11 (1991), 3, 327—340

Acta Mathematica Scientia
数学物理学报

LMS-LIKE ESTIMATION FOR TIME VARYING PARAMETERS*

H. F. Chen (陈翰馥) *L. Guo* (郭 雷) *J. F. Zhang* (张纪峰)

Inst. of Syst. Sci., Academia Sinica, Beijing, China

Dedicated to the 70th Birthday of Professor Wu Wen-tsün

Abstract

An LMS-like algorithm is applied for estimating the time-varying parameter θ_n in the linear model $y_n=\varphi_n^T\theta_n+v_n$, which is general in the sense that none of the probabilistic properties such as stationarity, Markov property, independence and ergodicity is imposed on any of the processes $\{y_n\}$, $\{\varphi_n\}$, $\{\theta_n\}$ and $\{v_n\}$. It is shown that the α-th moment of the estimation error is of order of the α-th moment of the observation noise and the parameter variation $w_n\triangleq\theta_n-\theta_{n-1}$.

§ 1. Introduction

For linear stochastic systems with constant parameters there has been made a great progress on the parameter estimation problem. Systems with constant parameters may be viewed as a first approximation to the real processes which, as a matter of fact, mostly are time-varying in practice, and such an approximation is not alwayi satisfied by the practitioners. By this reason for recent years a considerable attention has been paid to analysing systems with unknown time-varying parameters by researchers in the areas such as control theory, signal processing and time series analysis.

It is natural to expect that the first set of results on parameter estimation (or tracking) and adaptive control for systems with time-varying parameters is obtained under some statistical law assumptions on the regressor or on the observation noise or on the parameter itself. For example, it is assumed that the regressor is stationary and independent of the observation noise (Macchi, 1986), the observation noise is Gaussian (Kitagava and Gersh, 1985) and the parameter is a Markov process (Chen and Caines, 1990; Guo and Meyn, 1989; Ji and Chizeck, 1988). Without any doubt the assumptions made in the papers mentioned above are reasonable in certain circumstances, but they are restrictive in general. For example, the stationarity assumption on the regressor excludes the feedback control system from consideration, for which the regressor cannot be independent of the noise either.

In this paper we consider the following system with unknown time-varying parameter $\{\theta_k\}$:

* Received Jan. 22, 1991. Work supported by NSFC and TWAS RG MP 898-117.

$$y_k = \varphi_k^\tau \theta_k + v_k, \tag{1.1}$$

where y_k is the one-dimensional system output, φ_k is the r-dimensional regressor and v_k is the system noise.

Clearly, in the special case where

$$\varphi_k^\tau = [y_{k-1} \cdots y_{k-s} \quad u_{k-1} \cdots u_{k-t}]$$

and v_k is a moving average process System (1.1) turns to be the ordinary ARMAX model with time-varying parameters.

The problem stated in this paper is to on-line estimate or to track the time-varying parameter θ_k based on the observed data y_i and φ_i, $i \leqslant k$.

We would like to emphasize that System (1.1) is quite general in the class of linear models: For processes $\{y_k\}$, $\{\varphi_k\}$, $\{\theta_k\}$ and $\{v_k\}$ we do not make any assumption on their statistical relationship and do not require them to be processes of a restricted class such as stationary process, Markov process etc.

We characterize the observation error $\{v_k\}$ and the parameter variation

$$w_n = \theta_n - \theta_{n-1} \tag{1.2}$$

either by

$$\sigma_\alpha \triangleq \sup_{n>0} E(|v_n|^\alpha + \|w_n\|^\alpha) \tag{1.3}$$

or by

$$s_\alpha \triangleq \lim_{n>0} \sup \frac{1}{n} \sum_{i=0}^{n} (|v_i|^\alpha + \|w_i\|^\alpha), \quad \text{a. s.} \tag{1.4}$$

for some constant $\alpha > 0$.

In such a set-up of the problem, Guo (1990) provides a detailed analysis for estimation error $\tilde{\theta}_n \triangleq \theta_n - \hat{\theta}_n$, when $\hat{\theta}_n$ is calculated according to a Kalman-filter-like recursion for which a time-varying matrix adaptation gain is used. Guo's results are then strengthened (Zhang, Guo and Chen 1990). It is shown that α-th moment of the estimation error is of order σ_α or s_α in accordance with the average taken in the sample space sense or in the time domain sense, respectively.

In this paper we establish similar results for the estimate $\hat{\theta}_n$ produced by an LMS-like algorithm, which is characterized by its simplicity for computation:

$$\hat{\theta}_{n+1} = \hat{\theta}_n + \mu_n \varphi_n (y_n - \varphi_n^\tau \hat{\theta}_n), \tag{1.5}$$

where $\{\mu_n\}$ is bounded by a positive constant with $\mu_n \|\varphi_n\|^2 \leqslant 1$, $\forall n \geqslant 0$ and μ_n is measurable with respect to the σ-algebra generated by $\{v_i, \theta_i, \varphi_i, i \leqslant n\}$. It is easy to see that the algorithm (1.5) becomes the well-known LMS adaptive filter with step size μ considered by Widrow et al (1976), Macchi (1986), Bitmead and Anderson (1980), Benveniste et al (1987), if $\mu_k \equiv \mu \in (0, 1)$, and it turns to be the projection (or gradient) algorithm (see, e.g. Anderson et al, 1986; Chen and Guo, 1987) with step size μ if $\mu_k = \dfrac{\mu}{1 + \|\varphi_k\|^2}$.

In comparison with the previous work we note that the tracking error bounds are established when LMS is applied to tracking deterministic time-varying parameter. (Eweda and Macchi, 1985). However, they assumed that the time variation $\|\theta_n - \theta_{n-1}\|$ is bounded in k and $\{y_k, \varphi_k\}$ is an M-dependent sequence. For the constant parameter θ the convergence of LMS estimate is proved under the assumption that φ_k is purely nondeterminisitic and $\{\varphi_k, v_k\}$ is strictly stationary, while for the time-varying parameter an additional assumption is required that $\{\theta_n - \theta_{n-1}\}$ is stationary, zero mean and independent of $\{\varphi_k, v_k\}$ (Solo, 1990).

§ 2. Main Results

Recursively define

$$\Phi(n+1, m) = (I - \mu_n\varphi_n\varphi_n^\tau)\Phi(n, m), \ \Phi(m, m) = I, \ \forall n \geqslant m \geqslant 0. \tag{2.1}$$

From (1.1), (1.2), (1.5) and (2.1) it is easy to see that

$$\tilde{\theta}_{n+1} = (I - \mu_n\varphi_n\varphi_n^\tau)\tilde{\theta}_n + w_{n+1} - \mu_n\varphi_n v_n \tag{2.2}$$

$$= \Phi(n+1, 0)\tilde{\theta}_0 + \sum_{i=0}^{n}\Phi(n+1, i+1)\xi_{i+1} \tag{2.3}$$

where

$$\tilde{\theta}_n \triangleq \theta_n - \hat{\theta}_n \quad \text{and} \quad \xi_{n+1} \triangleq w_{n+1} - \mu_n\varphi_n v_n. \tag{2.4}$$

We always assume that

$$E\|\tilde{\theta}_0\|^\alpha < \infty.$$

From (2.3) we see that the tracking error $\{\tilde{\theta}_n\}$ strongly depends upon the behavior of $\{\Phi(n, m), \forall n \geqslant m \geqslant 0\}$. For analysing (2.3) the essential role is played by the "conditional richness" condition (Guo, 1990) and (Zhang, Guo and Chen, 1990):

There is a nondecreasing sequence $\{\mathscr{F}_n\}$ of σ-algebras such that $\mu_n \in \mathscr{F}_n$, $\varphi_n \in \mathscr{F}_n$ and

$$E\left\{\sum_{k=m+1}^{m+h}\mu_k\varphi_k\varphi_k^\tau \,\Big|\, \mathscr{F}_m\right\} \geqslant \frac{1}{\alpha_m} I \quad \text{a.s.,} \quad \forall m \geqslant 0, \tag{2.5}$$

where h is a positive integer and $\{\alpha_m, \mathscr{F}_m\}$ is an adapted nonnegative sequence satisfying $\alpha_m \geqslant 1$, and

$$\alpha_{m+1} \leqslant a\,\alpha_m + \eta_{m+1}, \quad \forall m \geqslant 0, \ E\alpha_0^{1+\delta} < \infty \tag{2.6}$$

where $a \in (0, 1)$ is a constant and $\{\eta_m, \mathscr{F}_m\}$ is an adapted nonnegative sequence such that

$$\sup_{m \geqslant 0} E[\eta_{m+1}^{1+\delta} \,|\, \mathscr{F}_m] \leqslant M \quad \text{a.s.} \tag{2.7}$$

with $\delta > 0$ and $M < \infty$ being constants.

Remark 1. In the case where $\mu_k = \dfrac{1}{1+\|\varphi_k\|^2}$, Condition (2.5) is an extension of the one introduced in (Guo, 1990). It is noted in (Zhang, Guo and Chen, 1990) that the conditional richness condition (2.5) is satisfied by a large class of processes. For example, if $\{\varphi_k\}$ is a ϕ-mixing process with

$$\inf_k \sup_{\|x\|=1} x^\tau E\varphi_k\varphi_k^\tau x > 0 \quad \text{and} \quad \sup_k E\|\varphi_k\|^4 < \infty,$$

then (2.5) holds. Also, if $\{\varphi_k\}$ is an output of a stable and output-controllable linear system, then (2.5) holds. Finally, (2.5) is obviously satisfied if its determi-