Metrika (1995) 42:331–339

Metrika

A Proof of Asymptotic Normality for some VARX Models

MOHAMED BOUTAHAR AND CLAUDE DENIAU

Dépt. de Mathématiques, Faculté des Sciences, Case 901, 163 Avenue de Luminy, 13288 Marseille Cedex 9, France and GREQAM, Vieille Charité, 13002 Marseille, France

Abstract: Here we present a proof of the asymptotic normality of least squares estimates for stable multivariate autoregressive models excited by a deterministic second order input signal.

Key Words: Conditional Lindeberg condition, least squares estimates, martingale difference, persistent excitation, stable autoregressive model, spectral measure.

1 Introduction

We establish the asymptotic normality of the ordinary least squares estimator (O.L.S.E.) of the parameter θ for the stable vectorial autoregressive model excited by a deterministic input signal (denoted by $VARX_d(p, s)$). There is a large number of similar works in the purely autoregressive case, for example (Chan (1988), Dickey & Fuller (1979), adn Touati (1990)). For ARX scalar models there are also some very interesting results: Crowder (1980) obtains the asymptotic normality without homoscedasticity of the model noise i.e.: $E(\varepsilon_n^2/\mathscr{F}_{n-1}) = \sigma_n^2$; Lai & Wei (1982) consider a multiple regression model with a convergent sequence σ_n^2, i.e.: $\lim_{n\to\infty} \sigma_n^2 = \sigma^2$ and obtain the same results. In a more general context (i.e. the model noise follows an $ARMA$ equation), Reinsel (1979) proves the asymptotic normality in a $VARX$ model.

The motivation of this paper is to give, under a slightly stronger assumption than in Lai & Wei (1982): i) a more simple proof of asymptotic normality; and ii) explicitly the covariance matrix of the limit distribution, which is given in the frequency domain.

The studied model is defined by:

$$\left(\mathbf{I}_d + \sum_{i=1}^{p} \mathbf{A}_i z^i\right) Y_n = \left(\sum_{i=1}^{s} \mathbf{B}_i z^i\right) U_n + \varepsilon_n , \qquad n \in \mathscr{N}^* , \tag{1}$$

z is the backward shift operator (i.e: $zX_n = X_{n-1}$), $U_n \in \mathscr{R}^r$, Y_n and ε_n (vectors of $\mathscr{R}^d$) are respectively the input signal, the observable output and the unobservable random noise at stage n; the matrices $(\mathbf{A}_j)_{1\le j\le p}$ and $(\mathbf{B}_j)_{1\le j\le s}$ are (d, d) an (d, r) respectively.

The associated regression model is:

$$Y_n = \theta\Phi_{n-1} + \varepsilon_n \ , \tag{2}$$

where

$$\theta = (-\mathbf{A}_1, \ldots, -\mathbf{A}_p, \mathbf{B}_1, \ldots, \mathbf{B}_s) \ ;$$

and

$$\Phi_n = (Y_n', \ldots, Y_{n-p+1}', U_n', \ldots, U_{n-s+1}')' \ .$$

We recall that the O.L.S.E. estimator of θ, denoted by θ_n, is a solution of the following equation:

$$\mathbf{P}_n\theta_n' = \sum_{k=1}^{n} \Phi_{k-1} Y_k' \ , \tag{3}$$

where θ_n' is the transpose of θ_n, $\mathbf{P}_n = \sum_{k=1}^{n} \Phi_{k-1}\Phi_{k-1}'$.

Taking (2) and (3) into account, the estimation error satisfies:

$$\mathbf{P}_n(\theta_n - \theta)' = \sum_{k=1}^{n} \Phi_{k-1}\varepsilon_k' \tag{4}$$

The regression vector (Φ_n) can be computed recursively by:

$$\Phi_n = \mathscr{A}\Phi_{n-1} + e_n$$
$$\text{with} \quad e_n = (\varepsilon_n', 0, \ldots, 0, U_n', 0, \ldots, 0)' \ , \qquad \mathscr{A} = \begin{pmatrix} \mathbf{A}_c & \mathbf{B} \\ \mathbf{O} & \mathbf{K} \end{pmatrix} , \tag{5}$$

where $\mathbf{A}_c$ is the companion matrix of $A(z) = \mathbf{I}_d + \sum_1^p \mathbf{A}_i z^i$,

$$\mathbf{B} = \begin{bmatrix} \mathbf{B}_1 & \mathbf{B}_2 & \cdots & \cdots & \mathbf{B}_s \\ \mathbf{O} & \cdots & \cdots & \cdots & \mathbf{O} \\ \vdots & & & & \vdots \\ \mathbf{O} & \cdots & \cdots & \cdots & \mathbf{O} \end{bmatrix} , \quad \mathbf{K} = \begin{bmatrix} \mathbf{O} & \cdots & \cdots & \cdots & \cdots & \mathbf{O} \\ \mathbf{O} & \mathbf{I}_r & \mathbf{O} & \cdots & \cdots & \mathbf{O} \\ \mathbf{O} & \mathbf{O} & \mathbf{I}_r & \mathbf{O} & \cdots & \mathbf{O} \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ \mathbf{O} & \cdots & \cdots & \mathbf{O} & \mathbf{I}_r & \mathbf{O} \end{bmatrix} .$$

2 Model Assumptions

($\mathbf{H}_1$): The model (1) is *stable*, i.e: the roots of $\det(A(z))$ are strictly out-side the unit disk of $\mathbb{C}$.
($\mathbf{H}_2$): The noise (ε_n) is a *martingale difference* sequence with respect to an increasing sequence of σ-fields $\mathbb{F} = (\mathscr{F}_n)$, (i.e: ε_n is $\mathscr{F}_n$-measurable and $E(\varepsilon_n/\mathscr{F}_{n-1}) = 0$ *a.s.*, for every n) such that:
a. There exists $\alpha > 2$, such that:

$$\sup_n E(\|\varepsilon_n\|^\alpha/\mathscr{F}_{n-1}) < \infty \ , \qquad a.s. \ ,$$

b. For all n, $E(\varepsilon_n \varepsilon_n'/\mathscr{F}_{n-1}) = \Gamma_\varepsilon$ *a.s.*
($\mathbf{H}_3$): The input signal (U_n) is deterministic and has an *empirical second moment* (called here covariance), i.e:

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n} U_k U'_{k+l} = \Gamma(l) \ , \qquad l \in \mathbb{Z} \ .$$

Remarks and Discussions:

i) With regards to the noise: The assumptions upon (ε_n) can be replaced by: (ε_n) is a sequence of *independent and identically distributed* (i.i.d.) random variables with covariance matrix Γ_ε.
ii) The assumption ($\mathbf{H}_3$) is satisfied by the two following signals:

a)

$$u_n = \begin{cases} 0 & n \neq 2^k \\ \sqrt{k} & n = 2^k \ , \end{cases} \qquad k \in \mathcal{N} \ , \qquad n \in \mathcal{N}^* \ ;$$

which is unbounded and has an empirical mean.
b) $u_n = (-1)^k, n \in [2^k, 2^{k+1}[, k \in \mathcal{N}, n \in \mathcal{N}^*$; which is bounded and doesn't have an empirical mean.

iii) The model (5) is not an autoregressive model; indeed (e_n) is not a martingale difference sequence and then the results obtained in Boutahar (1991) for random input signal cannot be applied.

3 Asymptotic Normality of the O.L.S.E.

The main result of this section is given by the following theorem:

Theorem: Under the assumptions $(\mathbf{H}_i)$, $i = 1, 2, 3$, *the O.L.S.E.* θ_n *satisfies*:

$$\frac{1}{\sqrt{n}}\mathbf{P}_n(\theta_n - \theta)' \overset{\mathscr{L}}{\to} T \ , \tag{6}$$

$T \sim N(0, \Gamma_\varepsilon \otimes \mathbf{P})$ *a gaussian matrix with zero mean and covariance* $\Gamma_\varepsilon \otimes \mathbf{P}$, *where*

$$\mathbf{P} = \int_{-\pi}^{\pi} L(e^{i\omega})d\xi_e(\omega)L(e^{-i\omega})'$$

and $L(z) = (\mathbf{I}_{dp+rs} - \mathscr{A}z)^{-1}$, $\xi_e(.)$ *is the spectral measure of* (e_n) ,

($\otimes$ *denotes the tensor product*).

Proof: By (4):

$$\frac{1}{\sqrt{n}}\mathbf{P}_n(\theta_n - \theta)' = \frac{1}{\sqrt{n}}T_n \tag{7}$$

where $T_n = \sum_{k=1}^{n} \Phi_{k-1}\varepsilon_k'$.
(T_n) is a $\mathbb{F}$-martingale with conditional variance:

$$\langle T_n \rangle \triangleq \sum_{k=1}^{n} E(\Delta T_k \Delta T_k' / \mathscr{F}_{k-1})$$

where $\Delta T_k = T_k - T_{k-1} = \Phi_{k-1}\varepsilon_k'$; since Φ_k is $\mathscr{F}_k$-measurable and $E(\varepsilon_k\varepsilon_k'/\mathscr{F}_{k-1}) = \Gamma_\varepsilon$, we deduce that

$$\langle T_n \rangle = \Gamma_\varepsilon \otimes \sum_{k=1}^{n} \Phi_{k-1}\Phi_{k-1}'$$

$$= \Gamma_3 \otimes \mathbf{P}_n \ .$$

By (5) and the stability assumption on the model we get $\Phi_k = L(z)e_k$. Applying the strong law of large number for martingales, we obtain:

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n} e_k e'_{k+l} = \operatorname{diag}(\delta_0^l \Gamma_\varepsilon, \mathbf{O}, \ldots, \mathbf{O}, \Gamma(l), \mathbf{O}, \ldots, \mathbf{O}) \quad a.s.$$

(δ_0^l is the Kronecker symbol); consequently the spectral measure of (e_n) exists (cf. Rauzy (1976)).

Hence

$$\frac{1}{n}\mathbf{P}_n = \frac{1}{n}\sum_{0}^{n-1} L(z)e_k e'_k L(z)' \to \mathbf{P} \tag{8}$$

and then

$$\frac{1}{n}\langle T_n \rangle \xrightarrow{\mathbb{P}} \Gamma_3 \otimes \mathbf{P} \tag{9}$$

It remains to show that $(T_n, \mathbb{F})$ satisfies the conditional Lindeberg condition:

$$\mathbf{V}_n \triangleq \frac{1}{n}\sum_{k=1}^{n} E(\|\Phi_{k-1}\varepsilon'_k\|^2 \mathbf{1}_{\{\|\Phi_{k-1}\varepsilon'_k\| > \delta\sqrt{n}\}}/\mathscr{F}_{k-1}) \xrightarrow{\mathbb{P}} 0 \ , \qquad \forall\delta > 0 \ ;$$

Now, for all random vector X and all σ-field $\mathscr{F}$, we have:

$$\forall\alpha' > 0 \ , \qquad \forall\delta > 0\colon E(\|X\|^2 \mathbf{1}_{\{\|X\| > \delta\}}/\mathscr{F}) \le \frac{1}{\delta^{\alpha'}} E(\|X\|^{2+\alpha'}/\mathscr{F}) \ ,$$

consequently

$$\begin{aligned}
\mathbf{V}_n &= \frac{1}{n}\sum_{k=1}^{n} E(\|\Phi_{k-1}\varepsilon'_k\|^2 \mathbf{1}_{\{\|\Phi_{k-1}\varepsilon'_k\| > \delta\}}/\mathscr{F}_{k-1}) \\
&= \sum_{k=1}^{n} E\left(\left(\frac{1}{\sqrt{n}}\|\Phi_{k-1}\|\,|\varepsilon_k|\right)^2 \mathbf{1}_{\{(1/\sqrt{n})\|\Phi_{k-1}\|\,|\varepsilon_k| > \delta\}}/\mathscr{F}_{k-1}\right) \\
&\le \frac{1}{\delta^{\alpha'}}\sum_{k=1}^{n} E\left(\left(\frac{1}{\sqrt{n}}\|\Phi_{k-1}\|\,|\varepsilon_k|\right)^{\alpha'+2} \Big/ \mathscr{F}_{k-1}\right) \ ,
\end{aligned}$$

if we choose $\alpha' = \alpha - 2$, where α is given by assumptions $(\mathbf{H}_2).a$, then

$$\mathbf{V}_n \leq K \frac{1}{n^{(2+\alpha')/2}} \sum_1^n \|\Phi_{k-1}\|^{\alpha'+2} ,$$

$$\leq K \left\{ \frac{1}{\sqrt{n}} \max_{1 \leq k \leq n} \|\Phi_{k-1}\| \right\}^{\alpha'} \frac{1}{n} \sum_1^n \|\Phi_{k-1}\|^2 ,$$

By (8) we have

$$\frac{1}{n} \sum_1^n \|\Phi_{k-1}\|^2 \overset{a.s}{\to} trace(\mathbf{P}) ,$$

therefore

$$\frac{1}{\sqrt{n}} \max_{1 \leq k \leq n} \|\Phi_{k-1}\| = o(1) , \qquad a.s.$$

Since $\alpha' > 0$, we conclude that

$$\mathbf{V}_n \overset{a.s}{\to} 0 ;$$

this and (9) imply

$$\frac{1}{\sqrt{n}} T_n \overset{\mathscr{L}}{\to} T ,$$

where $T \sim N(0, \Gamma_\varepsilon \otimes \mathbf{P})$; and this completes the proof □.

Using the covariance function of the input signal (U_n), let:

$$\mathbf{R} = \begin{bmatrix} \Gamma(0) & \Gamma'(1) & \dots & \dots & \Gamma'(dp+s-1) \\ \Gamma(1) & & \ddots & & \vdots \\ \vdots & & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \Gamma'(1) \\ \Gamma(dp+s-1) & \dots & \dots & \Gamma(1) & \Gamma(0) \end{bmatrix} .$$

Corollary: If, in addition to the assumptions $(\mathbf{H}_i)$, $i = 1, 2, 3$, *we assume also that the matrices* $\mathbf{A}_p$, Γ_ε *and* $\mathbf{R}$ *are regular, then*:

$$\sqrt{n}(\theta_n - \theta)' \xrightarrow{\mathscr{L}} \mathbf{P}^{-1}T \ , \tag{10}$$

This result is a direct consequence of the previous theorem and (8) when $\mathbf{P}$ is positive definite. To establish the positive definiteness of $\mathbf{P}$, it sufficient to show that:

$$\underline{\lim}\ \lambda_{\min}\left(\frac{1}{n}\mathbf{P}_n\right) > 0 \quad a.s. \ , \tag{11}$$

($\lambda_{\min}$ denotes here the smallest eigenvalue). To show (11), we use a classical result upon the transfer of excitation {Lai & Wei (1986), theorem 2} which implies that there exists $\rho > 0$ such that

$$\lambda_{\min}\left(\frac{1}{n}\mathbf{P}_n\right) \geq \rho\lambda_{\min}\left(\frac{1}{n}\mathbf{X}_n\right) \ ,$$

where $\mathbf{X}_n = \sum_{k=1}^{n} x_k x_k'$ and $x_n = (U_{n-1}', \ldots, U_{n-dp-s}', \varepsilon_{n-1}', \ldots, \varepsilon_{n-dp}')'$; moreover

$$\frac{1}{n}\mathbf{X}_n \xrightarrow{a.s.} \mathbf{X} = \operatorname{diag}(\mathbf{R}, \Gamma_\varepsilon, \ldots, \Gamma_\varepsilon) \ , \qquad a.s. \ ;$$

and the matrix $\mathbf{X}$ is obviously positive definite; then the desired conclusion holds. □

Example: Consider the scalar model:

$$y_n = ay_{n-1} + bu_{n-1} + \varepsilon_n \ , \qquad |a| < 1 \ . \tag{12}$$

The noise (ε_n) is a martingale difference sequence such that $E(\varepsilon_n^2/\mathscr{F}_{n-1}) = \sigma_\varepsilon^2 > 0$; the input signal (u_n) is deterministic given by

$$u_n = \sin(n\omega_1 + \varphi) \ , \qquad \omega_1 \in]0, \pi[\ , \qquad \varphi \text{ arbitrary} \ . \tag{13}$$

It is easy to show that (u_n) has a persistent excitaton of degree 2, and an intensity equivalent to n, (cf. Viano (1987)). It satisfies:

a) $|u_n| \leq 1$.
b)

$$\Gamma_n(l) \triangleq \frac{1}{n} \sum_{1}^{n} u_k u_{k+l} \to \Gamma(l) = \frac{1}{2} \cos(l\omega_1) \ .$$

If we denote by $\delta_{\omega_1}(.)$ the Dirac distribution at ω_1, we can easily show that the spectral measure of $e_n = (\varepsilon_n, u_n)'$ is given by:

$$d\xi_e(\omega) = \begin{bmatrix} \frac{\sigma_\varepsilon^2}{2\pi} d\omega & 0 \\ 0 & \frac{1}{2}\delta_{\omega_1}(\omega) d\omega \end{bmatrix} ,$$

Now $L(z) = \begin{bmatrix} \frac{1}{A(z)} & \frac{bz}{A(z)} \\ 0 & 1 \end{bmatrix}$, $A(z) = 1 - az$.

Then

$$\mathbf{P} = \int_{-\pi}^{\pi} \begin{bmatrix} \frac{1}{A(e^{i\omega})A(e^{-i\omega})} \frac{\sigma_\varepsilon^2}{2\pi} d\omega + \frac{b^2}{2A(e^{i\omega})A(e^{-i\omega})} \delta_{\omega_1}(\omega) d\omega & \frac{be^{i\omega}}{2A(e^{i\omega})} \delta_{\omega_1}(\omega) d\omega \\ \frac{be^{i\omega}}{2A(e^{i\omega})} \delta_{\omega_1}(\omega) d\omega & \frac{1}{2}\delta_{\omega_1}(\omega) d\omega \end{bmatrix} ,$$

and after computation of the integrals we obtain:

$$\mathbf{P} = \begin{bmatrix} \frac{\sigma_\varepsilon^2}{1 - a^2} + \frac{b^2/2}{(1 - a\cos(\omega_1))^2 + a^2 \sin(\omega_1)^2} & \frac{b(\cos(\omega_1) - a)/2}{(1 - a\cos(\omega_1))^2 + a^2 \sin(\omega_1)^2} \\ \frac{b(\cos(\omega_1) - a)/2}{(1 - a\cos(\omega_1))^2 + a^2 \sin(\omega_1)^2} & \frac{1}{2} \end{bmatrix} ,$$

which is positive definite.

Hence the asymptotic normality of the O.L.S.E. of $\theta = (a, b)'$ is obtained, the limiting distribution is a gaussian vector with zero mean and covariance $\sigma_\varepsilon^2 \mathbf{P}^{-1}$.

Remark: The assumptions about the regularity of the matrix **R** is realistic. Indeed, it is tied to the *richness* of the input signal; and is equivalent to a persistent excitation of degree $dp + s$ and an intensity of excitation equal to n (see Bay & Sastry (1987), Moore (1983), for more details).

Acknowledgements: The authors thank the referee for valuable suggestions.

References

[1] Bay EW, Sastry SS (1985) Persistency of excitation, sufficient richness and parameter convergence in discrete time adaptive control. System and control letters 6: 153–163

[2] Boutahar M (1991) Distribution asymptotique de l'estimateur des moindres carrés. Cas des modèles $ARX(p, s)$ instables. Stochastics and Stochastics Report 37: 105–126

[3] Chan NH, Wei CZ (1988) Limiting distributions of least squares estimates of unstable autoregressive processes. The Annals of Statistics 16, 1: 367–401

[4] Crowder M (1980) On the asymptotic properties of least squares in autoregression. The Annals of Statistics 8, 1: 132–146

[5] Dickey DA, Fuller WA (1979) Distribution of estimators for autoregressive times series with a unit root. J Amer Statist Ass 74: 427–431

[6] Lai TL, Wei CZ (1982) Least square estimates in stochastic regression models with application to identification and control of dynamic systems. The Annals of Statistics 10, 1: 154–186

[7] Lai TL, Wei CZ (1986) On the concept of excitaton in least square identification and adaptive control. Stochastics 16: 227–254

[8] Moore JB (1983) Persistence of excitation in extended least square. IEEE Trans Autom Control 28: 60–68

[9] Rauzy G (1976) Propriétés statistiques des suites arithmétiques. PUF

[10] Reinsel G (1979) FIML estimation of the dynamic simultaneous equations models with ARMA disturbances. Journal of Econometrics 9, 3: 263–291

[11] Touati A (1990) Loi limite de l'esimateur des moindres carrés dans le modèle autorégressif mixte. Ann de L'IHP 26

[12] Viano MC (1987) Degré d'excitation persistante et convergence presque sûre de l'estimateur des moindres carrés généralisé dans un modèle ARMAX. CRAS Tome 305, I, 6

Received 23.06.1993
Revised version 31.03.1994