Available online at www.sciencedirect.com

ScienceDirect

Journal of the Franklin Institute 347 (2010) 426–437

Journal of The Franklin Institute

www.elsevier.com/locate/jfranklin

The residual-based ESG algorithm and its performance analysis ☆

Yongsong Xiao[a], Dongqing Wang[b], Feng Ding[a,*]

[a]School of Communication and Control Engineering, Jiangnan University, Wuxi 214122, PR China
[b]College of Automation Engineering, Qingdao University, Qingdao 266071, PR China

Received 16 September 2008; received in revised form 25 March 2009; accepted 20 May 2009

Abstract

The performance of the residual-based extended stochastic gradient (ESG) algorithms for identifying CARMA models with disturbances is analyzed under weaker conditions on statistical properties of the noise. The paper derives the conditions under which the parameter estimation errors converge to zero. Three examples are given to show the advantages of the proposed algorithm.

Keywords: Recursive identification; Parameter estimation; Convergence properties; Stochastic gradient; Stochastic approximation

1. Introduction

The stochastic gradient (SG) or stochastic approximation and least squares (LS) algorithms are two important methods for system identification and parameter estimation. Compared with the least-squares method [1,2], the gradient search method has small computational load and wide applications in many areas, including control and optimization [3–7], function approximation suited to reinforcement learning [8], and option pricing based on stochastic approximation techniques [9].

☆This work was supported by the National Natural Science Foundation of China (60973043).
*Corresponding author. Tel.: +86 510 88637783; fax: +86 510 85910653.
E-mail addresses: ys_xiao82@163.com (Y. Xiao), dqwang64@163.com (D.Q. Wang), fding@jiangnan.edu.cn (F. Ding).

0016-0032/$32.00 © 2009 Published by Elsevier Ltd. on behalf of The Franklin Institute.
doi:10.1016/j.jfranklin.2009.05.008

A great deal of work has been published on convergence analysis of SG [10–14] and LS identification methods [15–17]. In this literature, an auxiliary model-based SG algorithm for dual-rate systems [18], a polynomial transform-based SG algorithm for dual-rate systems [19], a multi-innovation SG algorithm for linear regression systems [20,21] and a hierarchical SG algorithm for multivariable systems [22] were presented. Recently, Ding, Shi and Chen proposed the extended SG algorithm for Hammerstein nonlinear ARMAX systems based on the over-parameterization method [23], and Wang and Ding studied the extended SG algorithm for Hammerstein–Wiener models [1].

Exploring the properties of the SG algorithm under weaker conditions is still open [14] and also the goal in this paper. This paper studies the identification problem for the CARMA models with a deterministic disturbance and presents a residual-based extended stochastic gradient (ESG) algorithms and analyzes the performance of the proposed ESG algorithm by assuming that the process variables have non-zero mean values and time-varying variances. The proposed algorithm has less computational load than the least-squares algorithm and can improve the accuracy of the parameter estimates by introducing the forgetting factor.

This paper is organized as follows. Section 2 discusses the problem formulation. Section 3 analyzes the performance of the extended SG algorithm. Section 4 provides an illustrative example. Finally, concluding remarks are given in Section 5.

2. The system description

Consider the discrete-time system described by a CARMA model with a disturbance [14],

$$A(z)y'(t) = B(z)u'(t) + D(z)v'(t) + f', \tag{1}$$

where $\{y'(t)\}$ and $\{u'(t)\}$ are the system input and output sequences with mean values $E[y'(t)]=\mu_y$ and $E[u'(t)]=\mu_u$, $\{v'(t)\}$ is a random noise sequence with mean value $E[v'(t)]=\mu_v$ and time-varying variance $\sigma^2(t)$, f' is a deterministic disturbance, z^{-1} represents the unit backward shift operator: $z^{-1}y(t)=y(t-1)$, and $A(z)$, $B(z)$ and $D(z)$ are polynomials in z^{-1} with

$$\begin{aligned} A(z) &= 1 + a_1z^{-1} + a_2z^{-2} + \cdots + a_{n_a}z^{-n_a}, \\ B(z) &= b_1z^{-1} + b_2z^{-2} + \cdots + b_{nb}z^{-n_b}, \\ D(z) &= 1 + d_1z^{-1} + d_2z^{-2} + \cdots + d_{n_d}z^{-n_d}. \end{aligned}$$

Let

$$\begin{aligned} y(t) &:= y'(t) - \mu_y, \\ u(t) &:= u'(t) - \mu_u, \\ v(t) &:= v'(t) - \mu_v, \\ f &:= f' + B(1)\mu_u + D(1)\mu_v = A(1)\mu_y. \end{aligned} \tag{2}$$

Then $y(t)$, $u(t)$ and $v(t)$ have zero mean values. Eq. (1) can be rewritten as [14]

$$A(z)y(t) = B(z)u(t) + D(z)v(t) + f. \tag{3}$$

Without loss of generality, assume that $y(t)=0$, $u(t)=0$ and $v(t)=0$ for $t\leq 0$.