Statistique et analyse des données

WOJCIECH PIECZYNSKI

Time series and counting estimation

Statistique et analyse des données, tome 14, nº 3 (1989), p. 29-36.
<http://www.numdam.org/item?id=SAD_1989__14_3_29_0>

Statistique et Analyse des Données
1989 - Vol. 14 n°3 - p. 29 - 36

TIME SERIES AND COUNTING ESTIMATION

Wojciech PIECZYNSKI

Département Mathématiques et Systèmes de Communication

Ecole Nationale Supérieure des Télécommunications de Bretagne

B.P. 832 - 29285 BREST CEDEX - FRANCE

Résumé

Nous proposons une nouvelle méthode d'estimation bien adaptée au cas où l'ensemble des paramètres est fini et les observations non indépendantes. Elle est appliquée à l'estimation des coefficients d'un processus auto-régressif non asymptotiquement stationnaire.

Mots clés : processus auto-régressif, non stationnarité.

Classification AMS : 62 F 10, 62 M 10
Classification STMA : 04 030, 12 070

Abstract

Supposing the set of parameters finite and the observations dependent, we propose a new method of estimation which can be applied to non-asymptotically stationary auto-regressive processes.

Keywords : auto-regressive processes, non stationarity.

1. INTRODUCTION

In the first part, we introduce a new method of estimation (counting estimator) well suited to the case, in which the distribution of a process is defined by conditional distributions and the set of parameters finite. The construction of the counting estimator (CE) is quite natural and we can show, under certain conditions, its consistency at the exponential rate of convergence.

Manuscrit reçu le 20 mars 1989, révisé le 16 novembre 1989

In the second part, we study AR processes (i.e. verifying $X_n = \alpha_1 X_{n-1} + \alpha_2 X_{n-2} + ... + \alpha_p X_{n-p} + U_n$, with (U_n) white gaussian noise). A great number of works have been devoted to the problem of estimation of the parameter $\theta = (\alpha_1,...\alpha_p)$. In [5] P. Newbold presents a compilation of the main results accompanied by a vast bibliography. The maximum likelihood and the least squares constitute the two groups of proposed techniques. The contribution made by CE as regards these techniques is to be found at the hypotheses level : indeed, the study of the asymptotic behaviour of known estimators requires the asymptotic stationarity of X with which our hypothesis, which implies the convergence of CE, is incompatible. The proof of the result is based upon a technical lemma by J. Geffroy ([3]). In considering a second order autoregressive process, we show in particular that if the roots of the associated polynomial are real, distinct, and of a modulus superior to one, then the CE converges (the usual hypothesis consists of supposing them to be of a modulus inferior to one).

More generally, the use of the CE is interesting when the distribution of the process is defined by stationary (i.e. independent from n) conditional distributions, which facilitates the proof of its convergence, but the process is not asymptotically stationary, which makes applying the law for large numbers difficult.

2. COUNTING ESTIMATOR

Let us consider a sequence of real, not necessarily independent random variables $X = (X_1, X_2,...,X_n,...)$ and suppose that the distribution of X is defined by the conditional distributions $P_n(x^{(n-1)}, \theta)$, θ being a parameter belonging to a finite set $\Theta = \{\theta_1,\theta_2,...\theta_k\}$. Let us suppose that for any $n \in N$, $n \geq 2$ we can, after having observed $x^{(n-1)} = (x_1,..,x_{n-1})$, define a partition $[B_i]_{1 \leq i \leq k}$ (which depends on $x^{(n-1)}$) of R such that :

$$\forall 1 \leq i \leq k \quad P_n(x^{(n-1)}, \theta_i)(B_i) \geq \sup_{j \neq i} P_n(x^{(n-1)}, \theta_j)(B_i) \tag{1}$$

We can than consider a "partial" estimation of θ defined by :

$$[\tilde{\theta}_n(x^{(n-1)}, x_n) = \theta_i] \Leftrightarrow [x_n \in B_i] \tag{2}$$

this procedure defines, for any $x^{(n)} = (x_1, x_2,..., x_n)$, n values $\tilde{\theta}_1(x_1), \tilde{\theta}_2(x^{(2)}), ..., \tilde{\theta}_n(x^{(n)})$. It is then natural to choose as an estimated parameter the element $\theta \in \Theta$ the most frequently affected by the sequence $\tilde{\theta}_1, \tilde{\theta}_2,..., \tilde{\theta}_n$. To be more precise, by denoting

$$N_{n,\theta} = \text{Card}\{ 1 \leq j \leq n \,/\, \tilde{\theta}_j = \theta \} \tag{3}$$

let us put :

Definition

We call the counting estimator any estimator $(\hat{\theta}_n)$ verifying :

$$[\hat{\theta}_n = \theta_i] \Leftrightarrow [N_{n,\theta_i} = \sup_{1 \leq j \leq k} N_{n,\theta_j}] \tag{4}$$

Thus we obtain the estimated value by counting, for each $\theta \in \Theta$, the partial estimations which designate the latter, which justifies its being called the estimator.

3. ESTIMATION IN AR PROCESS

Let (U_n) be a real, centred, gaussian, stationary white noise, and (X_n) a process verifying :

$$X_n = \alpha_1 X_{n-1} + \alpha_2 X_{n-2} + ... + \alpha_p X_{n-p} + U_n$$

for $n \geq 1$. We suppose $X_0 = X_{-1} = ... = X_{1-p} = 0$ and wish to estimate the unknown parameter $\theta = (\alpha_1, ..., \alpha_p)$. For any $n \in N$, $n \geq p+1$ let us put $X_{n,p} = (X_{n-p-1}, X_{n-p}, ..., X_{n-1})$ and let us denote $<.,.>$ the scalar product in R^p.$X_{n,p}$ being a centred gaussian vector, $<X_{n,p}, x>$ is, for any $x \in R^p$, a centred gaussian random variable. By designating Var_θ the variance let us consider the following property (H) :

(H) For any $\theta \in \Theta$ and $x \in R^p$, $x \neq 0$ $\quad Var_\theta < X_{n,p}, x > \underset{n \to +\infty}{\longrightarrow} +\infty$

(H) is obvioulsy incompatible with the asymptotic stationarity. Let Θ be finite, $k = Card(\Theta)$. By denoting $(\hat{\theta}_n)$ the counting estimator, we can state the following result :

THEOREM

1 - There is at least one counting estimator.

2 - (H) implies the existence of two constants $a > 0$, $c > 0$, such that :

$$P_\theta [\hat{\theta}_n \neq \theta] \leq ae^{-cn}$$

for any $\theta \in \Theta$ and $n \geq 1$.

Proof.

1 - Given the nature of the process, $\tilde{\theta}_n$ depends only on the last p components of $x^{(n-1)} = (x_1,...,x_{n-1})$. Let us denote $x_{n,p} = (x_{n-p-1},....,x_{n-1})$ and, for any $\theta \in \Theta$, $P_{\theta,n,p}$ the probability distribution of $X_{n,p}$. $P_{\theta,n,p}$ admits a density with respect to Lebesgue's measure - there results that for all θ_1,θ_2 the set $B_{\theta_1,\theta_2} = \{x \in R^P / <x,\theta_1-\theta_2>=0\}$ is $P_{\theta,n,p}$ negligible. It is the same for $B = \bigcup_{\theta_1,\theta_2} B_{\theta_1,\theta_2}$. Let us denote ^cB the complementary of B.

The conditional distribution of X_n, given $X_{n,p} = x_{n,p}$ is the distribution of U_n translated from the parameter $\beta_\theta = <\theta, x_{n,p}>$ - thus, for any $x_{n,p} \in {}^cB$, each translation parameter is different from all the others. After suitably numbering the θ we can write : $\beta_{\theta_1} < \beta_{\theta_2} < ... < \beta_{\theta_k}$ ($x_{n,p}$ is fixed). U_n being a gaussian variable, the B_i defined by : $B_1 =]-\infty, \beta_{\theta_1}]$, $B_i =]\beta_{\theta_{i-1}}, \beta_{\theta_i}]$ for $2 \le i \le k-1$, $B_k =]\beta_{\theta_k}, +\infty[$ evidently verify (1) with strict inequality. $\tilde{\theta}_n$ is therefore well defined and we can determine a counting estimator.

2 - Let us put, for any $\theta \in \Theta$, $\varepsilon > 0$:

$$D_{\theta,\varepsilon} = \{x \in R^P / <\theta - \theta_1, x> \ge \varepsilon \text{ for each } \theta_1 \in \Theta, \theta_1 \neq \theta\} \qquad (5)$$

By defining the sets B_θ (which also depend on $x_{n,p}$) as above, and by denoting F the distribution function of U_n, we can write for any $\theta \neq \theta_1$:

$$x_{n,p} \in D_{\theta,\varepsilon} \Rightarrow \begin{cases} P_{\theta,x_{n,p}}(B_\theta) \ge F(\varepsilon/2) - F(-\varepsilon/2) \\ P_{\theta_1,x_{n,p}}(B_\theta) \le F(-\varepsilon/2) - F(-\varepsilon) \end{cases} \qquad (6)$$

where $P_{\theta,x_{n,p}}$ is the conditional distribution of X_n given $X_{n,p} = x_{n,p}$.

Furthermore, we have the following property :

$$\text{For any } \theta,\theta' \quad P_{\theta',n,p}(D_{\theta,\varepsilon}) \underset{n \to +\infty}{\longrightarrow} 1 \qquad (7)$$

In fact, we can write : $D_{\theta,\varepsilon} = \bigcap_{\theta_1 \neq \theta} D_{\theta,\theta_1,\varepsilon}$ with $D_{\theta,\theta_1,\varepsilon} = \{x \in R / <\theta-\theta_1, x> \ge \varepsilon\}$.

Denoting by Q_n the distribution of the gaussian r.v. $< X_{n,p}, \theta - \theta_1 >$ we have $P_{\theta,n,p}(D_{\theta,\theta_1,\varepsilon}) = 1 - Q_n[-\varepsilon,\varepsilon]$. Now, the dispertion of Q_n tends, by virtue of hypothesis (H), towards $+\infty$, so $Q_n[-\varepsilon,\varepsilon]$ tends to 0, from which we get (7). By supposing $\omega_1 = F(\varepsilon/2) - F(-\varepsilon/2)$, $\omega_0 = F(-\varepsilon/2) - F(-\varepsilon)$, (6) can be written as :

$$\text{For any } x_{n,p} \in D_{\theta,\varepsilon} \quad \begin{cases} P_{\theta,x_{n,p}}(B_\theta) \geq \omega_1 \\ \\ P_{\theta',x_{n,p}}(B_\theta) \leq \omega_0 \quad \text{for any } \theta' \neq \theta \end{cases} \tag{8}$$

Let us consider, for any $\theta \in \Theta$ and $n \geq p$ the application $\zeta_{\theta,n} : R^n \to \{0, 1\}$ defined by :

$$\zeta_{\theta,n}(x^{(n)}) = 1_{B_\theta}(x_n) \tag{9}$$

where $x^{(n)} = (x_1, ..., x_{n-1}, x_n)$ and B_θ is the borelian set verifying (1) for $x^{(n-1)}$ and θ. B_θ depends only on the last p components of $x_{(n-1)}$, so $\zeta_{\theta,n}(x^{(n)})$ depends only on the last p+1 components of $x^{(n)}$ - this results in the conditional expectation $E_\theta[\zeta_{\theta,n} / \zeta_{\theta,1}..,\zeta_{\theta,n-1}]$ being equal to $E_\theta[\zeta_{\theta,n} / \zeta_{\theta,n-p-1},...,\zeta_{\theta,n-1}]$. For any $A = \{\zeta_{\theta,n-p-1} = q_1,...,\zeta_{\theta,n-1} = q_p\}$ ($q_1,...,q_p$ in $\{0,1\}$) we have :

$$E_\theta\left[\zeta_{\theta,n} / A\right] = \frac{1}{P_{\theta,n,p}(A)} \int_A P_n\left(x^{(n-1)}, \theta\right) dP_{\theta,n,p}$$

$$\geq \frac{1}{P_{\theta,n,p}(A)} \int_{A \cap D_{\theta,\varepsilon}} P_n\left(x^{(n-1)}, \theta\right)(B_\theta)\, dP_{\theta,n,p}$$

$$\geq \frac{P_{\theta,n,p}\left(A \cap D_{\theta,\varepsilon}\right)}{P_{\theta,n,p}(A)}\, \omega_1$$

(7) implies that the above minorant tends, for any θ and A, towards ω_1.

The θ and A being in finite number we can affirm, by supposing $\omega'_1 = \omega_1 - (\omega_1 - \omega_0)/3$, that for any n superior to a certain n_0 :

$$E_\theta\, [\zeta_{\theta,n} \,/\, \zeta_{\theta,n-1-p},\ldots,\zeta_{\theta,n-1}] \geq \omega'_1 \qquad (10)$$

By supposing $\omega'_0 = \omega_0 + (\omega_1 - \omega_0)/3$ we could show in a similar way that for n large enough

$$\forall\ \theta' \neq \theta \qquad E_\theta\, [\, \zeta_{\theta',n} \,/\, \zeta_{\theta',n-p-1},\ldots,\zeta_{\theta',n-1}] \leq \omega'_0 \qquad (11)$$

We can then apply to the sequence $(\zeta_{\theta,n})$ J. Geffroy's lemma ([3] page 430) : by supposing $\gamma_n = \omega'_1 - \omega'_0$, $\alpha_n = \alpha = (\omega'_1 - \omega'_0)/2$ and $c = (\omega'_1 - \omega'_0)^2/4$, we can write for a certain $b > 0$

$$P_\theta\, (\zeta_{\theta,1}+\ldots+ \zeta_{\theta,n} > n\alpha) > 1 - be^{-cn} \qquad (12)$$

$$\forall\ \theta' \neq \theta \qquad P_\theta\, (\zeta_{\theta',1}+\ldots+\zeta_{\theta',n} \leq n\alpha) > 1 - be^{-cn} \qquad (13)$$

Let us denote that $N_{n,\theta} = \zeta_{\theta,1}+\ldots+ \zeta_{\theta,n}$. Given that $k = \text{Card}\,(\Theta)$, (12) and (13) imply :

$$P_\theta\, [(N_{n,\theta} > n\alpha) \cap (\bigcap_{\theta' \neq \theta} (N_{n,\theta'} \leq n\alpha))] \geq 1 - kbe^{-cn} \qquad (14)$$

Finally, given the definition of the counting estimator we can write

$$\{N_{n,\theta} > n\alpha\} \cap \{\bigcap_{\theta' \neq \theta} \{N_{n,\theta'} \leq n\alpha\}\} \subset \{\hat{\theta}_n = \theta\} \qquad (15)$$

(14) and (15) imply :

$$P_\theta(\hat{\theta}_n = \theta) \geq 1 - kbe^{-cn}$$

which ends the proof (with $a = kb$).

In order to illustrate the use of (H), let us consider a second order autoregressive process $X_n = \alpha_1 X_{n-1} + \alpha_2 X_{n-2} + U_n$. Let us suppose, without losing any degree of generality, that $\text{Var}\, U_n = 1$. Let us denote A the set of couples (α_1, α_2) for which the roots of the polynomial $p(z) = z^2 - \alpha_1 z - \alpha_2$ are real, distinct and of a modulus superior to one. We can state :

PROPOSITION

If $\Theta \subset A$ the hypothesis (H) hold.

Proof of the proposition.

Let $z_1 \neq z_2$ be the roots of the polynomial $p(z) = z^2 - \alpha_1 z - \alpha_2$. This gives :

$$X_n = U_n + c_1 U_{n-1} + c_2 U_{n-2} + ... + c_{n-1} U_1 \quad (16)$$

with

$$c_k = \left(z_1^{k+1} - z_2^{k+1}\right) / \left(z_1 - z_2\right) \quad (17)$$

(see [7]). This results in :

$$a_n = \text{Var}(X_n) = 1 + c^2_1 + ... + c^2_{n-1} \quad (18)$$
$$b_n = \text{Cov}(X_n, X_{n-1}) = c_1 + c_1 c_2 + ...c_{n-2} c_{n-1} \quad (19)$$

For any $(x,y) \in R^2$ the variance of $< X_{n+1,2},^t (x,y) > = < (X_{n-1}, X_n),^t (x,y)>$ $(p = 2)$ is :
$a_{n-1} x^2 + 2 b_n xy + a_n y^2 = (1 + c^2_1 + ...+ c^2_{n-2}) x^2 + 2 (c_1 + c_1 c_2 + ...+ c_{n-2} c_{n-1})xy +$

$$(1 + c^2_1 +... + c^2_{n-1}) y^2 = x^2 + 2 c_1 xy + (1 + c^2_2) y^2 + \sum_{k=1}^{n-2} (c_k x + c_{k+1} y)^2$$

Therefore, to show that (H) is verified, all that has to be done is show that the series $d_k = (c_k x + c_{k+1} y)^2$ diverges. Let us suppose that $|z_1| \geq |z_2|$ and let us put $p = z_2 / z_1$. Given (17), we get :

$$d_k = (z_1 - z_2)^{-2} z^{2}_1{}^{(k+1)} ((1 - p^{k+1}) x + z_1 (1 - p^{k+2}) y)^2 \quad (20)$$

Let us suppose $|z_1| > |z_2|$. $z^2_1{}^{(k+1)} \underset{k \to +\infty}{\longrightarrow} +\infty$ and

$((1 - p^{k+1}) x + z_1(1 - p^{k+2}) y)^2 \underset{k \to +\infty}{\longrightarrow} (x + z_1 y)^2$, therefore, if $x + z_1 y \neq 0$ the series diverges (see (20)). If $x + z_1 y = 0$ (20) implies : $d_k = (z_1 - z_2)^{-2} (z_1 z_2)^{k+1} (x + z_2 y)$. Divergence is obtained for $k \to +\infty$ since $|z_1 z_2| > 1$ and $x + z_2 y \neq 0$.Let us suppose $|z_1| = |z_2|$. So $z_1 = - z_2$ and $p = -1$. According to (20) :

(see (20)). If $x + z_1 y = 0$ (20) implies : $d_k = (z_1 - z_2)^{-2} (z_1 z_2)^{k+1} (x + z_2 y)$. Divergence is obtained for $k \to +\infty$ since $|z_1 z_2| > 1$ and $x + z_2 y \neq 0$.Let us suppose $|z_1| = |z_2|$. So $z_1 = - z_2$ and $p = -1$. According to (20) :

$$d_k = \begin{cases} 4 y^2 (z_1 - z_2)^{-2} z_1^{2(k+2)} = y^2 z^2{}_1^{(k+1)} & \text{if k is odd} \\ 4 x^2 (z_1 - z_2)^{-2} z_1^{2(k+1)} = x^2 z^2{}_1^{k} & \text{if k is even} \end{cases}$$

Divergence is obtained for $k \to +\infty$ since $(x,y) \neq (0,0)$.

References

1 - **Anderson, T.W.** - *The statistical analysis of time series.* Wiley, New-York, 1971.

2 - **Box, G.E.P. et Jenkins, G.M.** - *Time series analysis, forecasting and control.* Holden day, San Francisco, 1976.

3 - **Geffroy, J.** - Sufficient convergence conditions for some tests in the case of not necessarily independent or equidistributed data. *Recent dev. in Stat. J.R. Barra et al. editors. North Holland Publishing Compagny,* 1977.

4 - **Hannan, E.J.** - *Multiple time series.* Wiley, New York, 1970.

5 - **Newbold, P.** - Some recent developments in time series analysis. *Intern. Statist. Rev.* 49, 53-66, 1981.

6 - **Pieczynski, W.** - L'application de la méthode de la décantation dans un processus autorégressif - *C.R.A.S.*, Paris, t. 308, Série I, p. 583-585, 1989.

7 - **Prestley, M.B.** - *Spectral analysis and time series.* Academic Press, New York, 1981.