Metrika (1996) 43:57–67

Metrika

Least Squares Estimator for Regression Models with some Deterministic Time Varying Parameters

MOHAMED BOUTAHAR

Dept. de Math. Faculté des Sciences case 901, 163 Av. de Luminy 13288 Marseille Cedex 9, France

CLAUDE DENIAU

GREQAM Vieille Charité, 2 rue de la Charité 13002 Marseille, France

Abstract: Here we study the least squares estimates in some regression models. We assume that the evolution of the parameter is linearly explosive (i.e. polynomial), or stable (i.e. sinusoidal). We prove the strong consistency, and establish the rate of convergence.

Key Words: Martingale difference sequence, least squares, regression, rate of convergence.

1 Introduction

Models with time varying parameters are usually taken to be stochastic and used in applied works of adaptive control for example. Is the relevance of a regression model with deterministically changing parameters doubtful? Suppose we use a simple regression model $y_n = bu_{n-1} + \varepsilon_n$ instead of a true model with a trend evolution $y_n = (b + 0.001n)u_{n-1} + \varepsilon_n$ in an applied problem of control. For small values of n, error in model is not serious; but what happens if the evolution is assumed to continue indefinitely? It seems important to detect and identify quickly the trend coefficient. At least for this reason this type of model seems to be of practical importance.

To the best of the author's knowledge, there are few works about almost sure convergence of the least squares estimator (L.S.E.) of deterministic time varying parameters in stochastic regression models (see Chow (1983), Sant (1977)). Here we study the case in which the evolution of the parameter is given by $\theta_n = \beta_0\varphi_{n,0} + \beta_1\varphi_{n,1} + \cdots + \beta_r\varphi_{n,r}$, where $(\varphi_{n,i})_{0\le i\le r}$ is a given deterministic functional basis time varying and β_i are the parameters to be estimated. We consider two kind of evolutions, an explosive (polynomial) evolution, $\theta_n = \beta_0 + \beta_1 n + \cdots + \beta_r n^r$, a stable (sinusoidal) evolution, $\theta_n = A\sin(n\omega + \varphi)$.

Note that the model considered here can be regarded as a regression model with constant parameters and trending regressor, and the problem becomes a classical one of estimation of constant parameters (cf. Lai and Wei (1982),

0026-1335/96/43:1/57-67 $2.50

Davidov et al. (1992), Tsypkin and Bondarenko (1992)). The strong consistency of the L.S.E. can be deduced from many nice papers in the literature (Lai and Wei (1982), Duflo (1990), Chen and Guo (1991)). The purpose of the paper is to examine some particular cases of the functional basis $(\varphi_{n,i})_{0 \le i \le r}$ and to give a more precise rate of convergence which can not be obtained directly from the results in the previous papers. For the polynomial case we obtain a rate related to the degree of the trend, $|\hat{\beta}_k(n) - \beta_k|^2 = O(\log(n)/n^{2k+1})$; and for the sinusoidal evolution we obtain the rate $\log(n)/n$.

2 General Assumptions

Consider the scalar regression model with time varying parameters:

$$y_n = b_1(n)u_{n-1} + \varepsilon_n \ , \tag{1}$$

where y_n is an observed output, u_n is an observed input, and ε_n is an unobserved random perturbation at stage n. We assume that the parameter

$$b_1(n) = \beta_0 \varphi_{n,0} + \beta_1 \varphi_{n,1} + \cdots + \beta_r \varphi_{n,r} \ ,$$

in which $\varphi_{n,i}$ are elementary functions of n.

Noise (ε_n):
H.1: (ε_n) is a *martingale difference* sequence with respect to $F = (\mathscr{F}_n)$, (i.e. ε_n is $\mathscr{F}_n$-measurable and $E(\varepsilon_n/\mathscr{F}_{n-1}) = 0$ for all n) such that:

a.

$$\sup_n E(|\varepsilon_n|^\alpha/\mathscr{F}_{n-1}) < \infty \qquad \text{for some } \alpha > 2 \ ,$$

b. for all n, $E(\varepsilon_n^2/\mathscr{F}_{n-1}) = \sigma_\varepsilon^2 > 0$.

H'.1: (ε_n) is a sequence of i.i.d. random variables with variance $\sigma_\varepsilon^2 > 0$.

Input signal (u_n):
H.2: For all n, u_{n-1} is $\mathscr{F}_{n-1}$-measurable.

3 Main Result

We can write the model (1) into the following form:

$$y_n = \theta' \Phi_{n-1} + \varepsilon_n \tag{2}$$

where $\Phi_{n-1} = (\varphi_{n,0}, \varphi_{n,1}, \ldots, \varphi_{n,r})' u_{n-1}$ and $\theta = (\beta_0, \ldots, \beta_r)'$.

The least squares estimator $\hat{\theta}_n$ of θ is the solution of the fundamental equation:

$$\mathbf{P}_n \hat{\theta}_n = \sum_{k=1}^{n} \Phi_{k-1} y_k \tag{3}$$

where $\mathbf{P}_n = \sum_{k=0}^{n-1} \Phi_k \Phi_k'$.

Proposition 3.1: Under the assumptions H.1. and H.2., if we assume also that there is a sequence of invertible matrices (Δ_n) with limit 0 as n tends to infinity such that

$$\Delta_n \mathbf{P}_n \Delta_n' \xrightarrow{a.s.} \mathbf{P} \ , \qquad \text{where } \mathbf{P} \text{ is a positive definite matrix} \ . \tag{4}$$

Then the estimation errors $\tilde{\theta}_n = \hat{\theta}_n - \theta$ satisfy:

$$\|(\Delta_n^{-1})' \tilde{\theta}_n\|^2 = O(\log(\lambda_{\max}(\Delta_n^{-1}(\Delta_n^{-1})'))) \quad a.s. \tag{5}$$

and hence

$$\|\tilde{\theta}_n\|^2 = O(\log(\lambda_{\max}(\Delta_n^{-1}(\Delta_n^{-1})'))/\lambda_{\min}(\Delta_n^{-1}(\Delta_n^{-1})')) \quad a.s. \tag{6}$$

where $\lambda_{\min}(\cdot)$ and $\lambda_{\max}(\cdot)$ are the smallest and the largest eigenvalue.

Remark 3.2: If we replace H.1. by H'.1., then the rate of convergence is slightly modified i.e: the $O(\log(\lambda_{\max} \Delta_n^{-1}(\Delta_n^{-1})'))$ in (5) can be replaced by $o(f(\log(\lambda_{\max}(\Delta_n^{-1}(\Delta_n^{-1})')))$ for all f such that $\int \frac{1}{f(t)} dt < \infty$ (cf. Duflo 1990, th. 2.II.8).

Proof: Because of (4) $\lambda_{\min}(\Delta_n \mathbf{P}_n \Delta_n')$ is bounded away from zero and $\lambda_{\max}(\Delta_n \mathbf{P}_n \Delta_n')$ is bounded from above. Therefore, with some positive constants c_i, the following chain of inequalities holds a.s.

$$
\begin{aligned}
c_1 \|(\Delta_n^{-1})'\tilde{\theta}_n\|^2 &\leq \lambda_{\min}(\Delta_n \mathbf{P}_n \Delta_n')\|(\Delta_n^{-1})'\tilde{\theta}_n\|^2 \\
&\leq \tilde{\theta}_n' \Delta_n^{-1} \Delta_n \mathbf{P}_n \Delta_n' (\Delta_n^{-1})'\tilde{\theta}_n = \tilde{\theta}_n' \mathbf{P}_n \tilde{\theta}_n \\
&\leq c_2 \log(\lambda_{\max}(\mathbf{P}_n)) \text{ according to Lai and Wei 1982 (2.4)} \\
&\leq c_2 \log[\lambda_{\max}(\Delta_n \mathbf{P}_n \Delta_n')\lambda_{\max}(\Delta_n^{-1}(\Delta_n^{-1})')] \\
&\leq c_3 + c_2 \log[\lambda_{\max}(\Delta_n^{-1}(\Delta_n^{-1})')]
\end{aligned}
$$

As $\lambda_{\max}(\Delta_n^{-1}(\Delta_n^{-1})') \to \infty$, the result (5) follows. □

4 Particular Cases

4.1 Polynomial Evolution

If we assume that $\varphi_{n,k} = n^k$ then $b_1(n) = \beta_0 + \beta_1 n + \cdots + \beta_r n^r$. Denote $\hat{\theta}_n = (\hat{\beta}_0(n), \ldots, \hat{\beta}_r(n))'$, ($\hat{\beta}_k(n)$ is the least squares estimator of β_k, $0 \leq k \leq r$); $\tilde{\beta}_k(n) = \hat{\beta}_k(n) - \beta_k$.

Corollary 4.1: Under the assumptions H.1. and H.2., if we assume also that the regressor vector satisfies the following law of large numbers

$$
\frac{1}{n} \sum_{k=1}^{n} u_k^2 \xrightarrow{a.s.} \sigma_u^2 > 0 \ , \tag{7}
$$

then we have

$$
\|(\Delta_n^{-1})'\tilde{\theta}_n\|^2 = O(\log(n)) \quad a.s. \ , \tag{8}
$$

where

$$\Delta_n = diag(n^{-1/2}, n^{-3/2}, \ldots, n^{-(2r+1)/2}) \ ,$$

consequently for the coefficient β_k, $0 \le k \le r$, *we obtain the following rate*:

$$|\tilde{\beta}_k(n)|^2 = O\left(\frac{\log(n)}{n^{2k+1}}\right) \quad a.s. \tag{9}$$

To prove the corollary 4.1, we first establish the following

Lemma 4.2: Suppose that (Φ_n) *satisfies*:

$$\frac{1}{n}\sum_{k=1}^{n} \Phi_k \Phi_k' \xrightarrow{a.s.} \Gamma_\Phi \ ;$$

where Γ_Φ *is a deterministic matrix, then for all* $i \in \mathbb{N}$:

$$\frac{1}{n^{i+1}}\sum_{k=1}^{n} k^i \Phi_k \Phi_k' \xrightarrow{a.s.} \frac{\Gamma_\Phi}{i+1} \ ;$$

Proof: Let $R_0 = 0$, $R_n = \sum_{k=1}^{n} \Phi_k \Phi_k'$, then $\Phi_k \Phi_k' = R_k - R_{k-1}$.

$$\begin{aligned}
\frac{1}{n^{i+1}}\sum_{k=1}^{n} k^i \Phi_k \Phi_k' &= \frac{1}{n}\sum_{k=1}^{n-1}\left(\frac{k}{n}\right)^i (R_k - R_{k-1}) + \frac{1}{n}R_n - \frac{1}{n}R_{n-1} \ , \\
&= \frac{1}{n}\sum_{k=1}^{n-1}\left(\frac{k}{n}\right)^i R_k \left(1 - \left(\frac{k+1}{k}\right)^i\right) + \frac{1}{n}R_n \ , \\
&= \frac{1}{n}\sum_{k=1}^{n-1}\left(\frac{k}{n}\right)^i R_k \left(-\frac{i}{k} - o\left(\frac{1}{k}\right)\right) + \Gamma_\Phi + o(1) \ ,
\end{aligned}$$

since

$$\left(\frac{k+1}{k}\right)^i = 1 + \frac{i}{k} + o\left(\frac{1}{k}\right) \quad and \quad \frac{1}{n}R_n = \Gamma_\Phi + o(1) \ ,$$

then

$$\frac{1}{n^{i+1}}\sum_{k=1}^{n} k^i \Phi_k \Phi_k' = -\frac{i}{n}\sum_{k=1}^{n-1}\left(\frac{k}{n}\right)^i \frac{R_k}{k} + \frac{R_k}{k}o(1) + \Gamma_\Phi + o(1)$$

$$= -\frac{i}{n}\sum_{k=1}^{n-1}\left[\left(\frac{k}{n}\right)^i + o(1)\right](\Gamma_\Phi + o(1)) + \Gamma_\Phi + o(1)$$

$$= \frac{1}{i+1}\Gamma_\Phi \ .$$

Proof of corollary 4.1: We need to prove (4).
The (i, j) element of $\Delta_n \mathbf{P}_n \Delta_n'$ is given by:

$$\mathbf{p}_n(i,j) = \frac{1}{n^{i+j-1}}\sum_{k=1}^{n} k^{i+j-2} u_{k-1}^2 \ , \qquad 1 \le i, j \le r+1$$

hence by (7) and lemma 4.2

$$\mathbf{p}_n(i,j) \xrightarrow{a.s.} \frac{\sigma_u^2}{i+j-1} \ ,$$

$\mathbf{P}$ is an Hilbert matrix, therefore it is positive definite (cf. Choi (1983)). □

Remark 4.3: We can define a least squares estimator for $b_1(n)$ by:

$$\hat{b}_1(n) = \sum_{j=0}^{r} \hat{\beta}_j(n) n^j \ ,$$

and we find the waited result:

$$|\hat{b}_1(n) - b_1(n)|^2 = O\left(\frac{\log(n)}{n}\right) \quad a.s. \tag{10}$$

Possible Extension: The result of corollary 4.1 can be extended to a more general regression model:

$$y_n = \theta_n' \Phi_{n-1} + \varepsilon_n \ ;$$

where

$$\theta_n = \sum_{j=0}^{r} \beta_j n^j \qquad \text{with } \beta_j \in \mathbf{R}^d \ , \qquad 0 \le j \le r \ ;$$

To obtain the result of corollary 4.1 we need the assumption H.1. and Φ_{n-1} is $\mathscr{F}_{n-1}$-measurable satisfying a strong law of large numbers:

$$\frac{1}{n} \sum_{k=1}^{n} \Phi_k \Phi_k' \xrightarrow{a.s.} \Gamma_\Phi > 0 \ ,$$

then

$$\hat{\theta}_n = \sum_{j=0}^{r} \hat{\beta}_j(n) n^j$$

is a strongly consistent estimator of θ_n in the sense that

$$\|\hat{\theta}_n - \theta_n\|^2 = O\left(\frac{\log(n)}{n}\right) \quad a.s.$$

4.2 Sinusoidal Evolution

Here $b_1(n) = A \sin(n\omega + \varphi)$, $\omega \in]-\pi, \pi[- \{0\}$ is given, A and φ two unknown parameters. We can write $b_1(n) = \beta_0 \varphi_{n,0} + \beta_1 \varphi_{n,1}$, where $\beta_0 = A \cos(\varphi)$, $\beta_1 = A \sin(\varphi)$, $\varphi_{n,0} = \sin(n\omega)$, $\varphi_{n,1} = \cos(n\omega)$.
We obtain the following

Corollary 4.4: Under assumptions H.1. and H.2., if we assume also that (u_n) *satisfies*:

a.

$$\sup_n E(|u_n|^\beta / \mathscr{F}_{n-1}) < \infty \qquad \textit{for some } \beta > 2 \ ,$$

b. *for all n,* $E(u_n^2/\mathscr{F}_{n-1}) = \sigma_u^2 > 0$.

Then the estimation errors satisfy

$$\|\tilde{\theta}_n\|^2 = O\left(\frac{\log(n)}{n}\right) \quad a.s. \tag{11}$$

Proof:

$$\mathbf{P}_n = \sum_1^n \begin{pmatrix} \sin^2 k\omega & \sin k\omega \cos k\omega \\ \sin k\omega \cos k\omega & \cos^2 k\omega \end{pmatrix} u_{k-1}^2 \ ,$$

Let $w_k = u_k^2 - \sigma_u^2$, (w_k) is a martingale difference sequence such that

$$\sup_n E(|w_n|^\delta/\mathscr{F}_{n-1}) < \infty \ , \qquad \text{for } \delta = \beta/2 > 1 \ ,$$

then by theorem 3.3.3 of Stout (1974) we get that

$$\frac{1}{n}\sum_{k=1}^n w_k \xrightarrow{a.s.} 0 \ ,$$

consequently

$$\frac{1}{n}\mathbf{P}_n \xrightarrow{a.s.} \frac{\sigma_u^2}{2}\mathbf{I}_2 > 0 \tag{12}$$

hence (4) is proved with $\Delta_n = \frac{1}{\sqrt{n}}\mathbf{I}_2$. □

Remarks and Discussion:

1) Like in paragraph 4.1, we can define the least squares estimator of $b_1(n)$:

$$\hat{b}_1(n) = (\sin n\omega, \cos n\omega)\hat{\theta}_n \ ,$$

and we obtain

$$|\hat{b}_1(n) - b_1(n)|^2 = O\left(\frac{\log(n)}{n}\right) \quad a.s. \tag{13}$$

2) The last corollary can easily be extended to the case

$$b_1(n) = \sum_{i=0}^{r} A_i \sin(n\omega_i + \varphi_i) \ ;$$

where $\omega_i \neq \omega_j$, $\omega_i \in\,]-\pi, \pi[\, - \{0\}$, (A_i, φ_i), $0 \leq i \leq r$, are unknown parameters.

3) The rate $\log(n)/n$ is obtained for every functional basis $(\varphi_{n,i})_{0 \leq i \leq r}$ and input signal (u_n) which satisfy

$$\frac{1}{n} \sum_{k=1}^{n} \varphi_{k,i}\varphi_{k,j}u_{k-1}^2 \xrightarrow{a.s.} \mathbf{P}_{ij} \ .$$

where $\mathbf{P}$ is a positive definite matrix. This result is similar to those obtained in the frame of classical stable models with constant parameters (see for example Chen (1991), Dulfo (1990)) even for ARMA or ARMAX models.

5 Simulation Results

5.1 Polynomial Evolution

We simulate the following model:

$$y_n = (0.5 + 0.002n^2)u_{n-1} + \varepsilon_n \ ,$$

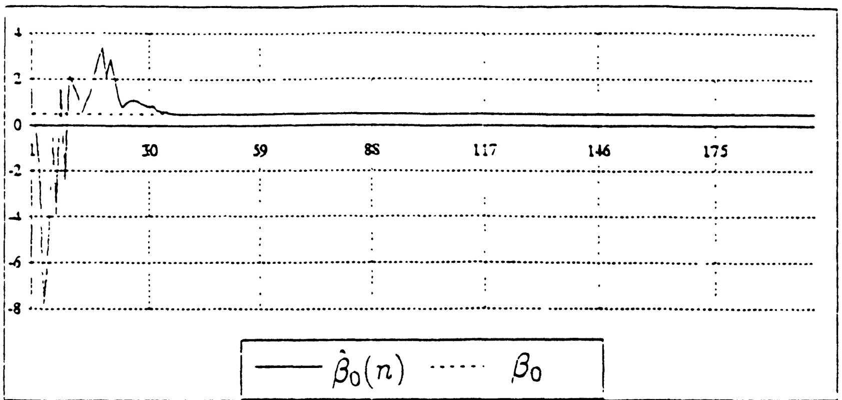

Fig. 1. β_0 and $\hat{\beta}_0(n)$, $n = 0$ to 200

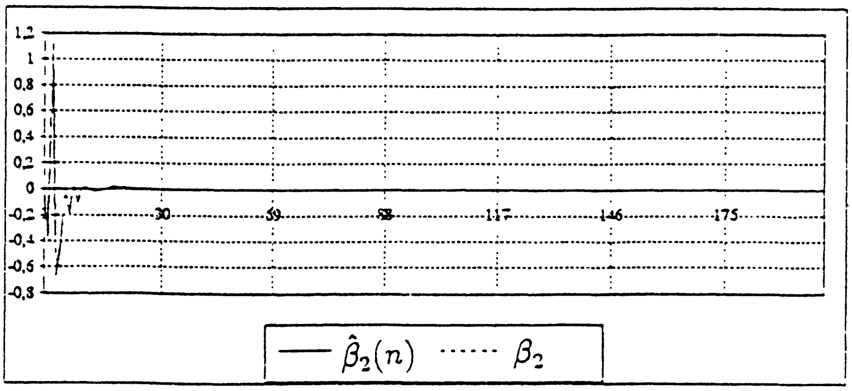

Fig. 2. β_2 and $\hat{\beta}_2(n)$, $n = 0$ to 200

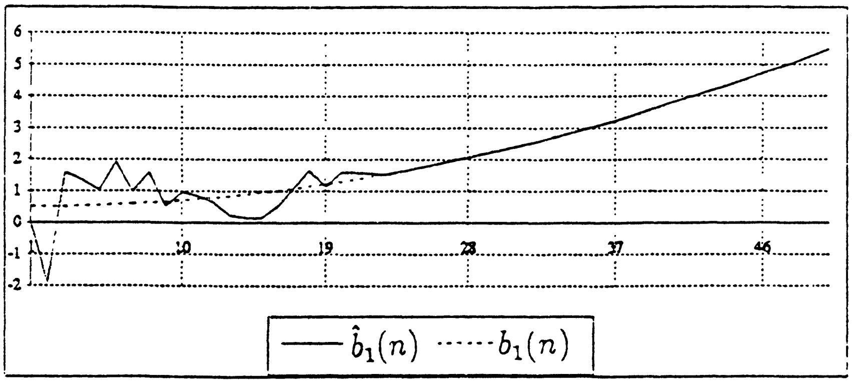

Fig. 3. $b_1(n)$ and $\hat{b}_1(n)$, $n = 0$ to 50

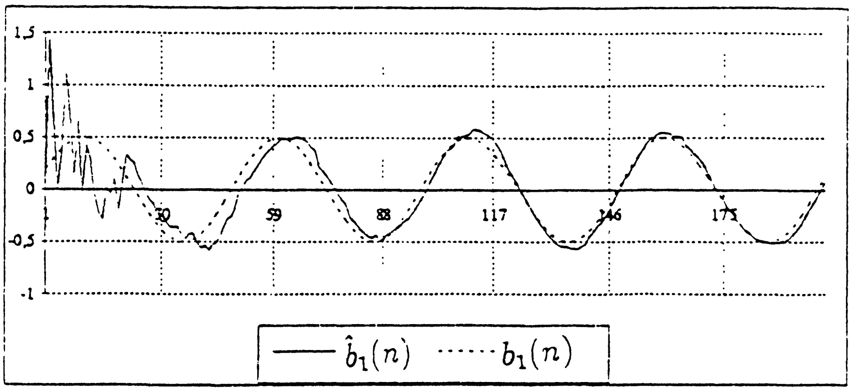

Fig. 4. $b_1(n)$ and $\hat{b}_1(n)$, $n = 0$ to 200

in which (ε_n) and (u_n) are independent gaussian white noises $N(0, 1)$ (hence *H'.1.* and *H.2.* are fulfilled).

Fig. 1 and Fig. 2 allow us to compare the respective rates of convergence of $\hat{\beta}_0(n)$ and $\hat{\beta}_2(n)$. Figure 3 shows the estimates of $b_1(n)$.

5.2 Sinusoidal Evolution

This simulated model is:

$$y_n = \frac{1}{2}\sin\left(\frac{2n\pi}{50} + \frac{\pi}{12}\right)u_{n-1} + \varepsilon_n \ ,$$

with the same assumptions as in 5.1.

Acknowledgements: The authors would like to thank the referee for valuable criticims and suggestions.

References

Boutahar M (1991) Asymptotic distribution of least squares estimates in unstable ARX(p, s) models. Stochastics and Stochastics Reports 37:105–126

Chen HF, Guo L (1991) Identification and stochastic adaptive control. Birkhäuser

Choi MD (1983) Tricks or treats with the hilbert matrix. Ann Math Month 90:301–312

Chow GC (1983) Econometrics. McGraw-Hill

Davidov G, Shavit A, Shinkin N, Koren Y (1992) Estimation of dynamical-varying parameters by internal model principle. IEEE Trans on Autom Contr 37(4):498–503

Duflo M (1990) Méthodes récursives aléatories. Masson

Lai TL, Wei CZ (1982) Least squares estimates in stochastic regression models with applications to identification and control of dynamics systems. The Annals of Statistics 10(1):154–166

Lai TL, Robbin H, Wei CZ (1978) Strong consistency of least squares estimates in Multiple Regression. Proc Nat Acad Sci USA, 3034–3036

Sant DT (1977) Generalized least squares applied to time varying parameter models. An of Econ and Soc Meas 6/3:301–310

Stout WF (1974) Almost sure convergence. Academic Press

Tsypkin YZ, Bondarenko MV (1992) An optimal algorithm for identification of rapidly time varying systems. IEEE Trans on Autom Contr 37 n (2):237–239

Received 28.12.1993
Revised version 07.09.1994