C. R. Acad. Sci. Paris, t. 314, Série I, p. 403-406, 1992

Statistique/*Statistics*

Estimation pondérée et poursuite pour les modèles ARMAX

Bernard BERCU

Résumé – Pour un modèle **ARMAX** vectoriel complexe, on étudie l'algorithme des moindres carrés pondéré qui améliore l'algorithme des moindres carrés usuel par le choix de pondérations convenables. Concernant les problèmes de poursuite adaptative, on assure à la fois la consistance de l'estimateur et l'optimalité du contrôle. On précise également les vitesses de convergence presque sûre.

Weighted estimation and tracking for ARMAX models

Abstract – *For a complex multivariate* **ARMAX** *model, we study the weighted least squares algorithm which improve the usual least squares algorithm by the choice of suitable ponderations. Concerning adaptive tracking problems, we ensure both strong consistency of the estimator and control optimality. We also precise almost sure rates of convergence.*

1. ESTIMATION PONDÉRÉE. – Soit $(\Omega, \mathscr{A}, P)$ un espace de probabilité muni d'une filtration $F=(\mathscr{F}_n)_{n\geqq 0}$. On considère le modèle **ARMAX** vectoriel complexe d'ordre (p, q, r)

$$A(R)Y_n = B(R)U_n + C(R)\varepsilon_n, \tag{1}$$

où Y, U et ε sont respectivement l'observation (d_1), le contrôle (d_2) et le bruit (d_1) associés au modèle. Pour l'opérateur retard R, on a

$$A(R) = I_{d_1} - A_1 R - \ldots - A_p R^p, \tag{2}$$

$$B(R) = B_1 R + \ldots + B_q R^q, \tag{3}$$

$$C(R) = I_{d_1} + C_1 R + \ldots + C_r R^r, \tag{4}$$

où A_i, B_j, C_k sont des matrices déterministes inconnues. La relation (1) peut s'écrire sous la forme

$$Y_{n+1} = {}^*\theta\,\Psi_n + \varepsilon_{n+1}, \tag{5}$$

avec ${}^*\theta = (A_1, \ldots, A_p, B_1, \ldots, B_q, C_1, \ldots, C_r)$ et ${}^t\Psi_n = ({}^tY_n^p, {}^tU_n^q, {}^t\varepsilon_n^r)$ où l'on pose ${}^tY_n^p = ({}^tY_n, \ldots, {}^tY_{n-p+1})$, ${}^tU_n^q = ({}^tU_n, \ldots, {}^tU_{n-q+1})$ et aussi ${}^t\varepsilon_n^r = ({}^t\varepsilon_n, \ldots, {}^t\varepsilon_{n-r+1})$. La suite régressive Ψ est inobservable et θ est le paramètre inconnu que l'on cherche à estimer. On suppose que le bruit ε vérifie

$$E(\varepsilon_{n+1} \mid \mathscr{F}_n) = 0, \qquad E(\varepsilon_{n+1}\,{}^*\varepsilon_{n+1} \mid \mathscr{F}_n) = \Gamma. \tag{6}$$

Soient $\hat{\varepsilon}_n$, Φ_n et $\hat{\theta}_n$ des estimateurs de ε_n, Ψ_n et θ respectivement. On estime ε_n par l'erreur *a posteriori*

$$\hat{\varepsilon}_n = Y_n - {}^*\hat{\theta}_n \Phi_{n-1}, \qquad {}^t\Phi_n = ({}^tY_n^p, {}^tU_n^q, {}^t\hat{\varepsilon}_n^r). \tag{7}$$

Soit $a=(a_n)$ une suite aléatoire adaptée à F, positive, décroissante et $\leqq 1$. On propose, pour estimer θ, l'estimateur des moindres carrés pondérés (**MCP**) introduit par Bercu et Duflo [4]

$$\hat{\theta}_{n+1} = \hat{\theta}_n + a_n S_n^{-1} \Phi_n^* (Y_{n+1} - {}^*\hat{\theta}_n \Phi_n) \tag{8}$$

Note présentée par Jean-Pierre KAHANE.

0764-4442/92/03140403 $ 2.00

où la valeur initiale $\hat{\theta}_0$ est choisie arbitrairement et

(9) $$S_n = \sum_{k=0}^{n} a_k \Phi_k {}^*\Phi_k + S$$

où S est une matrice déterministe, hermitienne et inversible. On utilise

(10) $$f_n = a_n {}^*\Phi_n S_n^{-1} \Phi_n, \qquad s_n = \sum_{k=0}^{n} \| \Phi_k \|^2 + s$$

où $s = \mathrm{tr}(S)$. On travaille également avec

(11) $$\Delta_n = \sum_{k=1}^{n} a_k f_k, \qquad \Delta = \lim_{n \to +\infty} \Delta_n.$$

On dit que $a = (a_n)$ est une pondération admissible si Δ est intégrable. On peut choisir, par exemple, $a_n = n^{-1-\gamma}$, $a_n = (\log s_n)^{-1-\gamma}$, $a_n = s_n^{-\gamma}$ avec $\gamma > 0$. La suite des erreurs de prédiction est $\pi = (\pi_n)$ avec

(12) $$\pi_n = {}^*\theta \Psi_n - {}^*\hat{\theta}_n \Phi_n.$$

2. Consistance.

Théorème 1. — *Pour l'algorithme* **MCP**, *on suppose que* C *est passif et* a *est une pondération admissible. Alors*

(13) $$\| S_n^{1/2} (\hat{\theta}_{n+1} - \theta) \|^2 = O(1) \quad p.s.$$

(14) $$E\left[\sum_{n=1}^{\infty} a_n \| {}^*(\hat{\theta}_{n+1} - \theta) \Phi_n \|^2 \right] < +\infty$$

(15) $$E\left[\sum_{n=1}^{\infty} a_n \| \Phi_n - \Psi_n \|^2 \right] < +\infty$$

Théorème 2. — *Dans le cadre du théorème* 1, *on a*

(16) $$E\left[\sum_{n=1}^{\infty} \| \hat{\theta}_{n+1} - \hat{\theta}_n \|^2 \right] < +\infty.$$

De plus, soit $b_n^{-1} = \sup \{ a_n, \mathrm{tr}(S_{n-1}^{-1} - S_n^{-1} \}$. *Alors, si* $\mathrm{tr}(S_n^{-1}) \leqq \mathrm{Cte}\, a_n$

(17) $$E\left[\sum_{n=1}^{\infty} b_n \| \hat{\theta}_{n+1} - \hat{\theta}_n \|^2 \right] < +\infty.$$

Corollaire 1. — *Soit* $c_n = \inf \{ a_n, \| \Phi_n \|^{-2} \}$ *si* $\| \Phi_n \| \neq 0$ *et* $c_n = a_n$ *sinon. Alors :*

(18) $$\sum_{n=1}^{\infty} c_n \| \pi_n \|^2 < +\infty.$$

De plus, si $s_n^{-1} = O(a_n)$, *comme pour l'algorithme du gradient*

(19) $$\sum_{n=1}^{\infty} \frac{\| \pi_n \|^2}{s_n} < +\infty.$$

Corollaire 2. — *Soit* $d_n = \inf \{ a_n, b_n \| \Phi_n \|^{-2} \}$ *si* $\| \Phi_n \| \neq 0$ *et* $d_n = a_n$ *sinon. Alors, si* $\mathrm{tr}(S_n^{-1}) \leqq \mathrm{Cte}\, a_n$

(20) $$\sum_{n=1}^{\infty} d_n \| \pi_n \|^2 < +\infty.$$

De plus, si $s_n^{-1}=O(a_n^2)$, *on a*

$$\sum_{n=1}^{\infty}\frac{b_n}{s_n}\|\pi_n\|^2<+\infty. \tag{21}$$

3. Poursuite. — On cherche à poursuivre une trajectoire de référence $y=(y_n)$. On utilise le contrôle adaptif de poursuite (**CAP**) qui vérifie la relation

$$y_{n+1}={}^*\hat{\theta}_n\,\Phi_n. \tag{22}$$

Il est p. s. défini si, pour $n\geqq 0$, la loi de ε_{n+1} conditionnelle à $\mathscr{F}_n$ est absolument continue par rapport à la mesure de Lebesgue. On suppose également que le bruit ε satisfait la loi forte des grands nombres.

3.1. *Optimalité.* — On utilise l'hypothèse de causalité suivante :

(A_1) $d_2\leqq d_1$ et la matrice B_1 est de rang d_2. De plus, si B_+ est l'inverse à gauche de B_1 et si $D(R)=B_+\,R^{-1}\,B(R)$, alors D est causal.

Théorème 3. — *Pour l'algorithme* **MCP**, *on suppose que* C *est passif et que* (A_1) *est satisfaite. On suppose aussi que* y_{n+1} *est* $\mathscr{F}_n$ *mesurable et que*

$$\sum_{k=1}^{n}\|y_k\|^2=O(n)\qquad p.s. \tag{23}$$

Alors, si a est admissible avec $s_n^{-1}=O(a_n)$, *le* **CAP** *est optimal.*

3.2. *Optimalité et consistance.* — On se contente d'un résultat pour une trajectoire y fortement excitante. Si le bruit ε admet un moment d'ordre >2, on peut également obtenir un résultat analogue avec un contrôle de poursuite excité et atténué semblable à celui de Guo et Chen [17]. Cette approche est développée par Bercu et Duflo ([4], [5]). On utilise l'hypothèse d'irréductibilité suivante :

(A_2) La matrice $B_+\,B_q$ est inversible et les polynômes matriciels $B_+\,A$, $B_+\,C$ et D sont premiers entre eux à gauche.

Théorème 4. — *On se place dans le cadre du théorème* 3. *On suppose que* (A_2) *est satisfaite et que* Γ *est inversible. On suppose que* y_{n+1} *est* $\mathscr{F}_{n-p-s}$ *et* $\mathscr{F}_{n-r-s}$ *mesurable avec*

$$\liminf\lambda_{\min}\left(\frac{1}{n}\sum_{k=p+s}^{n}y_k^{p+s+1}\,{}^*y_k^{p+s+1}\right)>0\quad p.s. \tag{24}$$

Alors, si a est admissible avec $s_n^{-1}=O(a_n^2)$, *le* **CAP** *est optimal avec la vitesse* $o(a_n)$. *De plus,* $\hat{\theta}_n$ *est un estimateur fortement consistant de* θ *avec la vitesse* $O((na_n)^{-1})$.

Remarque. — Si l'on privilégie la consistance, on choisit $a_n=(\log s_n)^{-1-\gamma}$ avec $\gamma>0$. On trouve alors la vitesse de convergence $n^{-1}(\log n)^{1+\gamma}$. Si l'on s'intéresse à l'optimalité, on prend $a_n=s_n^{-\gamma}$ avec $0<\gamma\leqq 1/2$. On obtient alors la vitesse de convergence $n^{-\gamma}$.

Note remise le 10 décembre 1991, acceptée le 18 décembre 1991.

Références bibliographiques

[1] K. J. Aström et B. Wittenmark, On self tuning regulators, *Automatica*, 9, 1973, p. 185-199.

[2] A. H. Becker, P. R. Kumar et C. Z. Wei, Adaptive control with the stochastic approximation algorithm, Geometry and convergence, *I.E.E.E. Trans. Automat. Control*, AC-30, 1985, p. 330-338.

[3] B. Bercu, Sur l'estimateur des moindres carrés généralisé d'un modèle ARMAX. Application à l'identification des modèles ARMA, *Ann. Inst. Henri Poincaré*, 27, 1991, p. 425-443.

[4] B. BERCU et M. DUFLO, Moindres carrés pondérés et poursuite, *Ann. Inst. Henri Poincarré*, 28, 1992 (à paraître).

[5] B. BERCU, *Weighted estimation and tracking for ARMAX models*, 1992, (à paraître).

[6] P. E. CAINES, Stochastic adaptive control: Randomly varying parameters and continually disturbed controls, *Control Science and Technology for the Progress of Society*, H. AKASHI éd., Pergamon, 1981, p. 925-930.

[7] P. E. CAINES et S. LAFORTUNE, Adaptive control with recursive identification for stochastic linear systems, *I.E.E.E. Trans. Automat. Control*, AC-29, 1984, p. 312-321.

[8] P. E. CAINES, *Linear stochastic systems*, John Wiley, New York, 1988.

[9] H. F. CHEN, Recursive system identification and adaptive control by use of the modified least squares algorithm, *S.I.A.M. J. Control and Optimization*, 22, 1984, p. 758-776.

[10] H. F. CHEN et P. E. CAINES, The strong consistency of the stochastic gradient algorithm of adaptive control, *I.E.E.E. Trans. Automat. Control*, AC-30, 1985, p. 189-192.

[11] H. F. CHEN. *Recursive estimation and control for stochastic systems*, John Wiley, New York, 1985.

[12] H. F. CHEN et L. GUO, Convergence rate of least squares identification and adaptive control for stochastic systems, *Int. J. Control*, 44, 1986, p. 1459-1476.

[13] H. F. CHEN et L. GUO, Asymptotically optimal adaptive control with consistent parameter estimates, *S.I.A.M. J. Control and Optimization*, 25, 1987, p. 558-575.

[14] H. F. CHEN et J. F. ZHANG, Convergence rates in stochastic adaptive tracking, *Int. J. Control*, 49, 1989, p. 1915-1935.

[15] M. DUFLO, *Méthodes récursives aléatoires*, Masson, Paris, 1990.

[16] G. C. GOODWIN, P. J. RAMADGE et P. E. CAINES, Discrete time stochastic adaptive control, *S.I.A.M. J. Control and Opt.*, 19, 1981, p. 829-853.

[17] L. GUO, H. F. CHEN, The Aström-Wittenmark self-tuning regulator revisited and ELS-based adaptive trackers, *I.E.E.E. Trans. Automat. Cont.*, 36, 1991, p. 802-812.

[18] P. R. KUMAR et L. PRALY, Self tuning trackers, *S.I.A.M. J. Control and Optimization*, 25, 1987, p. 1053-1071.

[19] P. R. KUMAR, Convergence of adaptive control schemes using least squares parameter estimates, *I.E.E.E. Trans. Automat. Cont.*, 35, 1990, p. 416-424.

[20] T. L. LAI et C. Z. WEI, Least squares estimates in stochastic regression models with applications to identification and control of dynamic systems, *Ann. Statistics*, 10, 1982, p. 154-166.

[21] T. L. LAI et C. Z. WEI, Extended least squares and their applications to adaptive control and prediction in linear systems, *I.E.E.E. Trans. Automat. Control*, AC-31, 1986, p. 898-906.

[22] T. L. LAI et C. Z. WEI, On the concept of excitation in least squares identification and adaptive control, *Stochastics*, 16, 1986, p. 227-254.

[23] K. S. SIN et G. C. GOODWIN, Stochastic adaptive control using a modified least squares algorithm, *Automat.*, 18, 1982, p. 315-321.

[24] V. SOLO, The convergence of AML, *I.E.E.E. Trans. Automat. Control*, AC-24, 1979, p. 958-962.

Laboratoire de Statistiques, Bât. n° 425, *Mathématiques*,
Université de Paris-Sud, 91405 *Orsay Cedex*.