Computers and Mathematics with Applications 58 (2009) 1190–1197

Contents lists available at ScienceDirect

Computers and Mathematics with Applications

journal homepage: www.elsevier.com/locate/camwa

The residual based interactive least squares algorithms and simulation studies

Yongsong Xiao [a,*], Yong Zhang [b], Jie Ding [a,*], Jiyang Dai [c]

[a] School of Communication and Control Engineering, Jiangnan University, Wuxi 214122, PR China

[b] School of Information Engineering, Inner Mongolia University of Science and Technology, Baotou 014010, PR China

[c] Key Laboratory of Nondestructive Testing (Ministry of Education), Nanchang Hangkong University, Nanchang 330063, PR China

ARTICLE INFO

Article history:
Received 6 March 2008
Received in revised form 15 February 2009
Accepted 27 February 2009

Keywords:
Recursive identification
Parameter estimation
Least squares
Controlled autoregressive moving average (C-ARMA) models
AR models
MA models

ABSTRACT

This paper presents a two-stage least squares based iterative algorithm, a residual based interactive least squares algorithm and a residual based recursive least squares algorithm for identifying controlled autoregressive moving average (C-ARMA) models. The simulation studies indicate that the proposed algorithms can effectively estimate the parameters of the C-ARMA models.

1. Introduction

The time series contains three basic models: autoregressive (AR) model, moving average (MA) model and autoregressive moving average (ARMA) model. This paper considers the least squares identification problem of controlled autoregressive moving average (C-ARMA) model.

Compared with least squares (LS) algorithm, the stochastic gradient (SG) algorithm has small computational load but slow convergence rate [1]. Recently, Ding, Yang and Liu analyzed the consistency of the multivariable SG algorithm [2]; Ding and Chen presented a hierarchical SG algorithm for multivariable systems [3] and an auxiliary model based SG algorithm for dual-rate systems [4]; Ding et al. studied the performances of the SG algorithms for dual-rate systems based on the polynomial transformation technique [5,6]. In order to improve the convergence rate of the SG algorithm, Ding and Chen developed a multi-innovation SG identification algorithm for linear regression model [7] and an extended stochastic gradient algorithm with a forgetting factor for Hammerstein nonlinear systems [8]; Ding and Wang discussed the gradient based identification algorithm for Hammerstein–Wiener ARMAX systems [9]. Finally, Zhang, Ding and Shi presented a multi-innovation SG parameter estimation based self-tuning control algorithm [10].

Because of the fast convergence rate of the least squares identification, it has received much attention in many areas, including signal processing [11], system identification and parameter estimation [12–19], adaptive control [5,20,21]. For example, Ding and Chen presented a hierarchical LS algorithm for multivariable systems [22], whose consistency was studied in [23]; Ding and Chen proposed an auxiliary model based LS algorithm for dual-rate systems [24]; Ding, Liu and Shi studied the performances of the polynomial transform based LS algorithm for dual-rate systems [25].

* Corresponding author.
E-mail addresses: ys_xiao82@163.com (Y. Xiao), djvsqxb@163.com (J. Ding).

0898-1221/$ – see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.camwa.2009.02.037

2. The system description

Consider a controlled autoregressive moving average (C-ARMA) model,

$$y(t) = B(z)u(t) + \frac{D(z)}{C(z)}v(t), \tag{1}$$

where $\{y(t)\}$ is the observation output, $\{u(t)\}$ is the control input, $\{v(t)\}$ is a white noise sequence with zero mean $\mathrm{E}[v(t)] = 0$ and variance $\sigma^2 := \mathrm{E}[v^2(t)]$ (E denotes the expectation operator) and independent of $u(t)$, z^{-1} represents the unit backward shift operator: $z^{-1}y(t) = y(t-1)$, and $B(z)$, $C(z)$ and $D(z)$ are polynomials in z^{-1} with

$$B(z) = b_1z^{-1} + b_2z^{-2} + \cdots + b_{n_b}z^{-n_b},$$
$$C(z) = 1 + c_1z^{-1} + c_2z^{-2} + \cdots + c_{n_c}z^{-n_c},$$
$$D(z) = 1 + d_1z^{-1} + d_2z^{-2} + \cdots + d_{n_d}z^{-n_d}.$$

Assume that the orders n_b, n_c and n_d are known. Let

$$w(t) := \frac{D(z)}{C(z)}v(t). \tag{2}$$

$w(t)$ is a correlated noise with zero mean value and variance:

$$\sigma_w^2 = \frac{\sigma^2}{2\pi \mathrm{j}} \oint_\gamma \frac{D(z)}{C(z)} \frac{D(z^{-1})}{C(z^{-1})} \frac{\mathrm{d}z}{z}, \quad \mathrm{j} = \sqrt{-1},$$

where γ is a unit circle with $|\gamma| = 1$.

Define the parameter vectors $\boldsymbol{\theta}$ and $\boldsymbol{\vartheta}$ and information vectors $\boldsymbol{\varphi}(t)$, $\boldsymbol{\phi}(t)$ and $\boldsymbol{\psi}(t)$ as

$$\begin{aligned}
&\boldsymbol{\theta} := \begin{bmatrix}\boldsymbol{b}\\ \boldsymbol{\vartheta}\end{bmatrix} \in \mathbb{R}^{n_b+n_c+n_d}, \qquad \boldsymbol{\varphi}(t) := \begin{bmatrix}\boldsymbol{\phi}(t)\\ \boldsymbol{\psi}(t)\end{bmatrix} \in \mathbb{R}^{n_b+n_c+n_d},\\
&\boldsymbol{b} := [b_1, b_2, \ldots, b_{n_b}]^{\mathrm{T}} \in \mathbb{R}^{n_b},\\
&\boldsymbol{\vartheta} := [c_1, c_2, \ldots, c_{n_c}, d_1, d_2, \ldots, d_{n_d}]^{\mathrm{T}} \in \mathbb{R}^{n_c+n_d},\\
&\boldsymbol{\phi}(t) := [u(t-1), u(t-2), \ldots, u(t-n_b)]^{\mathrm{T}} \in \mathbb{R}^{n_b},\\
&\boldsymbol{\psi}(t) := [-w(t-1), -w(t-2), \ldots, -w(t-n_c), v(t-1), v(t-2), \ldots, v(t-n_d)]^{\mathrm{T}} \in \mathbb{R}^{n_c+n_d},
\end{aligned} \tag{3}$$

where the superscript T represents the transpose of a matrix or vector.

From (2) and (1), we have the following identification models:

$$w(t) = \boldsymbol{\psi}^{\mathrm{T}}(t)\boldsymbol{\vartheta} + v(t), \tag{4}$$

$$y(t) = \boldsymbol{\phi}^{\mathrm{T}}(t)\boldsymbol{b} + w(t) \tag{5}$$

$$= \boldsymbol{\phi}^{\mathrm{T}}(t)\boldsymbol{b} + \boldsymbol{\psi}^{\mathrm{T}}(t)\boldsymbol{\vartheta} + v(t)$$

$$= \boldsymbol{\varphi}^{\mathrm{T}}(t)\boldsymbol{\theta} + v(t). \tag{6}$$

Note that $y(t)$ and $\boldsymbol{\phi}(t)$ are measured but $w(t)$, $\boldsymbol{\psi}(t)$ and $v(t)$ are unknown (unmeasured). The objective of this paper is to present the residual based identification algorithms to estimate the parameter vectors $\boldsymbol{\theta}$ or $\boldsymbol{b}$ and $\boldsymbol{\vartheta}$ from the observation data $\{u(t), y(t)\}$ or $\{y(t), \boldsymbol{\phi}(t)\}$.

3. The least squares estimate of *b*

Define the stacked vectors and matrices:

$$\boldsymbol{Y}(t) := \begin{bmatrix} y(1)\\ y(2)\\ \vdots\\ y(t)\end{bmatrix} \in \mathbb{R}^t, \qquad \boldsymbol{H}(t) := \begin{bmatrix} \boldsymbol{\phi}^{\mathrm{T}}(1)\\ \boldsymbol{\phi}^{\mathrm{T}}(2)\\ \vdots\\ \boldsymbol{\phi}^{\mathrm{T}}(t)\end{bmatrix} \in \mathbb{R}^{t\times n_b}, \tag{7}$$

$$\boldsymbol{W}(t) := \begin{bmatrix} w(1)\\ w(2)\\ \vdots\\ w(t)\end{bmatrix} \in \mathbb{R}^t, \qquad \boldsymbol{\Psi}(t) := \begin{bmatrix} \boldsymbol{\psi}^{\mathrm{T}}(1)\\ \boldsymbol{\psi}^{\mathrm{T}}(2)\\ \vdots\\ \boldsymbol{\psi}^{\mathrm{T}}(t)\end{bmatrix} \in \mathbb{R}^{t\times(n_c+n_d)}, \qquad \boldsymbol{V}(t) := \begin{bmatrix} v(1)\\ v(2)\\ \vdots\\ v(t)\end{bmatrix} \in \mathbb{R}^t.$$