Processus autorégressif à bruits gaussiens stationnaires.

Marius Soltane

Laboratoire Manceau de Mathématiques, Le Mans Université, Avenue O. Messiaen, 72085 Le Mans cedex 9, France.

Les processus autorégressifs sont très utilisés pour la modélisation dans divers domaines tels que l'économétrie ou bien l'écologie. Dans le cas classique où le processus est perturbé par un bruit blanc fort, les propriétés asymptotiques de l'estimateur des moindres carrés (LSE en abrégé) sont bien connues et en particulier, ce dernier est fortement consistant pour le paramètre autorégressif.

Lorsque les perturbations ne sont plus indépendantes, le LSE est perturbé et l'estimation n'est plus consistance pour le paramètre d'intérêt. Un récent travail [2] s'intéresse aux propriétés asymptotiques de l'estimateur du maximum de vraisemblance (MLE) lorsque les perturbations sont stationnaires et gaussiennes. La construction de l'estimateur résulte de la transformation linéaire du modèle observé développé dans [3].

Dans un récent travail [9], nous nous intéressons dans un premier temps au raffinement des propriétés établit dans [2] en obtenant la consistance forte du MLE pour le paramètre autorégressif ainsi qu'une loi du logarithme itérée permettant d'obtenir une estimation sur la vitesse de convergence de l'estimateur considéré. Dans un second temps nous cherchons la normalité locale asymptotique pour le ratio de vraisemblance (LAN en abrégé). Cette propriété permet de définir une notion d'efficacité asymptotique pour l'expérience considérée ainsi que la construction d'une procédure optimale pour tester la significativité du paramètre autorégressif.

Mots clés : processus autorégressif, contrôle ergodique, propriété LAN, efficacité asymptotique, test AUMPI.

An Autoregressive Process with Stationary Gaussian Noise.

Classical autoregressive processes driven by strong white noise were introduced by Box-Jenkins. Currently, models using autoregressive processes with dependant perturbations are widely used in various fields, particularly in econometrics and finance. The asymptotic behaviour of the least square estimator (LSE) is generally degraded for this type of process, and is not consistent for the autoregressive parameter. Some asymptotic properties of the maximum likelihood estimator (MLE) in that model were later studied in [2]. The present study is concerned with the the *a.s.* properties and the rate of convergence of the estimator, which have not been previously obtained. We also address the questions of the asymptotic efficiency for the MLE and the optimality of the significance test of the parameter driving the autoregressive dynamics.

Key words : Autoregressive process, Asymptotic efficiency, AUMPI test, ergodic control, LAN property.

1. Version longue :

Comme indiqué dans l'introduction, la présence de mémoire dans le processus de nuisance venant perturbé notre modèle autorégressif induit la non-consistance du LSE envers le paramétre autorégressif. Un tel résultat est illustré dans [1] et [8] où les auteurs considérent un processus autorgéressif perturbé par un processus $AR(1)$ (autorégressif d'ordre 1).
Nous nous intéressons dans cet exposé au processus (X_t) indexé sur $\mathbb{N}$ et vérifiant pour tout $t \in \mathbb{N}$,

$$X_n = \sum_{i=1}^{p} \theta_i X_{n-i} + \xi_n \tag{1.1}$$

où le processus de nuissance (ζ_n) est un processus gaussien centré et stationnaire.
Le modèle introduit ci-dessus fut étudié dans [2] où les auteurs considérent un filtre linéaire afin de blanchir le processus de nuissance (la transformation en question est présentée de manière détailée dans [3] et est assez technique). Ils transfèrent alors la problème d'estimation de θ du modèle (1.1) dans un modèle obtenu via le filtre linéaire et dans lequel la vraisemblance est connue grâce à l'indépendance des perburbations. Les auteurs obtiennent alors dans [2] la consistance faible du MLE envers le paramètre autorégressif ainsi que sa normalité asymptotique (ce dernier résultat sera mentionné ultérieurement). Plusieurs questions restent néanmoins en suspens à l'issue de cette étude : Est t-il possible d'obtenir la consistance forte du MLE envers le paramètre autorégressif comme cela est le cas pour le LSE dans un $AR(p)$ classique ? Peut t-on obtenir une estimation presque sûre de la vitesse de convergence du MLE ? Peut t-on construire une procédure optimale pour tester la significativité du paramètre autorégressif ?

L'objectif de cet exposé est de répondre aux problématiques ci-dessus. Le second modèle obtenu dans [2] permet d'écrire de manière explicite la fonction de vraisemblance et d'en déduire l'expression du MLE. Malheureusement le nouveau modèle n'est pas ergodigue qui est une hypothèse classique vérifiée dans un $AR(p)$ classique stable et permettant d'obtenir les résultats usuels et nottament ceux invoquant la convergence presque sûre (*p.s.* en abrégée). Nous nous intéressons donc dans un premier temps aux propriétés asymptotiques preque sûre du MLE en utilisant un contrôle ergodique pour obtenir les résultats sur la consistance forte et les vitesses de convergences *p.s.* (voir [9] pour plus de détails).
Dans un second temps, nous nous intéressons à l'obtention de la propriété LAN qui permet de définir une notion d'efficacité asymptotique pour les estimateurs (voir [7] pour plus de détails). La propriété LAN à déjà été obtenue pour une classe de processus gaussiens stationnaires dans [6] où les conditions portent sur la densité spectrale. Dans notre cas, la structure autorégressive nous permet directement de l'obtenir. Le second intérêt de la propriété LAN réside dans l'existence de tests optimaux pour la pamètre d'intérêt (qui est le pamètre autorégressif dans notre cas). Les propriétés des tests en questions furent traitées dans [5] sous réserve que les matrices de taux intervenant dans la propriété LAN soient diagonales. Nous

terminerons donc cet exposé par la construction d'un test (optimal dans un sens définit ultérieurement).

Le modèle en question peut se reformuler sous forme vectorielle de la façon suivante ($*$ désignant la transposition) :

$$Y_n = A_0 Y_{n-1} + b\xi_n, \tag{1.2}$$

avec

$$A_0 = \begin{pmatrix} \theta_1 & \theta_2 & \dots & \dots & \theta_p \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & 1 & 0 \end{pmatrix}, \; Y_n = (X_n, X_{n-1}, \dots, X_{n-p+1})^*$$

et $b = (1, 0_{1\times(p-1)})^*$. Nous retenons les hypothèses suivantes :

- (H_1) $\rho(A_0) < 1$, où ρ désigne le rayon spectral de A_0. L'espace paramétrique est donc $\Theta = \{\theta \in \mathbb{R}^p | \rho(A_0) < 1\}$.
- (H_2) La fonction d'autocovariance du processus de nuissance que nous notons r vérifie $r(n) = O(n^\alpha)$ où $\alpha > 0$.
- (H_3) Notons (β_n) la fonction d'autocorrélation partielle de (ξ_n), nous supposons que $\beta_n^2 = O(\frac{1}{n^\alpha})$ avec $\alpha > 1$.

Nous avons le premier résultat sur la consistance du MLE pour le paramètre autorégressif.

Théorème 1.1. *Supposons que (H_1), (H_2) et (H_3) sont satisfaites, alors,*

$$\widehat{\theta}_n \xrightarrow[n\to\infty]{p.s} \theta$$

Le second résultat donne la loi forte quadratique pour le MLE ainsi qu'une loi logarithme itéré qui sont également établit dans [1] et [8]. Le résultat est démontré en utilisant la loi forte quadratique pour les martingales (voir [4]).

Théorème 1.2. *Toujours sous (H_1), (H_2) et (H_3),*

$$\frac{1}{\log n} \sum_{k=1}^{n} (\widehat{\theta}_k - \theta)(\widehat{\theta}_k - \theta)^* \xrightarrow[n\to\infty]{p.s.} I(\theta)^{-1}. \tag{1.3}$$

où la matrice $I(\theta)$ est régulière et solution de l'équation de Lyapunov

$$I(\theta) = A_0^* I(\theta) A_0 + bb^*.$$

Pour conclure sur les propriétés presque sûre du MLE nous avons sous (H_1), (H_2) et (H_3), la proposition suivante

Proposition 1.1. *Pour tout $v \in \mathbb{R}^p$,*

$$\begin{aligned} \limsup_{n\to\infty} \left(\frac{n}{2\log\log n}\right)^{\frac{1}{2}} v^*(\widehat{\theta}_n - \theta) &= -\liminf_{n\to\infty} \left(\frac{n}{2\log\log n}\right)^{\frac{1}{2}} v^*(\widehat{\theta}_n - \theta) \\ &= (v^* I(\theta)^{-1} v)^{\frac{1}{2}} \; p.s. \end{aligned}$$

En conséquence,

$$\left\| \widehat{\theta}_n - \theta \right\|^2 = O\left(\frac{\log\log n}{n}\right) \; p.s. \tag{1.4}$$

Pour plus de clarté envers le lecteur, nous rappelons un certains de nombres de définitions et propriétés relatives à la propriété LAN.

Définition 1.1. *Une famille de mesures de probabilités $\mathbb{P}_\theta^{(n)}$ est dite LAN en $\theta_0 \in M \subset \mathbb{R}^d$ lorsque les condtions suivantes sur la ration de vraisemblance sont satisfaites :*

$$L_n(u) = \frac{d\mathbb{P}^{(n)}_{\theta_0+\phi_n(\theta_0)u}}{d\mathbb{P}^{(n)}_{\theta_0}}, \tag{1.5}$$

$$L_n(u) = exp\left(\langle u, Z_n(\theta_0)\rangle - \frac{1}{2}\langle u, J(\theta_0)u\rangle + R_n(\theta_0, u)\right), \tag{1.6}$$

où

$$Z_n(\theta_0) \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}(0, J(\theta_0)), \tag{1.7}$$

et

$$R_n(\theta_0, u) \xrightarrow[n\to\infty]{\mathcal{L}} 0 \tag{1.8}$$

sous $\mathbb{P}^{(n)}_{\theta_0}$.

La suite $(\phi_n(\theta_0))$ satisfait quand à elle

$$\phi_n(\theta_0) \xrightarrow[n\to\infty]{} 0. \tag{1.9}$$

Dans cette définition $u \in K \subset \mathbb{R}^d$, $\phi_n(\theta_0)$ sont des matrices régulières et il en est de même pour $J(\theta_0)$ qui est une matrice de taille $d \times d$.

Le théorème suivant nous donne la propriété LAN dans notre modèle où les matrices de taux sont diagonales et en $\frac{1}{\sqrt{n}}$.

Théorème 1.3. *Sous les notations de la définition précédente, si $\phi_n(\theta_0) = \frac{1}{\sqrt{n}} Id_p$, alors*

$$\log\left(\frac{\mathcal{L}(\theta_0 + \phi_n(\theta_0)u, X^{(n)})}{\mathcal{L}(\theta_0, X^{(n)})}\right) = \langle u, \frac{M_n}{\sqrt{n}}\rangle - \frac{1}{2}\langle u, I(\theta_0)u\rangle + R_n(\theta_0, u), \tag{1.10}$$

où $(\frac{M_n}{\sqrt{n}})$ satisfait condition (1.7) sous $\mathbb{P}^{(n)}_{\theta_0}$ avec $J(\theta_0) = I(\theta_0)$, et $R_n(\theta_0, u)$ satisfait (1.8) sous $\mathbb{P}^{(n)}_{\theta_0}$. Dans le présent théorème, $u \in B(0; R)$ pour tout $R > 0$. Ici $\mathcal{L}\left(\theta_0 + \phi_n(\theta_0)u, X^{(n)}\right)$ désigne la vraisemblance évaluée en $\theta_0 + \phi_n(\theta_0)u$.

Nous rappelons un résultat démontré dans [7]. Ce résultat permet de définir une notion d'efficacité aymptotique dans les expériences LAN lorsque la matrice de covariance est régulière.

Théorème 1.4. *Supposons que la famille de mesures $\left\{\mathbb{P}^{(n)}_\theta, \theta \in M \subset \mathbb{R}^d\right\}$ est LAN en θ_0. Alors pour tout $\delta > 0$,*

$$(1.11)\quad \liminf_{n\to\infty} \sup_{\|\phi_n(\theta_0)^{-1}(\theta-\theta_0)\|\leqslant\delta} \mathbb{E}_{\theta_0}^{(n)}\left(f\left(\phi_n(\theta_0)^{-1}(\widehat{\theta}_n-\theta)\right)\right) \geqslant \int_{\mathbb{R}^m} f\left(J(\theta_0)^{-\frac{1}{2}}x\right)\Phi_d(x)\,\mathrm{d}x,$$

pour tout estimateur $\widehat{\theta}_n$ et pour toute fonction coût f telle que f est continue, symétrique, quasi-convexe et $f(z)exp(-\frac{\|z\|^2}{2}) \to 0$ lorsque $\|z\| \to \infty$. Ici, Φ_d est la densité de la loi normale standard (centrée et réduite) d-dimensionnelle.

Nous souhaitons maintenant construire une procédure optimale pour tester la significativité du paramètre autorégressif. Pour plus de clarté, nous rappelons des définitions ainsi que des résultats formulés dans [5]. Nous supposons dans tout ce qui suit que la famille de mesures $\mathbb{P}_{\theta}^{(n)}$ est LAN en θ_0, et nous construisons enfin une procédure optimale pour tester $\theta = \theta_0$ contre $\theta \neq \theta_0$.

Définition 1.2. *Un test ϕ_n^1 est dit $AUMP(\alpha)$ (Asymptotiquement uniformément le plus puissant de niveau α) si*

$$(1.12)\qquad \limsup_{n\to\infty} \mathbb{E}_{\theta_0}^{(n)}(\phi_n^1) \leqslant \alpha,$$

et pour tout test ϕ_n^2 de niveau asymptotique α,

$$(1.13)\qquad \limsup_{n\to\infty} \mathbb{E}_{\theta_0+\phi_n(\theta_0)u}^{(n)}(\phi_n^2) \leqslant \liminf_{n\to\infty} \mathbb{E}_{\theta_0+\phi_n(\theta_0)u}^{(n)}(\phi_n^1).$$

Remarque 1.1. *Nous énonçons un Lemme démontré dans* [5] *pour formaliser la définition suivante. Nous formulons ce Lemme dans le cadre de notre modèle (i.e sans paramètres de nuissances), car dans notre cas, ils sont calculés (via l'algorithme de Durbin–Levinson).*

Lemme 1.1. *Sous les notations de la définition 1.1, pour tout test ϕ_n^1 et toute sous-suite n', nous pouvons extraire une sous-suite n'' de n' et constuire un test ϕ de $\mathbb{R}^p$ dans $[0;1]$ tels que $u \in K$,*

$$(1.14)\qquad \lim_{n''\to\infty} \mathbb{E}_{\theta_0+\phi_{n''}(\theta_0)u}^{(n'')}(\phi_{n''}^1) = \int_{\mathbb{R}^d} \phi(x)\Phi_d(x-J(\theta_0)u)\,\mathrm{d}x,$$

où Φ_d est définit comme dans le Théorème 1.4.

Nous allons maintenant donner un principe d'invariance par rotation pour les tests.

Définition 1.3. *Un test est ϕ_n^1 est AUMPI(α) si les conditions de la Définition 1.2 sont vérifiées, et pour tout sous suite n', le test correspondant ϕ (obtenu par le Lemme 1.1) satisfait $\phi(Ru) = \phi(u)$ pour toute rotation de $\mathbb{R}^d$ dans $\mathbb{R}^d$.*

Pour finir, nous explicitons un test AUMPI pour tester la significativité du paramètre autorégressif dans notre modèle.

Théorème 1.5. *Le test*

$$(1.15)\qquad \widetilde{\phi}_n = \mathbb{1}_{\left\{2\log\left(\frac{\mathcal{L}\left(\widehat{\theta}_n, X^{(n)}\right)}{\mathcal{L}\left(\theta, X^{(n)}\right)}\right)\geqslant C_\alpha\right\}}$$

est AUMPI(α)pour tester $\theta = \theta_0$ contre $\theta \neq \theta_0$, où C_α est le quantile d'ordre α du χ_p^2. Ici $\mathcal{L}\left(\widehat{\theta}_n, X^{(n)}\right)$ désigne la vraisemblance évaluée en $\widehat{\theta}_n$.

Références

[1] B. Bercu and F. Proïa. A sharp analysis on the asymptotic behavior of the Durbin-Watson statistic for the first-order autoregressive process. *ESAIM : Probability and Statistics*, 17 :500–530, 2013.

[2] A. Brouste, C. Cai, and M. Kleptsyna. Asymptotic properties of the MLE for the autoregressive process coefficients under stationary Gaussian noise. *Mathematical Methods of Statistics*, 23(2) :103–115, Apr 2014.

[3] A. Brouste and M. Kleptsyna. Kalman type filter under stationary noises. *Systems and Control Letters*, 61(12) :1229–1234, 2012.

[4] F. Chaabane and F. Maaouia. Théorèmes limites avec poids pour les martingales vectorielles. *ESAIM : Probability and Statistics*, 4 :137—-189, 2000.

[5] S. Choi, W. J. Hall, and A. Schick. Asymptotically uniformly most powerful tests in parametric and semiparametric models. *Ann. Statist.*, 24(2) :841–861, 04 1996.

[6] S. Cohen, F. Gamboa, C. Lacaux, and J-M. Loubes. LAN property for some fractional type Brownian motion. *ALEA : Latin American Journal of Probability and Mathematical Statistics*, 10(1) :91–106, 2013.

[7] I. A. Ibragimov and R. Z. Khasminskiĭ . Statistical estimation : Asymtotic theory. *Springer-Verlag*, 16, 1981.

[8] F. Proïa. Further results on the h-test of durbin for stable autoregressive processes. *Journal of Multivariate Analysis*, 118 :77 – 101, 2013.

[9] Marius Soltane. Asymptotic Efficiency in the Autoregressive Process driven by a Stationary Gaussian Noise. Working paper or preprint, October 2018.

E-mail address: `marius.soltane.etu@univ-lemans.fr`