$ *est* P^{θ} *martingale et* $h_s\left(X_s\right) - \gamma_s\theta$ *est une différence de martingale (Corollaire [2.3]). Si* $\Gamma_t \longrightarrow \infty$, *la condition nécessaire et suffissante pour la convergence vers* θ *est* $\Gamma_t^{-1}M_t^{\theta} \to 0$ *p.s. la condition* (c) *du Corollaire* 4.2 *est une condition standard pour avoir* $\Gamma_t^{-1}M_t^{\theta} \longrightarrow 0$ *(cf.[17],Ch.VII,§5 Théorème 4). La première partie de* (b) *sera obtenue si* $\gamma_t = \Delta\Gamma_t = \Gamma_t - \Gamma_{t-1} \geq 0$ *où* $\Gamma_t \geq \Gamma_{t-1} > 0$. *Donc* $\gamma_t/\Gamma_t = \Delta\Gamma_t/\Gamma_t = \frac{\Gamma_t - \Gamma_{t-1}}{\Gamma_t} = 1 - \frac{\Gamma_{t-1}}{\Gamma_t} > 0$ *et* $0 < \delta_1 \leq \frac{\Delta\Gamma_t}{\Gamma_t} \leq 2 - \delta, 0 < \delta < 1$. *D'autre part par le Lemme 3 d'annexe 1,* $\Gamma_t \nearrow$ *et* $\Gamma_t \to \infty$ *donc* $\sum\limits_{t=1}^{\infty}\frac{\gamma_t}{\Gamma_t} = \sum\limits_{t=1}^{\infty}\frac{\Delta\Gamma_t}{\Gamma_t} = \infty$.

Remarque 10 *si nous considérons le processus*

$$X_t = \theta X_{t-1} + Z_t, \qquad t \geq 1$$

où Z_t est une différence de martingale et $D_t := E_\theta\left(Z_t^2/\mathcal{F}_{t-1}\right) > 0$. Le choix $h_t = D_t^{-1}X_{t-1}X_t$ et $\Delta\Gamma_t = \gamma_t = D_t^{-1}X_{t-1}^2$ dans (2.3.2) donne l'estimateur des moindres carrés de θ: $\psi_t(\hat{\theta}_{t-1}) := h_t - \gamma_t\hat{\theta}_{t-1} = D_t^{-1}X_{t-1}X_t - D_t^{-1}X_{t-1}^2\hat{\theta}_{t-1}$, et l'equation d'estimation devient:

$$\sum_{s=1}^{t}\psi_s(\hat{\theta}_{s-1}) = \sum_{s=1}^{t}\left(D_t^{-1}X_{t-1}X_t - D_t^{-1}X_{t-1}^2\hat{\theta}_{t-1}\right) = 0$$

D'où

$$\hat{\theta}_t = (\sum_{s=1}^{t}D_s^{-1}X_{s-1}X_s)/\sum_{s=1}^{t}D_t^{-1}X_{t-1}^2$$

la condition (a) est satisfaite:

$$\begin{aligned} E_\theta\left(h_t/\mathcal{F}_{t-1}\right) &= E_\theta\left(D_t^{-1}X_{t-1}X_t/\mathcal{F}_{t-1}\right) = D_t^{-1}X_{t-1}E_\theta\left(X_t/\mathcal{F}_{t-1}\right) \\ &= D_t^{-1}X_{t-1}\left(\theta X_{t-1}\right) = \theta D_t^{-1}X_{t-1}^2 = \theta\gamma_t, \quad P^\theta p.s, \quad \forall t \geq 1 \end{aligned}$$

Pour la condition (c):

$$\begin{aligned} E_\theta\left[\left(h_t - \gamma_t\theta\right)^2/\mathcal{F}_{t-1}\right] &= E_\theta\left[\left(D_t^{-1}X_{t-1}X_t - \theta D_t^{-1}X_{t-1}^2\right)^2/\mathcal{F}_{t-1}\right] \\ &= E_\theta\left[\left(D_t^{-1}X_{t-1}Z_t\right)^2/\mathcal{F}_{t-1}\right] \\ &= D_t^{-2}X_{t-1}^2D_t = D_t^{-1}X_{t-1}^2 = \Delta\Gamma_t \end{aligned}$$

D'où (c) du Corollaire [2.3] est équivalente à $\sum\limits_{t=1}^{\infty}\frac{\Delta\Gamma_t}{\Gamma_t^2} < \infty$. Or (b) est satisfaite si $\Gamma_t \longrightarrow \infty$ (par le Lemme 3 dans l'annexe 1). Donc les estimateurs des moindres carrés sont convergents. Par exemple,si Z_t sont des v.a.r iid, donc $\Gamma_t = I_t = \sum\limits_{s=1}^{t}X_{s-1}^2 \longrightarrow \infty$ (cf.[17],Ch.VII,5.5), et $I_t - I_{t-1} = X_{t-1}^2 \geq 0$ alors par le Lemme 3 dans l'annexe 1 on a

$$\sum_{t=1}^{\infty}\frac{\Delta\Gamma_t}{\Gamma_t} = \infty, \;\; et \;\; \sum_{t=1}^{\infty}\frac{\Delta\Gamma_t}{\Gamma_t^2} < \infty, \quad P^\theta p.s$$

les conditions de Corollaire [2.3] sont ainsi satisfaites et $\hat{\theta}_t \longrightarrow \theta$, $\;P^\theta p.s.$

3. Processus AR(m)

Considérons un processus AR(m) defini par

$$X_i = \theta_1 X_{i-1} + ... + \theta_m X_{i-m} + Z_i = \theta^T X_{i-m}^{i-1} + Z_i,$$

où $X_{i-m}^{i-1} = (X_{i-1}, ..., X_{i-m})^T$, $\theta = (\theta_1, ..., \theta_m)^T$, et Z_i est une suite de v.a iid.

Nous definissons une suite d'estimateurs de la forme

$$\hat{\theta}_t = \hat{\theta}_{t-1} + \Gamma_t^{-1}(\hat{\theta}_{t-1})\psi_t(X_t - \hat{\theta}_{t-1} X_{t-m}^{t-1}) \tag{2.3.4}$$

où $\psi_t(z)$ et $\Gamma_t^{-1}(z); z \in \mathbb{R}^m$ sont un processus vectoriel et matriciel verifiants les conditions des sections précédentes. Supposons que la densité de Z_t est g alors la densité conditionelle est

$$f\left(\theta, x_t / x_1^{t-1}\right) = g\left(x_t - \theta^T x_{t-m}^{t-1}\right) = g \circ \left(x_t - \theta^T x_{t-m}^{t-1}\right)$$

et

$$\dot{f}^T\left(\theta, x_t / x_1^{t-1}\right) = \left(x_t - \theta^T x_{t-m}^{t-1}\right)' \left[g \circ \left(x_t - \theta^T x_{t-m}^{t-1}\right)\right]' = -x_{t-m}^{t-1} \acute{g}\left(x_t - \theta^T x_{t-m}^{t-1}\right)$$

Donc

$$\psi_t(z) = l_t(z) = \frac{\dot{f}^T(z)}{f(z)} = -\frac{\acute{g}(z)}{g(z)} X_{t-m}^{t-1} \tag{2.3.5}$$

et (2.3.4) donne l'estimateur de maximum du vraissemblance. Dans ce cas un choix possible de $\Gamma_t(z)$ est la matrice d'information conditionelle de Fisher: $\Gamma_t(z) = I_t(z)$ et

$$\begin{aligned}
I_t(z) &= \sum_{s=1}^{t} i_s(z) = \sum_{s=1}^{t} E_\theta\left(l_t(z) l_t^T(z) / \mathcal{F}_{t-1}\right) = \sum_{s=1}^{t} E_\theta\left(\psi_t(z)\psi_t^T(z) / \mathcal{F}_{t-1}\right) \\
&= \sum_{s=1}^{t} E_\theta\left[\left(-\frac{\acute{g}^T(z)}{g(z)} X_{t-m}^{t-1}\right)\left(-\frac{\acute{g}^T(z)}{g(z)} X_{t-m}^{t-1}\right)^T / \mathcal{F}_{t-1}\right] \\
&= \sum_{s=1}^{t} \int \left(-\frac{\acute{g}^T(z)}{g(z)}\right)^2 X_{t-m}^{t-1}\left(X_{t-m}^{t-1}\right)^T g(z) dz \\
&= \int \left(-\frac{\acute{g}^T(z)}{g(z)}\right)^2 g(z) dz \sum_{s=1}^{t} X_{t-m}^{t-1}\left(X_{t-m}^{t-1}\right)^T =: i^g \sum_{s=1}^{t} X_{t-m}^{t-1}\left(X_{t-m}^{t-1}\right)^T
\end{aligned}$$

Une classe d'estimateurs récursifs pour des processus AR(m) strictement stationnaire est dans

Campbell (1982,[2]). Ces estimateurs verifient

$$\hat{\theta}_t = \hat{\theta}_{t-1} + a_t \gamma(X_{t-m}^{t-1}) \phi(X_t - \hat{\theta}_{t-1}^T X_{t-m}^{t-1}), \tag{2.3.6}$$

où a_t est une suite de nombres positifs tel que $a_t \longrightarrow 0$, ϕ est une fonction scalaire limitée et $\gamma(u)$ est une fonction de vecteurs de la forme $uh(u)$ pour une fonction non-négative h de u (cf.[9]) La classe de type (2.3.6) est une sous classe de (2.3.4) et donc peut être étudiée avec les résultats des sections précédentes. Si la densité g est symétrique et ϕ impaire, continue au 0 nous reécrivons les conditions de corollaire [2.1] pour

$$\Gamma(\theta) = a_t^{-1} \mathbf{1}, \quad \psi_t(\theta) = X_{t-m}^{t-1} h(X_{t-m}^{t-1}) \phi\left(X_t - \theta^T X_{t-m}^{t-1}\right) \tag{2.3.7}$$

Nous avons

$$\begin{aligned} E_\theta\left[\phi\left(X_t - (\theta+u)^T X_{t-m}^{t-1}\right) / \mathcal{F}_{t-1}\right] &= E_\theta\left[\phi\left(Z_t - u^T X_{t-m}^{t-1}\right) / \mathcal{F}_{t-1}\right] \\ &= \int \phi\left(z - u^T X_{t-m}^{t-1}\right) g(z) dz \end{aligned}$$

Vérifions les conditions de lemme 4 dans l'annexe 1. Des conditions sur g et ϕ on a:

$$\int_{\mathbb{R}} |\phi(z-\omega)| \, g(z) dz \leq \int_{\mathbb{R}} (\sup_{z,\omega} |\phi(z-\omega)|) g(z) dz \leq M \int_{\mathbb{R}} g(z) dz = M$$

et par suit pour tout $\omega \neq 0$ nous avons: $\omega \int_{-\infty}^{+\infty} \phi(z-\omega) \, g(z) dz < 0$. Donc

$$G(\omega) := -\omega \int_{-\infty}^{+\infty} \phi(z-\omega) \, g(z) dz > 0$$

Comme $G(u^T X_{t-m}^{t-1})$, a_t, $h(X_{t-m}^{t-1})$ sont positives, nous avons donc

$$\begin{aligned} u^T\Gamma_t^{-1}(\theta+u)\,b_t(\theta,u) &= u^T a_t E_\theta\left[X_{t-m}^{t-1}h(X_{t-m}^{t-1})\phi\left(Z_t-u^TX_{t-m}^{t-1}\right)/\mathcal{F}_{t-1}\right] \\ &= -a_t h(X_{t-m}^{t-1})\left[-u^TX_{t-m}^{t-1}\int\limits_{-\infty}^{\infty}\phi\left(z-u^TX_{t-m}^{t-1}\right)g(z)dz\right] \\ &= -a_t h(X_{t-m}^{t-1})G(u^TX_{t-m}^{t-1})\leq 0 \end{aligned} \tag{2.3.8}$$

D'où $(C1)$ de corollaire [2.1]. D'autre part, ϕ etant bornée, nous avons

$$\begin{aligned} & E_\theta\left(\left\|\Gamma_t^{-1}(\theta+u)\,\psi_t(\theta+u)\right\|^2/\mathcal{F}_{t-1}\right) \\ = \;& E_\theta\left(\left\|a_tX_{t-m}^{t-1}h(X_{t-m}^{t-1})\phi\left(Z_t-u^TX_{t-m}^{t-1}\right)\right\|^2/\mathcal{F}_{t-1}\right) \\ = \;& a_t^2\left\|X_{t-m}^{t-1}\right\|^2h^2(X_{t-m}^{t-1})\int\limits_{\mathbb{R}}\left\|\phi\left(z-u^TX_{t-m}^{t-1}\right)\right\|^2g(z)dz \\ \leq \;& a_t^2\left\|X_{t-m}^{t-1}\right\|^2h^2(X_{t-m}^{t-1})\int\limits_{\mathbb{R}}C^\theta g(z)dz = C^\theta a_t^2\left\|X_{t-m}^{t-1}\right\|^2h^2(X_{t-m}^{t-1}) \\ = \;& C^\theta a_t^2\left\|X_{t-m}^{t-1}\right\|^2h^2(X_{t-m}^{t-1})+C^\theta a_t^2\left\|X_{t-m}^{t-1}\right\|^4h^2(X_{t-m}^{t-1}) \\ & -C^\theta a_t^2\left\|X_{t-m}^{t-1}\right\|^4h^2(X_{t-m}^{t-1}) \\ \leq \;& C^\theta a_t^2\left\|X_{t-m}^{t-1}\right\|^2h^2(X_{t-m}^{t-1})+C^\theta a_t^2\left\|X_{t-m}^{t-1}\right\|^4h^2(X_{t-m}^{t-1}) \\ = \;& C^\theta a_t^2\left\|X_{t-m}^{t-1}\right\|^2h^2(X_{t-m}^{t-1})\left(1+\left\|X_{t-m}^{t-1}\right\|^2\right) = B_t^\theta\left(1+\left\|X_{t-m}^{t-1}\right\|^2\right) \end{aligned}$$

où $B_t^\theta = C^\theta a_t^2\left\|X_{t-m}^{t-1}\right\|^2h^2(X_{t-m}^{t-1})$ pour $C^\theta>0$. Ainsi $(C2),(C3)$ sont satisfaites si

$$\sum_{t=1}^{\infty}a_th(X_{t-m}^{t-1})\inf_{\varepsilon\leq\|u\|\leq1/\varepsilon}G(u^TX_{t-m}^{t-1})=\infty,\sum_{t=1}^{\infty}a_t^2\left\|X_{t-m}^{t-1}\right\|^2h^2(X_{t-m}^{t-1})<\infty \quad p.s \tag{2.3.9}$$

Ces conditions sont vraies si X_t est un processus strictement stationnaire. Par exemple, si $a_t=\frac{1}{t}\longrightarrow 0$, $h(x)\neq 0$, $\forall x\neq 0$ et g continue et $\inf G(x)>0$, $\forall x\neq 0$. Alors pour tout $x\neq 0$: $h(x)\inf\limits_{\varepsilon\leq\|u\|\leq1/\varepsilon}G(u^Tx)>0$. Par le Lemme 4 dans l'annexe 1, nous avons $(A2)$: $\sup\limits_{\varepsilon\leq|\omega|\leq1/\varepsilon}\omega\int\limits_{-\infty}^{\infty}\phi(z-\omega)\,g(z)dz<0$, qui implique que $\inf\limits_{\varepsilon\leq|\omega|\leq1/\varepsilon}-\omega\int\limits_{-\infty}^{\infty}\phi(z-\omega)\,g(z)dz>0$, et par conséquent, $\inf\limits_{\varepsilon\leq\|u\|\leq1/\varepsilon}G(u^Tx)>0,\forall x\neq 0$, Par la loi forte des grands nombres pour les AR strictement stationnaires:

$$\lim_{t\longrightarrow\infty}\frac{1}{t}\sum_{s=1}^{t}h\left(X_{s-m}^{s-1}\right)\inf_{\varepsilon\leq\|u\|\leq1/\varepsilon}G(u^TX_{s-m}^{s-1})=E(h\left(X_{1-m}^{0}\right)\inf_{\varepsilon\leq\|u\|\leq1/\varepsilon}G(u^TX_{1-m}^{0}))>0$$

et

$$\lim_{t\longrightarrow\infty}\frac{1}{t}\sum_{s=1}^{t}\left\|X_{s-m}^{s-1}\right\|^2h^2(X_{s-m}^{s-1})=E(\left\|X_{1-m}^{0}\right\|^2h^2(X_{t-1}^{0}))>0$$

par lemme 5 dans l'annexe 1 donne la condition (2.3.9). Donc $\hat{\theta}_t\longrightarrow\theta\ P^{\theta}p.s.$

4. Cas Processus AR(1)

Considérons le processus

$$X_t=\theta X_{t-1}+Z_t,\qquad t\geq1$$

où Z_t de la loi de Student à α degrée de liberté. La densité de Z_t est

$$g(x)=C_\alpha\left(1+\frac{x^2}{\alpha}\right)^{-(\alpha+1)/2}\text{, où }C_\alpha=\Gamma\left(\left(\alpha+1\right)/2\right)/\left(\sqrt{\pi\alpha}\Gamma\left(\alpha/2\right)\right)$$

On a $\frac{\acute{g}(z)}{g(z)}=-(\alpha+1)\frac{z}{\alpha+z^2}$. Par (2.3.5), nous avons

$$\frac{\dot{f}_t^T\left(\theta,X_t/X_{t-1}\right)}{f_t\left(\theta,X_t/X_{t-1}\right)}=-X_{t-1}\frac{\acute{g}\left(X_t-\theta X_{t-1}\right)}{g\left(X_t-\theta X_{t-1}\right)}=\left(\alpha+1\right)X_{t-1}\frac{X_t-\theta X_{t-1}}{\alpha+\left(X_t-\theta X_{t-1}\right)^2}$$

L'information de Fisher est $I_t = i^g \sum\limits_{s=1}^{t} X_{t-1}^2$, où

$$\begin{aligned}
i^g &= \int \left(\frac{\acute{g}\,(z)}{g(z)}\right)^2 g(z)dz = \int \left(-\left(\alpha+1\right)\frac{z}{\alpha+z^2}\right)^2 C_\alpha \left(1+\frac{z^2}{\alpha}\right)^{-\frac{\alpha+1}{2}} dz \\
&= C_\alpha \left(\alpha+1\right)^2 \int \frac{z^2 dz}{\alpha^2 \left(1+\frac{z^2}{\alpha}\right)^{\frac{\alpha+5}{2}}} \\
&= C_\alpha \frac{\left(\alpha+1\right)^2}{\sqrt{\alpha}} \int \frac{u^2 du}{\left(1+u^2\right)^{(\alpha+5)/2}} \text{ (ch.v } u = \frac{z}{\sqrt{\alpha}}) \\
&= C_\alpha \frac{\left(\alpha+1\right)^2}{\sqrt{\alpha}} \int\limits_0^\infty \frac{\sqrt{v}dv}{\left(1+v\right)^{(\alpha+5)/2}} \\
&= C_\alpha \frac{\left(\alpha+1\right)^2}{\sqrt{\alpha}} \frac{\Gamma\left(\frac{3}{2}\right)\Gamma\left(\frac{\alpha+5}{2}-\frac{3}{2}\right)}{\Gamma\left(\frac{\alpha+5}{2}\right)} = \frac{\alpha+1}{\alpha+3}
\end{aligned}$$

Par conséquent, l'estimateur de maximum de vraissemblance est

$$\hat{\theta}_t = \hat{\theta}_{t-1} + I_t^{-1}(\hat{\theta}_{t-1})\left(\alpha+1\right)X_{t-1}\frac{X_t - \hat{\theta}_{t-1}X_{t-1}}{\alpha + \left(X_t - \hat{\theta}_{t-1}X_{t-1}\right)^2}, \qquad t \geq 1 \tag{2.3.10}$$

Nous notons que pour $I_t = I_{t-1} + i^g X_{t-1}^2$, (2.3.10) donne une procedure récursive de type (2.3.6) avec une suite stochastique $a_t = I_t^{-1}$. Maintenant ψ_t est de la forme de (2.3.7) avec $h(u) = 1$; $\phi(z) = \left(\alpha+1\right)z/\left(\alpha+z^2\right)$, et g est symétrique. Par conséquent, pour obtenir la convergence vers θ il suffit de verifier les conditions (2.3.9) qui s'écrivent comme suit

$$\sum_{t=1}^{\infty} \frac{1}{I_t} \inf_{\varepsilon \leq |u| \leq 1/\varepsilon} G\left(uX_{t-1}\right) = \infty, \sum_{t=1}^{\infty} \frac{X_{t-1}^2}{I_t^2} < \infty, \qquad P^\theta p.s \tag{2.3.11}$$

Nous avons $\forall\theta \in \mathbb{R} : I_t\left(\theta\right) = \sum\limits_{s=1}^{t} i_s\left(\theta\right) = i^g \sum\limits_{s=1}^{t} X_{s-1}^2 \longrightarrow \infty$ (cf.[17],Ch.VII, 5.5) et $\Delta I_t = i^g X_{t-1}^2$, par le Lemme 4 dans l'annexe 1, avec $d_n = I_n, \varepsilon = 1$ la deuxieme condition de (2.3.11) devient

$$\sum_{t=1}^{\infty} \frac{X_{t-1}^2}{I_t^2} = \frac{1}{i^g} \sum_{t=1}^{\infty} \frac{\Delta I_t}{I_t^2} < \infty$$

où $X_{s-1}^2 > 0$ et non-décroissante ($\Delta I_t = i^g X_{t-1}^2 > 0; i^g > 0$).

Supposons que $|\theta| < 1$, par le Lemme 4 dans l'annexe 1, nous avons $G\left(ux\right) = -ux\Phi\left(ux\right)$, et d'aprés $(A2)$, $\sup\limits_{\varepsilon\leq|u|\leq 1/\varepsilon}(-G(ux)) = -\inf\limits_{\varepsilon\leq|u|\leq 1/\varepsilon} G(ux) < 0$, donc $\inf\limits_{\varepsilon\leq|u|\leq\ /\varepsilon} G(ux) = \delta > 0$, $\forall x \neq 0$. Alors par la loi forte des grands nombres

$$\frac{1}{t}I_t \longrightarrow c = \frac{\sigma^2}{1-\theta^2} > 0 \text{ et} \frac{1}{t}\sum_{t=1}^{t} \inf_{\varepsilon\leq|u|\leq 1/\varepsilon} G(uX_{s-1}) \to E(\inf_{\varepsilon\leq|u|\leq 1/\varepsilon} G(uX_1)) > 0 \ \ p.s$$

d'où $\lim \frac{1}{I_t}\sum\limits_{t=1}^{t} \inf\limits_{\varepsilon\leq|u|\leq 1/\varepsilon} G(uX_{s-1}) > 0\, p.s$, et par lemme 5 dans l'annexe 1 $\sum\limits_{t=1}^{\infty} \frac{1}{I_t} \inf\limits_{\varepsilon\leq|u|\leq 1/\varepsilon} G(uX_{t-1}) = \infty$. Donc la première condition de(2.3.11) est satisfaite.

Remarque 11 *Nous avons montré que l'estimateur récursif* (2.3.10) *converge vers* θ *p.s, si* $|\theta| < 1$*. Il est à noter que la deuxieme condition de* (2.3.11) *et* (2.3.9) *est verifiée pour tout* $\theta \in \mathbb{R}$ *qui garantit* $(C3)$ *de corollaire [2.1]. De même* (2.3.8) *implique que* $(C1)$ *a lieu. Par suit on a* $\hat{\theta}_t \to \theta$ P^θ*p.s (cf remarque 8).*

Chapitre 3

Vitesse de convergence et développement asymptotique

3.1 introduction

Considérons une classe d'estimateur $(\hat{\theta}_n)$ défini récursivement par

$$\hat{\theta}_n = \hat{\theta}_{n-1} + \Gamma_t^{-1}(\hat{\theta}_{n-1})\psi_n(\hat{\theta}_{n-1}), \quad n \geq 1 \tag{3.1.1}$$

où:

ψ_n est une fonction vectorielle quelconque.

Γ_n est un processs gaussien matriciel (peut être déterministe).

$\hat{\theta}_0 \in \mathbb{R}^m$ est un point initial.

Dans cette partie nous donnons des conditions suffisantes (différentes de celles du Ch.2) vitesse de cette classe d'estimateurs. En plus la classe de M-estimateurs contient des estimateurs ayant une représentation asymptotique suivante: sous des conditions d'ergodicité il existe une suite d'estimateurs solution de (3.1.1) vérifiant

$$a_n^{1/2}(\theta)(\theta_n - \theta) = a_n^{-1/2}(\theta)\sum_{i=1}^{n}\psi_i(\theta) + o_{p^\theta}(1) \tag{3.1.2}$$

ou encore

$$\theta_n = \theta + a_n^{-1}(\theta) \sum_{i=1}^{n} \psi_i(\theta) + o_{p^\theta}(1)$$

où $a_n(\theta) = a_n(\theta, \psi_n)$ est un processus matriciel réel déterministe et det $a_t(\theta) \neq 0$.

3.2 Vitesse de convergence

la vitesse est donnée par le résultat suivant. Ce résultat est similaire au Corollaire [1.1] Ch.1 (a part le coefficient 2 dans la condition $(R2)$).

Proposition 12 *Soit* $(a_t(\theta))_{t\geq 1}$ *un processus scalaire prévisible non-décroissant et* $a_t(\theta) \longrightarrow_{t\longrightarrow\infty} \infty$. *Notons* $\Delta a_t(\theta) = a_t(\theta) - a_{t-1}(\theta)$ *et supposons que*

$(R1)$ $\lim\limits_{t\longrightarrow\infty} \frac{\Delta a_t(\theta)}{a_{t-1}} = 0;$

$(R2)$ *il existe une matrice* C_θ *symétrique et définie positive tele que*

$$\left\langle C_\theta \Delta_{t-1}, \Gamma_t^{-1}(\theta + \Delta_{t-1})\, b_t(\theta, \Delta_{t-1})\right\rangle \leq -\lambda_t(\theta) \langle C_\theta \Delta_{t-1}, \Delta_{t-1}\rangle$$

où $(\lambda_t(\theta))_{t\geq 1}$ *est un processus scalaire prévisible satisfait*

$$\sum_{s=1}^{\infty} \left[\frac{\Delta a_t(\theta)}{a_t(\theta)} - 2\lambda_t(\theta)\right]^+ < \infty$$

$(R3)$ *pour tout* $0 < \varepsilon < 1$

$$\sum_{s=1}^{\infty} a_t^{\varepsilon}(\theta)\, E_\theta \left(\left\|\Gamma_t^{-1}(\theta + \Delta_{t-1})\, \psi_t(\theta + \Delta_{t-1})\right\|^2 / \mathcal{F}_{t-1}\right) < \infty$$

Alors pour tout $\delta \in]0, 1/2[$ $a_t(\theta)^\delta (\hat{\theta}_t - \theta) \longrightarrow 0$.

Preuve: Nous verifions les conditions du lemme 1 Ch 1 avec $C_t(\theta) = C_\theta (a_t(\theta))^{2\delta}$, $\delta \in]0, 1/2[$ nous notons

$$\mathcal{P}_t := a_t^{2\delta} E_\theta \left(\left[\Gamma_t^{-1}(\theta + \Delta_{t-1})\, \psi_t(\theta + \Delta_{t-1})\right]^T C_t(\theta) \left[\Gamma_t^{-1}(\theta + \Delta_{t-1})\, \psi_t(\theta + \Delta_{t-1})\right] / \mathcal{F}_{t-1}\right)$$

et

$$r_t = \frac{\Delta a_t^{2\delta} - 2a_t^{2\delta}\lambda_t}{a_{t-1}^{2\delta}}$$

On a $V_t(u) := \langle C_t(\theta) u, u\rangle = \left\langle C_\theta (a_t(\theta))^{2\delta} u, u\right\rangle$ et alors pour le processus $\mathcal{K}_t$ défini dans (1.3.3) (Ch 1), nous avons par $(R2)$

$$\begin{aligned}
\mathcal{K}_t \quad : \quad & = \Delta V_t(\Delta_{t-1}) + 2\left\langle C_t(\theta)\Delta_{t-1}, \Gamma_t^{-1}(\theta+\Delta_{t-1}) b_t(\theta, \Delta_{t-1})\right\rangle \\
& \quad + E_\theta\left(\left[\Gamma_t^{-1}(\theta+\Delta_{t-1})\psi_t(\theta+\Delta_{t-1})\right]^T C_t(\theta)\left[\Gamma_t^{-1}(\theta+\Delta_{t-1})\psi_t(\theta+\Delta_{t-1})\right]/\mathcal{F}_{t-1}\right) \\
& = \Delta a_t^{2\delta}\langle C_\theta\Delta_{t-1}, \Delta_{t-1}\rangle + 2a_t^{2\delta}\left\langle C_\theta\Delta_{t-1}, \Gamma_t^{-1}(\theta+\Delta_{t-1}) b_t(\theta,\Delta_{t-1})\right\rangle + \mathcal{P}_t \\
& \leq \Delta a_t^{2\delta}\langle C_\theta\Delta_{t-1}, \Delta_{t-1}\rangle + 2a_t^{2\delta}\left(-\lambda_t\langle C_\theta\Delta_{t-1}, \Delta_{t-1}\rangle\right) + \mathcal{P}_t \quad \text{(par } (R2))\\
& \leq \left(\Delta a_t^{2\delta} - 2a_t^{2\delta}\lambda_t\right)\langle C_\theta\Delta_{t-1}, \Delta_{t-1}\rangle + \mathcal{P}_t \\
& \leq a_{t-1}^{2\delta}\left(\frac{\Delta a_t^{2\delta} - 2a_t^{2\delta}\lambda_t}{a_{t-1}^{2\delta}}\right)\langle C_\theta\Delta_{t-1}, \Delta_{t-1}\rangle + \mathcal{P}_t \\
& = r_t\left\langle a_{t-1}^{2\delta} C_\theta\Delta_{t-1}, \Delta_{t-1}\right\rangle + \mathcal{P}_t
\end{aligned}$$

Comme les processus $C_\theta, \mathcal{P}_t$ sont définies positives alors $[\mathcal{P}_t]^+ = [\mathcal{P}_t]$ et donc

$$\begin{aligned}
(1+V_{t-1}(\Delta_{t-1}))^{-1}[\mathcal{K}_t]^+ \quad & = \quad \left(1+\left\langle a_{t-1}^{2\delta} C_\theta\Delta_{t-1}, \Delta_{t-1}\right\rangle\right)^{-1}[\mathcal{K}_t]^+ \\
& \leq \quad \left(1+\left\langle a_{t-1}^{2\delta} C_\theta\Delta_{t-1}, \Delta_{t-1}\right\rangle\right)^{-1}\left([r_t]^+\left\langle a_{t-1}^{2\delta} C_\theta\Delta_{t-1}, \Delta_{t-1}\right\rangle + \mathcal{P}_t\right) \\
& = \quad \frac{\left\langle a_{t-1}^{2\delta} C_\theta\Delta_{t-1}, \Delta_{t-1}\right\rangle}{1+\left\langle a_{t-1}^{2\delta} C_\theta\Delta_{t-1}, \Delta_{t-1}\right\rangle}[r_t]^+ + \frac{1}{1+\left\langle a_{t-1}^{2\delta} C_\theta\Delta_{t-1}, \Delta_{t-1}\right\rangle}\mathcal{P}_t \\
& \leq \quad [r_t]^+ + \mathcal{P}_t
\end{aligned}$$

car $\frac{\left\langle a_{t-1}^{2\delta} C_\theta\Delta_{t-1}, \Delta_{t-1}\right\rangle}{1+\left\langle a_{t-1}^{2\delta} C_\theta\Delta_{t-1}, \Delta_{t-1}\right\rangle} \leq 1$, et $\frac{1}{1+\left\langle a_{t-1}^{2\delta} C_\theta\Delta_{t-1}, \Delta_{t-1}\right\rangle} \leq 1$. Donc on a

$$\sum_{t=1}^{\infty}(1+V_{t-1}(\Delta_{t-1}))^{-1}[\mathcal{K}_t]^+ \leq \sum_{t=1}^{\infty}[r_t]^+ + \sum_{t=1}^{\infty}\mathcal{P}_t$$

Par $(R3)$ avec $\varepsilon = 2\delta$, nous avons $\sum_{t=1}^{\infty} \mathcal{P}_t < \infty$. Montrons que $\sum_{t=1}^{\infty} [r_t]^+ < \infty$. Comme $\Delta a_t^{2\delta} = a_t^{2\delta} - a_{t-1}^{2\delta}$, nous peuvons écrire r_t comme suit

$$\begin{aligned} r_t &= \left(\Delta a_t^{2\delta} - 2a_t^{2\delta}\lambda_t\right)\left(a_{t-1}^{-2\delta}\right) \\ &= \left(a_t^{2\delta} - a_{t-1}^{2\delta} - 2a_t^{2\delta}\lambda_t\right)\left(a_{t-1}^{-2\delta}\right) \\ &= \left(a_t a_{t-1}^{-1}\right)^{2\delta} - 1 - 2\lambda_t \left(a_t a_{t-1}^{-1}\right)^{2\delta} \\ &= \left(a_t a_{t-1}^{-1}\right)^{2\delta}(1 - 2\lambda_t) - 1 \end{aligned}$$

Par le développement limité $(1+x)^{2\delta} = 1 + 2\delta x + O\left(x^2\right)$ on a

$$\left(a_t a_{t-1}^{-1}\right)^{2\delta} = \left(\frac{a_{t-1} + a_t - a_{t-1}}{a_{t-1}}\right)^{2\delta} = \left(1 + \frac{\Delta a_t}{a_{t-1}}\right)^{2\delta} = 1 + 2\delta\frac{\Delta a_t}{a_{t-1}} + \delta_t^{(1)}$$

où $\delta_t^{(1)} = O\left(\Delta a_t / a_{t-1}\right)^2 \longrightarrow_{t\longrightarrow\infty} 0$ par $(R1)$. Notons par

$$\eta_t = \frac{\Delta a_t}{a_t} - 2\lambda_t$$

et du fait $[\eta_t]^+ \geq \eta_t$ on a

$$\begin{aligned} r_t &= \left(a_t a_{t-1}^{-1}\right)^{2\delta}(1 - 2\ \lambda_t) - 1 = \left(a_t a_{t-1}^{-1}\right)^{2\delta}\left(1 + \eta_t - \frac{\Delta a_t}{a_t}\right) - 1 \\ &\leq \left(a_t a_{t-1}^{-1}\right)^{2\delta}\left(1 + \eta_t^+ - \frac{\Delta a_t}{a_t}\right) - 1 \\ &= \left(1 + 2\delta\frac{\Delta a_t}{a_{t-1}} + \delta_t^{(1)}\right)\left(1 + \eta_t^+ - \frac{\Delta a_t}{a_{t-1}}\right) - 1 \\ &= 1 + \eta_t^+ - \frac{\Delta a_t}{a_t} + 2\delta\frac{\Delta a_t}{a_{t-1}} + 2\delta\frac{\Delta a_t}{a_{t-1}}\eta_t^+ - 2\delta\frac{\Delta a_t}{a_{t-1}}\frac{\Delta a_t}{a_t} + \delta_t^{(1)} \\ &\quad +\delta_t^{(1)}\eta_t^+ - \delta_t^{(1)}\frac{\Delta a_t}{a_t} - 1 \end{aligned}$$

Mais nous avons

$$-\frac{\Delta a_t}{a_t} = -\frac{\Delta a_t a_{t-1}}{a_t a_{t-1}} = -\frac{\Delta a_t\left(a_t - \Delta a_t\right)}{a_t a_{t-1}} = -\frac{\Delta a_t}{a_{t-1}} + \frac{\Delta a_t}{a_t}\frac{\Delta a_t}{a_{t-1}}$$

et

$$\begin{aligned}
-\frac{\Delta a_t}{a_t}+2\delta\frac{\Delta a_t}{a_{t-1}}-2\delta\frac{\Delta a_t}{a_{t-1}}\frac{\Delta a_t}{a_t} &= -\frac{\Delta a_t}{a_{t-1}}+\frac{\Delta a_t}{a_t}\frac{\Delta a_t}{a_{t-1}} \\
&\quad +2\delta\frac{\Delta a_t}{a_{t-1}}-2\delta\frac{\Delta a_t}{a_{t-1}}\frac{\Delta a_t}{a_t} \\
&= -\left(1-2\delta\right)\frac{\Delta a_t}{a_{t-1}}+\left(1-2\delta\right)\frac{\Delta a_t}{a_t}\frac{\Delta a_t}{a_{t-1}}
\end{aligned}$$

Par suit

$$r_t \leq -\left(1-2\ \delta\right)\frac{\Delta a_t}{a_{t-1}}+\delta_t^{(1)}+\left(1-2\delta\right)\frac{\Delta a_t}{a_t}\frac{\Delta a_t}{a_{t-1}}-\frac{\Delta a_t}{a_t}\delta_t^{(1)}+\eta_t^{+}+2\delta\eta_t^{+}\frac{\Delta a_t}{a_{t-1}}+\eta_t^{+}\delta_t^{(1)}$$

Or

$$\begin{aligned}
& \delta_t^{(1)}+\left(1-2\delta\right)\frac{\Delta a_t}{a_t}\frac{\Delta a_t}{a_{t-1}}-\frac{\Delta a_t}{a_t}\delta_t^{(1)} \\
= \ & \frac{\Delta a_t}{a_{t-1}}\left(\left(\frac{\Delta a_t}{a_{t-1}}\right)^{-1}\delta_t^{(1)}-\frac{\Delta a_t}{a_t}\left(\frac{\Delta a_t}{a_{t-1}}\right)^{-1}\delta_t^{(1)}+\left(1-2\delta\right)\frac{\Delta a_t}{a_t}\right) \\
= \ & \frac{\Delta a_t}{a_{t-1}}\left(\left(\frac{\Delta a_t}{a_{t-1}}\right)^{-1}\delta_t^{(1)}\left(1-\frac{\Delta a_t}{a_t}\right)+\left(1-2\delta\right)\frac{\Delta a_t}{a_t}\right) \\
= \ & :\frac{\Delta a_t}{a_{t-1}}\delta_t^{(2)}
\end{aligned}$$

En posant $\delta_t^{(3)} := \eta_t^{+}+2\delta\eta_t^{+}\frac{\Delta a_t}{a_{t-1}}+\eta_t^{+}\delta_t^{(1)}=\eta_t^{+}(1+2\delta\frac{\Delta a_t}{a_{t-1}}+\delta_t^{(1)})$, on aboutit à

$$r_t \leq -\left(1-2\delta\right)\frac{\Delta a_t}{a_{t-1}}+\frac{\Delta a_t}{a_{t-1}}\delta_t^{(2)}+\delta_t^{(3)}=\frac{\Delta a_t}{a_{t-1}}\left(-\left(1-2\delta\right)+\delta_t^{(2)}\right)+\delta_t^{(3)}$$

Par $(R1)$ on a $\delta_t^{(1)}\longrightarrow 0$ et $\delta_t^{(2)}\longrightarrow 0$. Donc de $1-2\delta>0$ on a $\frac{\Delta a_t}{a_{t-1}}\left(-\left(1-2\delta\right)+\delta_t^{(2)}\right)\leq 0$, et par conséquent et pour $t>t_0$, nous avons

$$\left|\delta_t^{(3)}\right| = \eta_t^{+}\left|1+2\delta\frac{\Delta a_t}{a_{t-1}}+\delta_t^{(1)}\right| < \eta_t^{+}(1+\varepsilon)$$

et par $(R2)$ on a donc $\sum\limits_{t>t_0}^{\infty}\left|\delta_t^{(3)}\right| \leq (1+\varepsilon)\sum\limits_{t>t_0}^{\infty}\eta_t^{+} < \infty$ Par conséquent

$$\sum_{t=1}^{\infty}[r_t]^{+} \leq \sum_{t>t_0}^{\infty}\left|\delta_t^{(3)}\right| < \infty$$

ce qui donne

$$\sum_{t=1}^{\infty}\left(1+V_{t-1}\left(\Delta_{t-1}\right)\right)^{-1}\left[\mathcal{K}_t\right]^{+} < \infty$$

Les conditions de lemme 1 Ch 1 sont satisfaites et $V_t\left(\Delta_t\right) \longrightarrow 0$ et $a_t^{\delta}(\hat{\theta}_t - \theta) \longrightarrow 0$ pour tout $\delta \in \left]0, 1/2\right[$. ■

Corollaire [3.1] Considérons le cas iid avec $\psi_t\left(\theta\right) = \psi\left(\theta, X_t\right)$, et $\Gamma_t\left(\theta\right) = t\gamma\left(\theta\right)$. Supposons que $\hat{\theta} \longrightarrow \theta$ et

$(B1)$ il existe une matrice C_θ symétrique et définie positive telle que

$$\left\langle C_\theta u, \gamma^{-1}\left(\theta+u\right)E_\theta\psi\left(\theta+u, X_1\right)\right\rangle \leq -\frac{1}{2}\left\langle C_\theta u, u\right\rangle$$

pour u assez petit

$(B2)$ $E_\theta\left\|\gamma^{-1}\left(\theta+u\right)\psi\left(\theta+u\right)\right\|^2 = O\left(1\right)$ quand $u \longrightarrow 0.$

Alors, pour tout $\delta \in \left]0, 1/2\right[$, $t^{\delta}(\hat{\theta}_t - \theta) \longrightarrow 0$ $p.s.$

Preuve: On choisit $a_t\left(\theta\right) = t$ qui est un processus scalaire prévisible croissant et $a_t\left(\theta\right) = t \longrightarrow \infty$; $\Delta a_t\left(\theta\right) = t - (t-1) = 1$. Vérifions les conditions de la proposition precedente. Nous avons $\lim\limits_{t\longrightarrow\infty}\frac{\Delta a_t}{a_{t-1}} = \lim\limits_{t\longrightarrow\infty}\frac{1}{t-1} = 0$ donc $(R1)$ est satisfaite. Posons $\lambda_t\left(\theta\right) = 1/2t$, qui est un processus scalaire prévisible et

$$\sum_{t=1}^{\infty}\left[\frac{\Delta a_t}{a_t} - 2\lambda_t\right]^{+} = \sum_{t=1}^{\infty}\left[\frac{1}{t} - 2\frac{1}{2t}\right]^{+} = 0$$

Par $(B1)$, nous avons

$$\begin{aligned}\left\langle C_\theta\Delta_{t-1}, \Gamma_t^{-1}\left(\theta+\Delta_{t-1}\right)b_t\left(\theta, \Delta_{t-1}\right)\right\rangle &= \left\langle C_\theta\Delta_{t-1}, \frac{1}{t}\gamma^{-1}\left(\theta+\Delta_{t-1}\right)E^{\theta}\psi\left(\theta+\Delta_{t-1}, X_1\right)\right\rangle \\ &\leq -\frac{1}{2t}\left\langle C_\theta\Delta_{t-1}, \Delta_{t-1}\right\rangle = -\lambda_t\left(\theta\right)\left\langle C_\theta\Delta_{t-1}, \Delta_{t-1}\right\rangle\end{aligned}$$

Donc $(R2)$ est satisfaite. Pour $(R3)$, soit $0<\varepsilon<1$, par $(B2)$ nous avons

$$\begin{aligned}
& \sum_{t=1}^{\infty} a_t^{\varepsilon}(\theta)\, E_\theta\left(\left\|\Gamma_t^{-1}(\theta+\Delta_{t-1})\,\psi_t(\theta+\Delta_{t-1})\right\|^2/\mathcal{F}_{t-1}\right) \\
= & \sum_{t=1}^{\infty} t^{\varepsilon} E_\theta\left(\left\|\frac{1}{t}\gamma^{-1}(\theta+\Delta_{t-1})\,\psi(\theta+\Delta_{t-1},X_1)\right\|^2/\mathcal{F}_{t-1}\right) \\
= & \sum_{t=1}^{\infty} t^{\varepsilon-2} E_\theta\left(\left\|\gamma^{-1}(\theta+\Delta_{t-1})\,\psi(\theta+\Delta_{t-1},X_1)\right\|^2/\mathcal{F}_{t-1}\right) \\
\leq & \; k\sum_{t=1}^{\infty} t^{\varepsilon-2}<\infty
\end{aligned}$$

D'où la condition $(R3)$. Finalement $\forall\delta\in\left]0,1/2\right[$, $a_t(\theta)^{\delta}\,(\hat{\theta}_t-\theta)=t^{\delta}(\hat{\theta}_t-\theta)\longrightarrow 0$. ■

3.3 Développement asymptotique des estimateurs

Dans cette partie nous donnons des développements asymptotiques des M-estimateurs formule des estmateurs recursifs

$$\hat{\theta}_t=\hat{\theta}_{t-1}+\Gamma_t^{-1}(\hat{\theta}_{t-1})\psi_t(\hat{\theta}_{t-1})$$

ou avec $\Delta_t=\hat{\theta}_t-\theta$

$$\hat{\theta}_t-\theta=\hat{\theta}_{t-1}-\theta+\Gamma_t^{-1}(\theta+\Delta_{t-1})\,\psi_t(\theta+\Delta_{t-1})$$

Le but est de donner un développement asymptotique de $\hat{\theta}_t$ dont le terme principal est une statistique linéaire de la forme suivante

$$\hat{\theta}_t^{*}=\theta+\Gamma_t^{-1}(\theta)\sum_{s=1}^{t}\psi_s(\theta) \tag{3.3.1}$$

Notons

$$C_t(\theta,u)=\left\{\begin{array}{ll} -\Gamma_t(\theta)\,\Gamma_t^{-1}(\theta+u)\,b_t(\theta,u)\,\acute{u}/\left\|u\right\|^2 & \text{si: } u\neq 0 \\ \Delta\Gamma_t(\theta) & \text{si: } u=0 \end{array}\right\} \tag{3.3.2}$$

et nous avons

$$\begin{aligned}
\Delta_t &= \Delta_{t-1} + \Gamma_t^{-1}(\theta + \Delta_{t-1})\,\psi_t(\theta + \Delta_{t-1}) \\
&= \Delta_{t-1} + \Gamma_t^{-1}(\theta + \Delta_{t-1})\,b_t(\theta, \Delta_{t-1}) + \Gamma_t^{-1}(\theta + \Delta_{t-1})\,\psi_t(\theta + \Delta_{t-1}) \\
&\quad -\Gamma_t^{-1}(\theta + \Delta_{t-1})\,b_t(\theta, \Delta_{t-1}) \\
&= \big(1 + (\Gamma_t^{-1}(\theta + \Delta_{t-1})\,b_t(\theta, \Delta_{t-1})\,\Delta_{t-1}^T/\Delta_{t-1}^T\Delta_{t-1})\big)\,\Delta_{t-1} \\
&\quad +\Gamma_t^{-1}(\theta + \Delta_{t-1})\,\psi_t(\theta + \Delta_{t-1}) - \Gamma_t^{-1}(\theta + \Delta_{t-1})\,b_t(\theta, \Delta_{t-1}) \\
&= \Big(1 + (\Gamma_t^{-1}(\theta)\,\Gamma_t(\theta)\,\Gamma_t^{-1}(\theta + \Delta_{t-1})\,b_t(\theta, \Delta_{t-1})\,\Delta_{t-1}^T/\left\|\Delta_{t-1}\right\|^2)\Big)\,\Delta_{t-1} \\
&\quad +\Gamma_t^{-1}(\theta + \Delta_{t-1})\,(\psi_t(\theta + \Delta_{t-1}) - b_t(\theta, \Delta_{t-1})) \\
&= \big(1 - \Gamma_t^{-1}(\theta)\,C_t(\theta, \Delta_{t-1})\big)\,\Delta_{t-1} \\
&\quad +\Gamma_t^{-1}(\theta)\,\Gamma_t(\theta)\,\Gamma_t^{-1}(\theta + \Delta_{t-1})\,(\psi_t(\theta + \Delta_{t-1}) - b_t(\theta, \Delta_{t-1})) \\
&= \big(1 - \Gamma_t^{-1}(\theta)\,C_t(\theta, \Delta_{t-1})\big)\,\Delta_{t-1} + \Gamma_t^{-1}(\theta)\,\varepsilon_t^\theta \qquad (3.3.3)
\end{aligned}$$

oú $\varepsilon_t^\theta = \Gamma_t(\theta)\,\Gamma_t^{-1}(\theta + \Delta_{t-1})\,(\psi_t(\theta + \Delta_{t-1}) - b_t(\theta, \Delta_{t-1}))$ est une différence de martingale.

Soit $\Delta_0^* = 0$, et pour $t \geq 1$, notons $\Delta_t^* = \hat{\theta}_t^* - \theta$. La différence Δ_t^* satisfait la relation

$$\Delta_t^* = \hat{\theta}_t^* - \theta = \Gamma_t^{-1}(\theta) \sum_{s=1}^{t} \psi_s(\theta) \qquad (3.3.4)$$

et la récurence suivante

$$
\begin{aligned}
\Delta_t^* &= \Gamma_t^{-1}(\theta)\sum_{s=1}^{t}\psi_s(\theta) \\
&= \Gamma_t^{-1}(\theta)\left[\Gamma_{t-1}(\theta)\Gamma_{t-1}^{-1}(\theta)\sum_{s=1}^{t-1}\psi_s(\theta)+\psi_t(\theta)\right] \\
&= \Gamma_t^{-1}(\theta)\left(\Gamma_{t-1}(\theta)\Delta_{t-1}^*+\psi_t(\theta)\right) \\
&= \Gamma_t^{-1}(\theta)\Gamma_{t-1}(\theta)\Delta_{t-1}^*+\Gamma_t^{-1}(\theta)\psi_t(\theta) \\
&= \left(1-\Gamma_t^{-1}(\theta)\Gamma_t(\theta)+\Gamma_t^{-1}(\theta)\Gamma_{t-1}(\theta)\right)\Delta_{t-1}^*+\Gamma_t^{-1}(\theta)\psi_t(\theta) \\
&= \left[1-\Gamma_t^{-1}(\theta)\left(\Gamma_t(\theta)-\Gamma_{t-1}(\theta)\right)\right]\Delta_{t-1}^*+\Gamma_t^{-1}(\theta)\psi_t(\theta) \\
&= \left(1-\Gamma_t^{-1}(\theta)\Delta\Gamma_t(\theta)\right)\Delta_{t-1}^*+\Gamma_t^{-1}(\theta)\psi_t(\theta) \\
&= \left(1-\Gamma_t^{-1}(\theta)\Delta\Gamma_t(\theta)\right)\Delta_{t-1}^*+\Gamma_t^{-1}(\theta)\varepsilon_t^* \qquad (3.3.5)
\end{aligned}
$$

où $\varepsilon_t^\theta=\psi_t(\theta)$ est martingale-différence. Par comparaison entre les équations (3.3.3) et (3.3.5), nous peuvons obtenus le résultat suivant à les relations asymptotiques entre $\hat{\theta}_t$ et $\hat{\theta}_t^*$.

Théorème 13 *Supposons qu'il existe une suite non-aléatoire de matrices diagonales inversibles $A_t(\theta)$, telles que*

(E) $A_t(\theta)\Gamma_t^{-1}(\theta)A_t(\theta)\longrightarrow\eta_t(\theta)$ *où* $\eta_t(\theta)$ *est une matrice aléatoire de* $\det\eta_t(\theta)\neq 0$;

(L1)

$$\lim_{t\longrightarrow\infty}A_t^{-1}(\theta)\sum_{s=1}^{t}\left(\Delta\Gamma_s(\theta)-C_s(\theta,\Delta_{s-1})\right)\Delta_{s-1}=0 \quad P^\theta\text{-probabilité}$$

(L2)

$$\lim_{t\longrightarrow\infty}\sum_{s=1}^{t}E_\theta\left(\left\|A_t^{-1}(\theta)\mathcal{E}_s(\theta)\right\|^2/\mathcal{F}_{t-1}\right)=0 \quad P^\theta\text{-probabilité}$$

où

$$\mathcal{E}_s(\theta)=\Gamma_s(\theta)\Gamma_s^{-1}(\theta+\Delta_{s-1})\left(\psi_s(\theta+\Delta_{s-1})-b_s(\theta,\Delta_{s-1})\right)-\psi_s(\theta)$$

Alors $A_t(\theta)(\hat{\theta}_t^*-\hat{\theta}_t)\longrightarrow 0$ P^θ*-probabilité.*

Preuve: Notons $\delta_t = \hat{\theta}_t - \hat{\theta}_t^*$. Par soustraction (3.3.5) de (3.3.3) on a

$$\begin{aligned}
\delta_t &= \hat{\theta}_t - \hat{\theta}_t^* = (\hat{\theta}_t - \theta) - (\hat{\theta}_t^* - \theta) = \Delta_t - \Delta_t^* \\
&= \left(1 - \Gamma_t^{-1}(\theta)\, C_t(\theta, \Delta_{t-1})\right) \Delta_{t-1} + \Gamma_t^{-1}(\theta)\, \varepsilon_t^\theta \\
&\quad - \left(1 - \Gamma_t^{-1}(\theta) \Delta\Gamma_t(\theta)\right) \Delta_{t-1}^* - \Gamma_t^{-1}(\theta)\, \varepsilon_t^* \\
&= \Delta_{t-1} - \Gamma_t^{-1}(\theta)\, C_t(\theta, \Delta_{t-1})\, \Delta_{t-1} + \Gamma_t^{-1}(\theta)\, (\varepsilon_t^\theta - \varepsilon_t^*) \\
&\quad -\Delta_{t-1}^* + \Gamma_t^{-1}(\theta)\, \Delta\Gamma_t(\theta)\, \Delta_{t-1}^* \\
&= \Delta_{t-1} - \Delta_{t-1}^* - \Gamma_t^{-1}(\theta)\, C_t(\theta, \Delta_{t-1})\, \Delta_{t-1} + \Gamma_t^{-1}(\theta)\, (\varepsilon_t^\theta - \varepsilon_t^*) \\
&\quad +\Gamma_t^{-1}(\theta)\, \Delta\Gamma_t(\theta)\, \Delta_{t-1}^* + \Gamma_t^{-1}(\theta)\, \Delta\Gamma_t(\theta)\, \Delta_{t-1} - \Gamma_t^{-1}(\theta)\, \Delta\Gamma_t(\theta)\, \Delta_{t-1} \\
&= \delta_{t-1} + \Gamma_t^{-1}\left(\Delta\Gamma_t - C_t(\theta, \Delta_{t-1})\right) \Delta_{t-1} + \Gamma_t^{-1}(\varepsilon_t^\theta - \varepsilon_t^*) + \Gamma_t^{-1}\Delta\Gamma_t(\Delta_{t-1}^* - \Delta_{t-1}) \\
&= \delta_{t-1} - \Gamma_t^{-1}\Delta\Gamma_t\delta_{t-1} + \Gamma_t^{-1}(\varepsilon_t^\theta - \varepsilon_t^*) + \Gamma_t^{-1}\left(\Delta\Gamma_t - C_t(\theta, \Delta_{t-1})\right) \Delta_{t-1} \\
&= \left(1 - \Gamma_t^{-1}\Delta\Gamma_t\right) \delta_{t-1} + \Gamma_t^{-1}(\varepsilon_t^\theta - \varepsilon_t^*) + \Gamma_t^{-1}\left(\Delta\Gamma_t - C_t(\theta, \Delta_{t-1})\right) \Delta_{t-1}
\end{aligned}$$

Posons

$$M_t := \sum_{s=1}^{t} \left[\varepsilon_s - \varepsilon_s^*\right] \text{ et } \mathcal{H}_t := \sum_{s=1}^{t} \left[\Delta\Gamma_s - C_s(\theta, \Delta_{s-1})\right] \Delta_{s-1}$$

Alors on a

$$\delta_t = \Gamma_t^{-1}\left(M_t + \mathcal{H}_t + \delta_0\right), \qquad t \geq 1$$

Nous avons aussi

$$\begin{aligned}
\delta_t - \delta_{t-1} &= \Gamma_t^{-1}\left(M_t + \mathcal{H}_t + \delta_0\right) - \Gamma_{t-1}^{-1}\left(M_{t-1} + \mathcal{H}_{t-1} + \delta_0\right) \\
&= \Gamma_t^{-1}\left\{M_{t-1} + \left(\varepsilon_t - \varepsilon_t^*\right) + \mathcal{H}_{t-1} + \left(\Delta\Gamma_t - C_t\left(\theta, \Delta_{t-1}\right)\right)\Delta_{t-1} + \delta_o\right\} \\
&\quad -\Gamma_{t-1}^{-1}\left\{M_{t-1} + \mathcal{H}_{t-1} + \delta_0\right\} \\
&= \left(\Gamma_t^{-1} - \Gamma_{t-1}^{-1}\right)\left\{M_{t-1} + \mathcal{H}_{t-1} + \delta_0\right\} + \Gamma_t^{-1}\left(\varepsilon_t - \varepsilon_t^*\right) \\
&\quad +\Gamma_t^{-1}\left(\Delta\Gamma_t - C_t\left(\theta, \Delta_{t-1}\right)\right)\Delta_{t-1} \\
&= \left(\Gamma_t^{-1}\Gamma_{t-1} - Id\right)\Gamma_{t-1}^{-1}\left\{M_{t-1} + \mathcal{H}_{t-1} + \delta_0\right\} + \Gamma_t^{-1}\left(\varepsilon_t - \varepsilon_t^*\right) \\
&\quad +\Gamma_t^{-1}\left(\Delta\Gamma_t - C_t\left(\theta, \Delta_{t-1}\right)\right)\Delta_{t-1} \\
&= -\Gamma_t^{-1}\left(\Gamma_t - \Gamma_{t-1}\right)\delta_{t-1} + \Gamma_t^{-1}\left(\varepsilon_t - \varepsilon_t^*\right) + \Gamma_t^{-1}\left(\Delta\Gamma_t - C_t\left(\theta, \Delta_{t-1}\right)\right)\Delta_{t-1} \\
&= -\Gamma_t^{-1}\Delta\Gamma_t\delta_{t-1} + \Gamma_t^{-1}\left(\varepsilon_t - \varepsilon_t^*\right) + \Gamma_t^{-1}\left(\Delta\Gamma_t - C_t\left(\theta, \Delta_{t-1}\right)\right)\Delta_{t-1}
\end{aligned}$$

Montrons que $A_t(\hat{\theta}_t - \hat{\theta}_t^*) = A_t\delta_t = A_t\Gamma_t^{-1}A_tA_t^{-1}\left(M_t + \mathcal{H}_t + \delta_0\right) \longrightarrow 0$ en P^θ- probabilité.

Par (E) on a: $A_t\Gamma_t^{-1}A_t \longrightarrow \eta_t$ et $A_t^{-1}\delta_0 \longrightarrow 0$. Par $(L1)$ on a $A_t^{-1}\mathcal{H}_t \longrightarrow 0$ en probabilité. Reste à montrer que $A_t^{-1}M_t \longrightarrow 0$ en probabilité. Notons le vecteur différence de martingale $M_t = (M_t^{(j)})$ et le crochet $\left\langle M^{(j)}\right\rangle_t$ est

$$\begin{aligned}
\left\langle M^{(j)}\right\rangle_t &= \left\langle \sum_{s=1}^{t}\left[\varepsilon_s^{(j)} - \varepsilon_s^{*(j)}\right]\right\rangle = \sum_{s=1}^{t}\left\langle \varepsilon_s^{(j)} - \varepsilon_s^{*(j)}\right\rangle \\
&= \sum_{s=1}^{t}\left\langle \varepsilon_s^{(j)} - \varepsilon_s^{*(j)}, \varepsilon_s^{(j)} - \varepsilon_s^{*(j)}\right\rangle = \sum_{s=1}^{t} E_\theta\left(\left(\varepsilon_s^{(j)} - \varepsilon_s^{*(j)}\right)^2 / \mathcal{F}_{s-1}\right) \\
&= \sum_{s=1}^{t} E_\theta\left(\left(\mathcal{E}_s^{(j)}\right)^2 / \mathcal{F}_{s-1}\right)
\end{aligned}$$

où $\mathcal{E}_s$ est défini dans $(L2)$.Par $(L2)$ et A_t diagonale, nous avons

$$\begin{aligned}
0 &= \lim \sum_{s=1}^{t} E_\theta\left(\left\|A_t^{-1}(\theta)\,\mathcal{E}_s(\theta)\right\|^2 / \mathcal{F}_{s-1}\right) = \lim \sum_{s=1}^{t} E_\theta\left[\left(A_t^{-1(jj)}\mathcal{E}_s^{(j)}\right)^2 / \mathcal{F}_{s-1}\right] \\
&= \lim \left(A_t^{-1(jj)}\right)^2 \sum_{s=1}^{t} E_\theta\left[\left(\mathcal{E}_s^{(j)}\right)^2 / \mathcal{F}_{s-1}\right] \\
&= \lim \left(A_t^{-1(jj)}\right)^2 \left\langle M^{(j)}\right\rangle_t \quad \text{en probabilité}
\end{aligned}$$

Maintenant, l'inégalité de Lenglart-Rebolledo (cf.[10],Ch.1,§9) donne, pour tout $k > 0$, $\varepsilon > 0$

$$P^{\theta}\left\{\left(M_t^{(j)}\right)^2 \geq k^2\left(A_t^{-1(jj)}\right)^{-2}\right\} \leq \frac{\varepsilon}{k} + P^{\theta}\left\{\left\langle M^{(j)}\right\rangle_t \geq \varepsilon\left(A_t^{-1(jj)}\right)^{-2}\right\}$$

Par suit

$$P^{\theta}\left\{\left(M_t^{(j)}\right)^2 \geq k^2\left(A_t^{-1(jj)}\right)^{-2}\right\} \longrightarrow 0 \text{ quand } t \longrightarrow \infty$$

c'est à dire

$$A_t^{-1(jj)} M_t^{(j)} \longrightarrow_{t \longrightarrow \infty}^{P^{\theta}} 0$$

Finalement

$$A_t \delta_t = A_t(\theta)\,(\hat{\theta}_t - \hat{\theta}_t^*) \longrightarrow_{t \longrightarrow \infty}^{P^{\theta}} 0$$

■

Remarque 14 *Les résultats dans cette section donnent des conditions suffisantes pour la convergence des suites de la forme $A_t(\theta)\,\Delta_t$. Par conséquent, nous pouvons utiliser ceci pour trouver une condition suffisante à (L1) sous la forme suivante:*

(LL) les élèments de $A_t(\theta)$ sont des processus non-décroissants, $A_t^{(jj)}(\theta) \longrightarrow \infty$ ($A_t(\theta)$ est diagonale), et

$$A_t^{-2}(\theta) \sum_{s=1}^{t} A_s(\theta)\left[\Delta\Gamma_s(\theta) - C_s(\theta, \Delta_{s-1})\right]\Delta_{s-1} \longrightarrow 0$$

Proposition 15 *On a la condition (LL) implique (L1).*

Preuve: Notons

$$\begin{aligned}
\chi_s &= A_s\left[\Delta\Gamma_s(\theta) - C_s(\theta, \Delta_{s-1})\right]\Delta_{s-1} \\
\mathcal{G}_t &= A_t^{-1} \sum_{s=1}^{t}\left[\Delta\Gamma_s(\theta) - C_s(\theta, \Delta_{s-1})\right]\Delta_{s-1} = A_t^{-1} \sum_{s=1}^{t} A_s^{-1}\chi_s
\end{aligned}$$

Nous avons la formule suivante

$$\begin{aligned}
\sum_{s=1}^{t} D_s \Delta c_s &= \sum_{s=1}^{t} D_s \left(c_s - c_{s-1}\right) = D_t c_t + \sum_{s=1}^{t-1} D_s c_s - \sum_{s=1}^{t} D_s c_{s-1} \\
&= D_t c_t - (\sum_{s=1}^{t} D_s c_{s-1} - \sum_{s=1}^{t} D_{s-1} c_{s-1}) \\
&= D_t c_t - \sum_{s=1}^{t} \left(D_s - D_{s-1}\right) c_s \\
&= D_t c_t - \sum_{s=1}^{t} \Delta D_{s+1} c_s \quad \text{avec} \quad c_0 = 0 = D_0 \qquad (3.3.6)
\end{aligned}$$

Nous utilisons cette formule pour $c_s = \sum\limits_{m=1}^{s} \chi_m$ et $D_s = A_s^{-1}$, nous obtenons

$$\sum_{s=1}^{t} A_s^{-1} \chi_s = A_t^{-1} \sum_{s=1}^{t} \chi_s - \sum_{s=1}^{t} \Delta A_s^{-1} \sum_{m=1}^{s-1} \chi_m$$

c'est à dire

$$\mathcal{G}_t = A_t^{-2} \sum_{s=1}^{t} \chi_s - A_t^{-1} \sum_{s=1}^{t} \Delta A_s^{-1} \sum_{m=1}^{s-1} \chi_m$$

Mais

$$\Delta A_s^{-1} = A_s^{-1} - A_{s-1}^{-1} = -A_s^{-1} \left(A_s - A_{s-1}\right) A_{s-1}^{-1} = -\Delta A_s A_s^{-1} A_{s-1}^{-1}$$

Par conséquent

$$\begin{aligned}
\mathcal{G}_t &= A_t^{-2} \sum_{s=1}^{t} \chi_s - A_t^{-1} \sum_{s=1}^{t} \left(-\Delta A_s A_s^{-1} A_{s-1}^{-1}\right) \sum_{m=1}^{s-1} \chi_m \\
&= A_t^{-2} \sum_{s=1}^{t} \chi_s + A_t^{-1} \sum_{s=1}^{t} \Delta A_s \left[A_s^{-1} A_{s-1}^{-1} \sum_{m=1}^{s-1} \chi_m \right]
\end{aligned}$$

Par (LL) nous avons $A_t^{-2} \sum\limits_{s=1}^{t} \chi_s \longrightarrow 0$ et $A_s^{-1} A_{s-1}^{-1} \sum\limits_{m=1}^{s-1} \chi_m \longrightarrow 0$. Par un *Lemme de Toeplitz* nous avons

$$A_t^{-1} \sum_{s=1}^{t} \Delta A_s \left\{ A_s^{-1} A_{s-1}^{-1} \sum_{m=1}^{s-1} \chi_m \right\} \longrightarrow 0$$

Donc $\mathcal{G}_t \longrightarrow 0$ qui est $(L1)$, donc (LL) implique $(L1)$. ∎

Remarque 16 i) *La condition* (E) *est du type ergodique. Elle a lieu si* $\Gamma_t(\theta)$ *est non-aléatoire ou si* $\Gamma_t(\theta)/t \longrightarrow \gamma(\theta)$ *pour une matrice inversible non-aléatoire.*

ii) *La condition* $(L2)$ *est une condition de continuté des fonctions* $\psi_t(\theta)$ *et* $\Gamma_t(\theta)$ *par rapport à* θ.

3.4 Exemples

Cas iid

Considére le cas d'observations iid $X_1, X_2, ...$ de densité $f(\theta, x)$, $\theta \in \mathbb{R}^m$. Supposons que $\psi(\theta, z)$ est une fonction d'estimation: $\int \psi(\theta, z) f(\theta, z) u(dz) = 0$. Nous définissons l'estimateur récursif $\hat{\theta}_t$ par

$$\hat{\theta}_t = \hat{\theta}_{t-1} + \frac{1}{t}\gamma^{-1}(\hat{\theta}_{t-1})\psi(\hat{\theta}_{t-1}, X_t), \qquad t \geq 1 \tag{3.4.1}$$

où $\gamma(\theta)$ est une matrice non-aléatoire, telle que $\gamma^{-1}(\theta)$ existe pour tout $\theta \in \mathbb{R}^m$.

Supposons que

$$j_\psi(\theta) = \int \psi(\theta, z) \psi^T(\theta, z) f(\theta, z) \mu(dz) < \infty$$

et considérons les conditions suivantes:

(I) pour tout $0 < \varepsilon < 1$,

$$\sup_{\varepsilon \leq \|u\| \leq 1/\varepsilon} u^T \gamma^{-1}(\theta + u) \int \psi(\theta + u, x) f(\theta, x) \mu(dx) < 0$$

(II) pour tout $u \in \mathbb{R}^m$,

$$\int \left\|\gamma^{-1}(\theta + u) \psi(\theta + u, x)\right\|^2 f(\theta, x) \mu(dx) \leq K_\theta(1 + \|u\|^2)$$

où $K_\theta > 0$.

(III) γ est continue

(IV)

$$\lim_{u \longrightarrow 0} \int \|\psi(\theta + u, x) - \psi(\theta, x)\|^2 f(\theta, x) \mu(dx) = 0$$

(V)

$$\int \psi\left(\theta+u,x\right)f\left(\theta,x\right)\mu\left(dx\right)=-\gamma\left(\theta+u\right)u+\alpha^{\theta}\left(u\right)$$

où $\alpha^{\theta}\left(u\right)=o(\left\|u\right\|^{1+\varepsilon})$ quand $u\longrightarrow 0$ et $\varepsilon>0$.

Nous avons le corollaire suivant

Corollaire [3.2] Supposons que les conditions $(I),(V)$ sont satisfaites. Alors l'estimateur $\hat{\theta}_t$ est convergent et $t^{\delta}(\hat{\theta}_t-\theta)\longrightarrow 0\ P^{\theta}p.s$, pour tout $0<\delta<1/2$. De plus $\hat{\theta}_t$ est asymptotiquement gaussien:

$$\mathcal{L}\left(t^{1/2}(\hat{\theta}_t-\theta)/P^{\theta}\right)\longrightarrow^{\omega}\mathcal{N}\left(0,\gamma^{-1}\left(\theta\right)j_{\psi}\left(\theta\right)\gamma^{-1}\left(\theta\right)\right)$$

Preuve: Comme $\Gamma_t\left(\theta\right)=t\gamma\left(\theta\right)$ et $b_t\left(\theta,u\right)=b\left(\theta,u\right)=\int\psi\left(\theta+u,z\right)f\left(\theta,z\right)\mu\left(dz\right)$, nous avons du Corollaire [2.1]. Ch 2 que (I) et (II) impliquent $(C1),(C2)$ et $(C3)$ de Théorème 7 que $\hat{\theta}_t\longrightarrow\theta$.

Alors, (II) implique $(B2)$ du Corollaire [3.1]. En effet la condition (II): $\forall u\in\mathbb{R}^m$,

$$\int\left\|\gamma^{-1}\left(\theta+u\right)\psi\left(\theta+u,x\right)\right\|^2 f\left(\theta,x\right)\mu\left(dx\right)\leq K_{\theta}(1+\left\|u\right\|^2)$$

ou encore

$$E_{\theta}\left(\left\|\gamma^{-1}\left(\theta+u\right)\psi\left(\theta+u,x\right)\right\|^2\right)\leq K_{\theta}(1+\left\|u\right\|^2)=O\left(1\right)$$

quand $u\longrightarrow 0$. D'où $(B2)$.La condition (V) implique $(B1)$ du Corollaire [3.1], avec $C_{\theta}=1$. En effet (V):

$$E^{\theta}\left(\theta+u,X_1\right)=\int\psi\left(\theta+u,x\right)f\left(\theta,x\right)\mu\left(dx\right)=-\gamma\left(\theta+u\right)u+\alpha^{\theta}\left(u\right)$$

implique que

$$\begin{aligned}
\left\langle u, \gamma^{-1}(\theta+u) E^{\theta}(\theta+u, X_1)\right\rangle &= \left\langle u, \gamma^{-1}(\theta+u)\left(-\gamma(\theta+u) u+\alpha^{\theta}(u)\right)\right\rangle \\
&= \langle u, -u\rangle + \left\langle u, \gamma^{-1}(\theta+u) \alpha^{\theta}(u)\right\rangle \\
&= -\|u\|^2 + \left\langle u, \gamma^{-1}(\theta+u) o(\|u\|^{1+\varepsilon})\right\rangle \\
&= -\|u\|^2 + o(\|u\|^{1+\varepsilon}) \left\langle u, \gamma^{-1}(\theta+u)\right\rangle \\
&\leq -\frac{1}{2}\|u\|^2 = -\frac{1}{2}\langle u, u\rangle \text{ quand } u \longrightarrow 0
\end{aligned}$$

D'où $(B1)$ avec $C_\theta = 1$.

Par conséquent les conditions $(B1)$ et $(B2)$ du corollaire [3.1] sont satisfaites et donc pour tout $0 < \delta < 1/2, t^{\delta}(\hat{\theta}_t - \theta) = t^{\delta}\Delta_t \longrightarrow 0$ $p.s.$

Vérifions les conditions du Théorème 13 avec $A_t = \sqrt{t}I$. La condition (E) est triviale. Pour vérifier $(L1)$ il suffit de montrer avec $\Gamma_t(\theta) = t\gamma(\theta)$ que

$$\begin{aligned}
& A_t^{-1}(\theta) \sum_{s=1}^{t} \left(\Delta\Gamma_s(\theta) - C_s(\theta, \Delta_{s-1})\right) \Delta_{s-1} \\
={}& \frac{1}{\sqrt{t}} \sum_{s=1}^{t} \left(s - (s-1)\gamma(\theta) - C(\theta, \Delta_{s-1})\right) \Delta_{s-1} \\
={}& \frac{\sqrt{t}}{t} \sum_{s=1}^{t} \left(\gamma(\theta) - C(\theta, \Delta_{s-1})\right) \Delta_{s-1} \\
={}& \frac{1}{t} \sum_{s=1}^{t} \left[\gamma(\theta) - C(\theta, \Delta_{s-1})\right] \sqrt{s}\Delta_{s-1} \to 0
\end{aligned}$$

où par (3.3.2) nous avons pour $u \neq 0$

$$\begin{aligned}
C_t(\theta, u) &= C(\theta, u) = -\Gamma_t(\theta)\Gamma_t^{-1}(\theta+u) b_t(\theta, u) u^T / \|u\|^2 \\
&= -t\gamma(\theta) \frac{1}{t} \gamma^{-1}(\theta+u) b(\theta, u) u^T / \|u\|^2 \\
&= -\gamma(\theta)\gamma^{-1}(\theta+u) \int \psi(\theta+u, z) f(\theta, z) \mu(dz) u^T / \|u\|^2
\end{aligned}$$

et pour $u = 0$, $C_t(\theta, 0) = \Delta\Gamma_t(\theta) = \Gamma_t(\theta) - \Gamma_{t-1}(\theta) = (t - (t-1))\gamma(\theta) = \gamma(\theta)$. La condition

(V) donne

$$\begin{aligned}
& [\gamma(\theta) - C(\theta, \Delta_{s-1})] \sqrt{s}\Delta_{s-1} \\
= & \left[\gamma(\theta) + \gamma(\theta)\gamma^{-1}(\theta + \Delta_{s-1}) \int \psi(\theta + \Delta_{s-1}, z) f(\theta, z) \mu(dz) \Delta_{s-1}^T / \|\Delta_{s-1}\|^2\right] \sqrt{s}\Delta_{s-1} \\
= & \sqrt{s}\left[\gamma(\theta) + \gamma(\theta)\gamma^{-1}(\theta + \Delta_{s-1}) \left(-\gamma(\theta + \Delta_{s-1}) + \alpha^\theta(\Delta_{s-1})\right)\right] \\
= & \sqrt{s}\left[\gamma(\theta) - \gamma(\theta) + \gamma(\theta)\gamma^{-1}(\theta + \Delta_{s-1}) \alpha^\theta(\Delta_{s-1})\right] \\
= & \sqrt{s}\gamma(\theta)\gamma^{-1}(\theta + \Delta_{s-1}) \alpha^\theta(\Delta_{s-1}) \\
= & (\sqrt{s}\|\Delta_{s-1}\|^{1+\varepsilon}) \left(\gamma(\theta)\gamma^{-1}(\theta + \Delta_{s-1}) \alpha^\theta(\Delta_{s-1}) / \|\Delta_{s-1}\|^{1+\varepsilon}\right) \\
= & \sqrt{s}\|\Delta_{s-1}\|^{1+\varepsilon} \delta_s
\end{aligned}$$

où $\delta_s = \gamma(\theta)\gamma^{-1}(\theta + \Delta_{s-1}) \alpha^\theta(\Delta_{s-1}) / \|\Delta_{s-1}\|^{1+\varepsilon}$. On a par (III) et (V) et $\Delta_{s-1} \longrightarrow 0$: $\delta_s \rightarrow_{s \rightarrow \infty} 0$. Par suit puisque

$$\begin{aligned}
[\gamma(\theta) - C(\theta, \Delta_{s-1})] \sqrt{s}\Delta_{s-1} &= \sqrt{s}\|\Delta_{s-1}\|^{1+\varepsilon} \delta_s \\
&= \sqrt{\frac{s}{s-1}} \left((s-1)^{\frac{1}{2(1+\varepsilon)}} \|\Delta_{s-1}\|\right)^{1+\varepsilon} \delta_s \longrightarrow 0
\end{aligned}$$

car pour $\frac{1}{2(1+\varepsilon)} < 1/2$: $(s-1)^{\frac{1}{2(1+\varepsilon)}} \Delta_{s-1} \rightarrow 0$ $p.s.$ Par un Lemme de Toeplitz, nous avons $\frac{1}{t} \sum\limits_{s=1}^{t} [\gamma(\theta) - C(\theta, \Delta_{s-1})] \sqrt{s}\Delta_{s-1} \longrightarrow 0$. D'où la condition $(L1)$.

Verifions la condition $(L2)$. Pour le processus $\mathcal{E}_s(\theta)$ nous avons

$$\begin{aligned}
\|\mathcal{E}_s(\theta)\|^2 &= \left\|\Gamma_s(\theta)\Gamma_s^{-1}(\theta + \Delta_{s-1})(\psi_s(\theta + \Delta_{s-1}) - b_s(\theta, \Delta_{s-1})) - \psi_s(\theta)\right\|^2 \\
&= \left\|\gamma(\theta)\gamma^{-1}(\theta + \Delta_{s-1})(\psi(\theta + \Delta_{s-1}, X_s) - b(\theta, \Delta_{s-1})) - \psi(\theta, X_s)\right\|^2
\end{aligned}$$

en utilisant $(x-y)^2 \leq 2x^2 + 2y^2$, nous obtenons

$$\begin{aligned}
\|\mathcal{E}_s(\theta)\|^2 \leq\ & 2\left\|\gamma(\theta)\gamma^{-1}(\theta + \Delta_{s-1})\psi(\theta + \Delta_{s-1}, X_s) - \psi(\theta, X_s)\right\|^2 \\
& +2\left\|\gamma(\theta)\gamma^{-1}(\theta + \Delta_{s-1}) b(\theta, \Delta_{s-1})\right\|^2
\end{aligned}$$

Par (V), nous avons

$$b\left(\theta,\Delta_{s-1}\right)=\int\psi\left(\theta+\Delta_{s-1},x\right)f\left(\theta,x\right)\mu\left(dx\right)=-\gamma\left(\theta+\Delta_{s-1}\right)\Delta_{s-1}+\alpha^{\theta}\left(\Delta_{s-1}\right)\longrightarrow 0$$

Par (III): $\gamma\left(\theta\right)\gamma^{-1}\left(\theta+\Delta_{s-1}\right)\rightarrow 1$ et par suit par (IV) nous avons $E_{\theta}\left(\left\|\mathcal{E}_{s}\left(\theta\right)\right\|^{2}/\mathcal{F}_{s-1}\right)\longrightarrow 0.$ Donc par un Lemme de Toeplitz

$$\sum_{s=1}^{t}E_{\theta}\left(\left\|A_{t}^{-1}\left(\theta\right)\mathcal{E}_{s}\left(\theta\right)\right\|^{2}/\mathcal{F}_{s-1}\right)=\frac{1}{t}\sum_{s=1}^{t}E_{\theta}\left(\left\|\mathcal{E}_{s}\left(\theta\right)\right\|^{2}/\mathcal{F}_{s-1}\right)\longrightarrow 0$$

Donc $(L2)$ est satisfaite.

Donc les conditions du Théorème 13 sont verifies pour $A_{t}\left(\theta\right)=\sqrt{t}$. La formule du dévoloppement asymptotique (3.1.2) donne pour $\Gamma_{t}\left(\theta\right)=t\gamma\left(\theta\right)$:

$$\begin{aligned}\hat{\theta}_{t}-\theta &= \Gamma_{t}^{-1}\left(\theta\right)\sum_{s=1}^{t}\psi(\theta,X_{s})+o_{p^{\theta}}\left(1\right)\\ \sqrt{t}(\hat{\theta}_{t}-\theta) &= \frac{1}{\sqrt{t}\gamma\left(\theta\right)}\sum_{s=1}^{t}\psi(\theta,X_{s})+o_{p^{\theta}}\left(1\right)\end{aligned}$$

où $o_{p^{\theta}}\left(1\right)\stackrel{P^{\theta}}{\longrightarrow}0$. Par le théorème centrale limite pour les v.a $\psi(\theta,X_{s})$ iid centrées et de matrice de covariance $j_{\psi}\left(\theta\right)$ on aboutit à:

$$\sqrt{t}(\hat{\theta}_{t}-\theta)\Longrightarrow\mathcal{N}(0,\gamma^{-1}\left(\theta\right)j_{\psi}\left(\theta\right)\gamma^{-1}\left(\theta\right))$$

■

Annexe 1

Résultats auxiliaires

Lemme 2 *Soit* $\mathcal{F}_0, \mathcal{F}_1, ...$ *une suite non-décoissante de* σ*-algebre et,* $X_n, \beta_n, \xi_n, \zeta_n \in \mathcal{F}_n$, $n \geq 0$ *de variables aléatoires non-négatives, telles que*

$$E\left(X_n/\mathcal{F}_{n-1}\right) \leq X_{n-1}\left(1+\beta_{n-1}\right)+\xi_{n-1}-\zeta_{n-1}, \qquad n \geq 1$$

Alors

$$\left\{\sum_{i=1}^{\infty} \xi_{i-1} < \infty\right\} \cap \left\{\sum_{i=1}^{\infty} \beta_{i-1} < \infty\right\} \subseteq \{X \longrightarrow\} \cap \left\{\sum_{i=1}^{\infty} \zeta_{i-1} < \infty\right\}, p.s$$

où $\{X \longrightarrow\}$ *dénote* $\lim\limits_{n \longrightarrow \infty} X_n$ *existe et finie.*

Preuve: Pour la preuve de ce lemme cf. [12]. ■

Lemme 3 *Soit* d_n *est une suite non-décroissante des nombres positives telle que* $d_n \longrightarrow \infty$.

Alors,

(a) $\sum\limits_{n=1}^{\infty} \frac{\Delta d_n}{d_n} = \infty,$

(b) $\sum\limits_{n=1}^{\infty} \frac{\Delta d_n}{d_n^{1+\varepsilon}} < \infty, \quad \forall \varepsilon > 0$

Preuve: Pour la preuve de (a) (cf.[17],Lemme 2,§3,Ch. IV).

(b)

$$\begin{aligned} 0 &\leq \sum_{n=1}^{N} \frac{\Delta d_n}{d_n^{1+\varepsilon}} \leq \sum_{n=1}^{N} \int_0^1 \frac{\Delta d_n dt}{\left(d_{n-1} + t\Delta d_n\right)^{1+\varepsilon}} = \sum_{n=1}^{N} \frac{-1}{\varepsilon\left(d_{n-1} + t\Delta d_n\right)^{\varepsilon}} \Big|_0^1 \\ &= \sum_{n=1}^{N} \frac{1}{\varepsilon}\left(\frac{1}{d_{n-1}^{\varepsilon}} - \frac{1}{d_n^{\varepsilon}}\right) = \frac{1}{\varepsilon}\left(\frac{1}{d_0^{\varepsilon}} - \frac{1}{d_N^{\varepsilon}}\right) \longrightarrow \frac{1}{\varepsilon d_0^{\varepsilon}} < \infty \end{aligned}$$

■

Lemme 4 *Supposons que $g \neq 0$ est une fonction paire positive sur $\mathbb{R}$, et $g \downarrow 0$ sur $\mathbb{R}_+$. Supposons aussi que ϕ est une fonction impaire mesurable dans $\mathbb{R}$, telle que $\phi(z) > 0$ pour $z > 0$, et $\int_{\mathbb{R}} |\phi(z-\omega)|\, g(z)dz < \infty$ pour tout $\omega \in \mathbb{R}$.*

Alors pour tout $\omega \neq 0$,

$$\omega \int_{-\infty}^{\infty} \phi(z-\omega)\, g(z)dz < 0 \tag{A.1}$$

De plus si $g(z)$ est continue, alors pour tout $\varepsilon \in \left]0,1\right[$,

$$\sup_{\varepsilon \leq |\omega| \leq 1/\varepsilon} \omega \int_{-\infty}^{\infty} \phi(z-\omega)\, g(z)dz < 0 \tag{A.2}$$

Preuve: Pour la preuve cf.[15] ■

Lemme 5 *Supposons que d_n, c_n, et c sont des variables aléatoires, telles que $d_n > 0$, $c_n \geq 0$, $c > 0$, et $d_n \longrightarrow \infty$ quand $n \longrightarrow \infty$. Alors*

$$\frac{1}{d_n} \sum_{i=1}^{n} c_i \longrightarrow c, \quad \textit{en probabilité}$$

implique que

$$\sum_{n=1}^{\infty} \frac{c_n}{d_n} = \infty \quad p.s. \ , \ \text{et} \quad \sum_{n=1}^{\infty} \frac{c_n}{d_n^2} < \infty \quad p.s.$$

Preuve: Notons, $\xi_n = \frac{1}{d_n} \sum_{i=1}^{n} c_i$. Comme $\xi_n \longrightarrow c$ en probabilité, il s'ensuit qu'il existe une sous suite ξ_{in} de ξ_n, telle que $\xi_{in} \longrightarrow c$ p.s.

Maintenant, supposons que $\sum_{n=1}^{\infty} \frac{c_n}{d_n} < \infty$ dans un ensemble A de probabilité positive Alors il s'ensuit d'aprés le Lemme de Kronecker (cf.[17],Lemme 2,§3,Ch.IV) que $\xi_n \longrightarrow 0$ dans A. Alors il s'ensuit de $\xi_{in} \longrightarrow 0$ dans A, que $c = 0$ dans A. Ce qui contredit que $c > 0$ p.s. Donc $\sum\limits_{n=1}^{\infty} \frac{c_n}{d_n} = \infty$ p.s..

Pour la deuxième implication (cf.[17],Lemme 2,§3,Ch.IV). ■

Annexe 2

Simulation

Nous étudions un processus AR (1) défini par

$$X_t = \theta X_{t-1} + Z_t, \qquad t = 1, 2, ...$$

où $X_0 = 0$, et Z_t sont iid de densité gaussienne $\mathcal{N}(0,1)$. L'estimateur des moindres carrés défini par (1.4.11):

$$\hat{\theta}_t = \hat{\theta}_{t-1} + \frac{1}{I_t}\left(X_t X_{t-1} - X_{t-1}^2 \hat{\theta}_{t-1}\right) \text{ où } I_t = I_{t-1} + X_{t-1}^2$$

La figure 1 représente une réalisation de $\hat{\theta}_t$ pour $\theta = 0.5$ et trois valeurs initiales : $\hat{\theta}_0 = -0.1, 0.1$ et 0.7 .

La taille de l'echantillon est 100. Dans le cas de la loi normale nous remarquons que la méthode des moindres carrées et du maximum du vraissemblance donnent le même estimateur défini par (1.4.11). En général cet estimateur vérifie l'équation (2.3.4):

$$\hat{\theta}_t = \hat{\theta}_{t-1} + \Gamma_t^{-1}(\hat{\theta}_{t-1})\psi_t(X_t - \hat{\theta}_{t-1}X_{t-1})$$

D'aprés (2.3.5), $\Gamma_t = i_g I_t$, $\psi_t(z) = l_t(z) = \frac{\dot{f}^T(z)}{f(z)} = -\frac{\acute{g}(z)}{g(z)}X_{t-1}$, et l'estimateur devient

$$\hat{\theta}_t = \hat{\theta}_{t-1} - I_t^{-1} i_g^{-1} X_{t-1} \frac{g'(z)}{g(z)}$$

où $g(z)$ est la densité gaussienne $\mathcal{N}(0,1)$ de Z_t, et $g'(z)/g(z) = -z$, $i_g = \int \left(\frac{-g'(z)}{g(z)}\right)^2 g(z)dz = 1$.

L'estimateur MV devient

$$\hat{\theta}_t = \hat{\theta}_{t-1} + \frac{1}{I_t}\left(X_t X_{t-1} - X_{t-1}^2 \hat{\theta}_{t-1}\right)$$

qui est aussi l'estimateur MC.

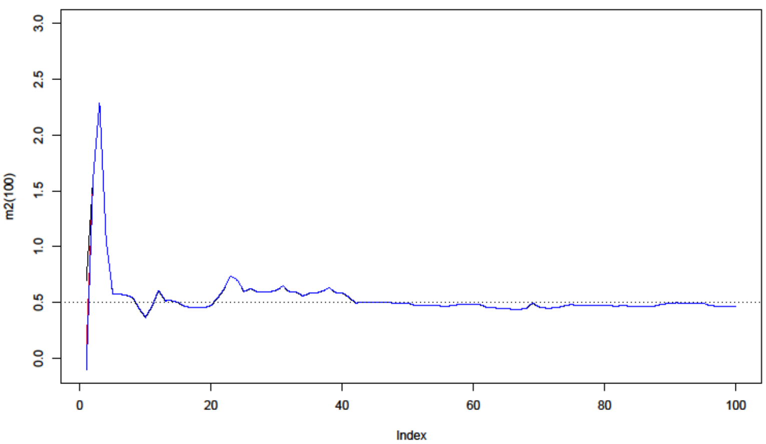

Figure 1.Trajectoire de $\hat{\theta}_t$

La figure 2 représente la realisation de $\hat{\theta}_t$ defini par (2.3.10):

$$\hat{\theta}_t = \hat{\theta}_{t-1} + I_t^{-1}(\hat{\theta}_{t-1})\,(\alpha+1)\,X_{t-1}\frac{X_t - \hat{\theta}_{t-1}X_{t-1}}{\alpha + \left(X_t - \hat{\theta}_{t-1}X_{t-1}\right)^2}, \qquad t \geq 1$$

qui est l EMV avec un bruit blanc de loi de Student à $\alpha = 3$ degrée de liberté,avec $I_t = I_{t-1} + i_g X_{t-1}^2$ où $i_g = \alpha + 1/\alpha + 3$ pour $\theta = 0.5$ et trois valeurs initiales $\hat{\theta}_0 = -0.1$, 0.1 et 0.7 et le nombre d'observation est 40.

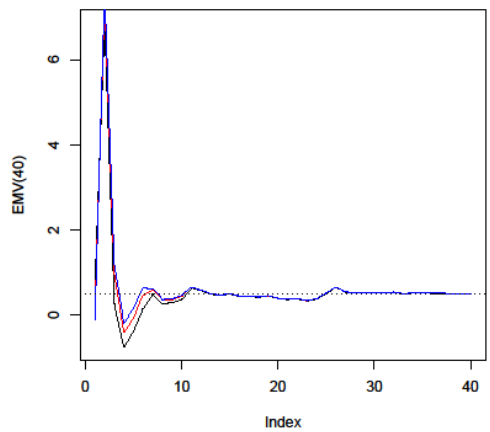

Figure 2. Trajectoire de $\hat{\theta}_t$

La figure 3 représente une realisation de (2.3.10) et (1.4.11) pour un processus AR(1) de bruit blanc de loi de Student à $\alpha = 3$ degrée de liberté et $\theta = 0.4$, $\hat{\theta}_0 = 0.1$, $n = 300$.

Dans les figures 3 et 4 l'EMV est en noir et l'EMC est en rouge.

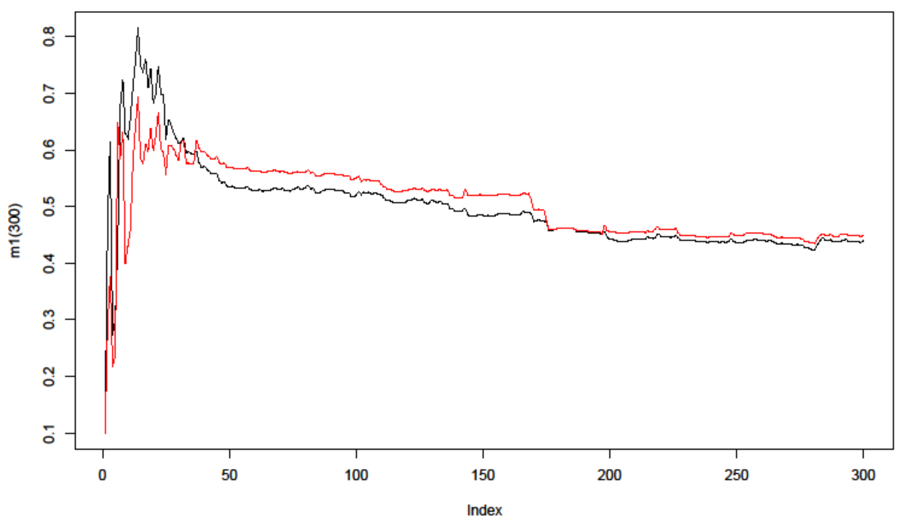

Figure 3. Trajectoires de EMV et EMC

La figure 4.représente l'erreur quadratique des estimateurs définis par (2.3.10) et (1.4.11).

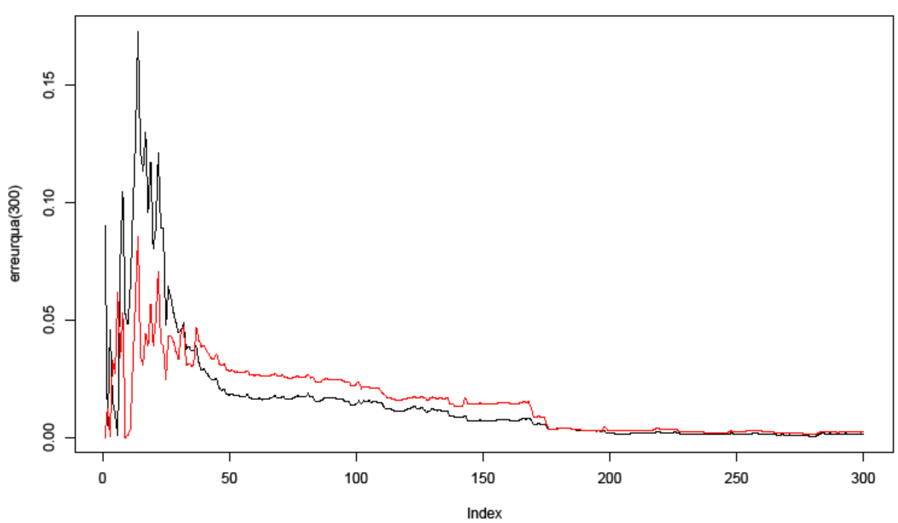

Figure 4. Erreur quadratique de (2.3.10) et (1.4.11).

Conclusion

Dans ce mémoire , nous avons présenté une synthèse de trois articles de Teo Sharia.

Nous avons développés les principaux résultats de convergence et les vitesses de convergence des estimateurs définis de manière recursive avec aussi des exemples.

Nous avons consacré une Annexe pour traiter des simulations numériques sur des estmateurs récursifs pour le paramètre d'un processus AR(1). Les résultats obtenus sont trés satisfaisants.

Bibliographie

[1] Anderson T. W. (1959). On asymptotic distributions of estimates of parameters of stochastic difference equations Ann. Math. Statist. **30**, 676-687.

[2] Campbell K. (1982). Recursive computation of M-estimates for the parameters of a finite autoregressive process. Ann. Statist. **10**, 442-453.

[3] Fabian V. (1978).On asymptotically efficient recursive estimation. Ann.Statist. **6**, 854- 867.

[4] Feigin P. D. (1981). Conditional exponential families and a representation theorem for asymptotic inference. Ann. Statist. **9**. 597-603.

[5] Gladyshev E. G. (1965). On stochastic approximation. Theory Probab. Appl. **10**, 297-300.

[6] Khas'minskii R. Z and Nevelson M. B. (1972). Stochastic Approximation and Recursive Estimation. Nauka, Moscow.

[7] Lazrieva N. Sharia T. and Toronjadze T. (1997). The Robbins-Monro type stochastic differential equations. **I.** Convergence of solutions. Stochastics and Stochastic Reports. **61**, 67-87.

[8] Lazrieva N. Sharia T.and Toronjadze T. (2003). The Robbins-Monro type stochastic differential equations. **II**. Asymptotic behaviour of solutions.Stochastics and Stochastic Reports. **75,**153-180.

[9] Leonov S. L and (1988). On recurrent estimation of autoregression parameters. Avtomatika Telemekhanika. **5**, 105-116.

[10] Liptser R. S and Shiryayev A. N. (1989). Theory of Martingales.Kluwer, Dordrecht.

[11] Robbins H. Monro S. (1951). A stochastic approximation method. Ann. Statist. **22**, 400-407.

[12] Robbins H and Siegmund D. (1971). A convergence theorem for nonnegative almost supermartingales and some applications. Optimizing Methods in Statistic, ed. J. S. Rustagi Academic Press, New York. 233-257.

[13] Sharia, T. (1998). On the recursive parameter estimation for the istical model. Stochastic Processes Appl. **73,2,**151-172.

[14] Sharia, T. (2003). On recursive parametric estimation theory. Technical Report, Royal Holloway, University of London (RHUL-MA).

[15] Sharia, T. (2007). Recursive parameter estimation: convergence. Statistical Inference for Stochastic Processes (to appear).

[16] Sharia, T. (2007). Rate of convergence in recursive parameter estimation procedures. Georgian Mathematical Journal, **14**,761-776.

[17] Shiryayev, A.N.(1984). Probability .Springer-Verlag, New York.

[18] White, J.S.(1958).The limiting distribution of the serial correlation coefficient in the explosive case. Ann. Math. Stat. **29**,1188-1197.