Research Article

Multi-Innovation Stochastic Gradient Identification Algorithm for Hammerstein Controlled Autoregressive Autoregressive Systems Based on the Key Term Separation Principle and on the Model Decomposition

Huiyi Hu, 1 Xiao Yongsong, 1 and Rui Ding 2

1 Key Laboratory of Advanced Process Control for Light Industry (Ministry of Education), Jiangnan University, Wuxi 214122, China
2 School of Internet of Things Engineering, Jiangnan University, Wuxi 214122, China

Correspondence should be addressed to Rui Ding; rding12@126.com

Received 10 June 2013; Revised 22 August 2013; Accepted 6 September 2013

Academic Editor: Reinaldo Martinez Palhares

Copyright © 2013 Huiyi Hu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

An input nonlinear system is decomposed into two subsystems, one including the parameters of the system model and the other including the parameters of the noise model, and a multi-innovation stochastic gradient algorithm is presented for Hammerstein controlled autoregressive autoregressive (H-CARAR) systems based on the key term separation principle and on the model decomposition, in order to improve the convergence speed of the stochastic gradient algorithm. The key term separation principle can simplify the identification model of the input nonlinear system, and the decomposition technique can enhance computational efficiencies of identification algorithms. The simulation results show that the proposed algorithm is effective for estimating the parameters of IN-CARAR systems.

1. Introduction

There exist many nonlinear systems in process control [1–3]. A nonlinear system can be modeled by input nonlinear systems [4] and output nonlinear systems [5], input-output nonlinear systems [6], feedback nonlinear systems [7], and so on. Input nonlinear systems, which are called Hammerstein systems [8], include input nonlinear equation error type systems and input nonlinear output error type systems. Recently, many identification algorithms have been developed for input nonlinear systems, such as the iterative methods [9–11], the separable least squares methods [12, 13], the blind methods [14], the subspace methods [15], and the overparameterization methods [16, 17]. Some methods require paying much extra computation.

The stochastic gradient (SG) algorithm is widely applied to parameter estimation. For example, Wang and Ding presented an extended SG identification algorithm for Hammerstein-Wiener ARMAX systems [18], but it is well known that the SG algorithm has slower convergence rates. In order to improve the convergence rate of the SG algorithm, Xiao et al. presented a multi-innovation stochastic gradient parameter estimation algorithm for input nonlinear controlled autoregressive (IN-CAR) models using the over-parameterization method [19]; Chen et al. proposed a modified stochastic gradient algorithm by introducing a convergence index in order to improve the convergence speed of the parameter estimation [20]; Han and Ding developed a multi-innovation stochastic gradient algorithm for multi-input single-output systems [21]; Liu et al. studied the performance of the stochastic gradient algorithm for multivariable systems [22].

The decomposition identification techniques include matrix decomposition and model decomposition. Hu and Ding presented a least squares based iterative identification algorithm for controlled moving average systems using the matrix decomposition [23]; Ding derived an iterative least squares algorithm to estimate the parameters of output error systems, and the matrix decomposition can enhance computational efficiencies [24]. Ding also divided a Hammerstein nonlinear system into two subsystems based on the
model decomposition and presented a hierarchical multi-innovation stochastic gradient algorithm for Hammerstein nonlinear systems [25].

This paper discusses identification problems of input nonlinear controlled autoregressive (IN-CARAR) systems or Hammerstein controlled autoregressive autoregressive (H-CARAR) systems, which is one kind of input nonlinear equation error type systems. The basic idea is using the key term separation principle [26] and the decomposition technique [24] to derive a multi-innovation stochastic gradient identification algorithm, which is different from the work in [19, 21, 25].

The rest of this paper is organized as follows. Section 2 gives the identification model for the IN-CARAR systems. Section 3 introduces the SG algorithm for the IN-CARAR system. Section 4 deduces a multi-innovation SG algorithm for Hammerstein nonlinear controlled autoregressive autoregressive (H-CARAR) systems, which is one kind of input nonlinear controlled autoregressive autoregressive (IN-CARAR) system, that is, an input nonlinear controlled autoregressive autoregressive (IN-CARAR) system or Hammerstein controlled autoregressive autoregressive (H-CARAR) system, which consists of a nonlinear block and a linear dynamic subsystem. It is worth noting that Xiao and Yue discussed a stochastic noise with zero mean. Here,

\[ u(t) = \sum_{i=1}^{n_c} c_i w(t-i) + v(t) \]

Assume \( y(t) = 0, u(t) = 0, \) and \( v(t) = 0 \) for \( t \leq 0. \) \( a_i, b_j, \) and \( c_i \) are the parameters to be estimated from measured input-output data \( \{u(t), y(t)\} \).

Define the parameter vectors:

\[
\theta := [\theta_n^T, \theta_s^T]^T \in \mathbb{R}^{n_s+n_l+n_v},
\]

\[
\varphi(t) := \begin{bmatrix} \phi_n(t) \\ \phi_s(t) \end{bmatrix} \in \mathbb{R}^{n_s+n_l+n_v},
\]

Equation (2) can be written as

\[
w(t) = [1 - C(z)] w(t) + v(t)
\]

where \( f(u(t)) := [f_1(u(t)), f_2(u(t)), \ldots, f_{n_l}(u(t))] \in \mathbb{R}^{1 \times n_l} \) and

\[
\gamma := [\gamma_1, \gamma_2, \ldots, \gamma_{n_l}]^T \in \mathbb{R}^{n_l} \text{ is the parameters vector of the nonlinear part.}
\]
Using the key term separation principle [26], (1) can be written as
\[ y(t) = [1 - A(z)] y(t) + [B(z) - 1] \bar{u}(t) + \bar{w}(t) \]
\[ = - \sum_{i=1}^{n_1} \alpha_i y(t-i) + \sum_{i=1}^{n_2} b_i \bar{u}(t-i) + \sum_{i=1}^{n_3} \gamma_i \hat{f}(u(t)) + w(t) \]
\[ = q_n^T(t) \theta_n + w(t) \]
\[ = q_n^T(t) \theta_n + \nu(t) \]
\[ = q_n^T(t) \theta_n + v(t) \]
(9)
(10)
(11)
This is the identification model of the IN-CARAR system.

The Stochastic Gradient Algorithm

According to [1] and based on the identification model in (11), we can obtain the stochastic gradient (SG) algorithm:
\[ \hat{\theta}(t) = \hat{\theta}(t-1) + F(t) e(t), \]
\[ e(t) = y(t) - \tilde{\phi}^T(t) \hat{\theta}(t-1), \]
\[ r(t) = r(t-1) + \|\phi(t)\|^2, \quad r(0) = 1, \]
\[ \tilde{\phi}(t) = \begin{bmatrix} \varphi_1(t) \\ \varphi_n(t) \end{bmatrix}, \]
\[ \tilde{\phi}_1(t) = \begin{bmatrix} -y(t-1), -y(t-2), \ldots, -y(t-n_0) \end{bmatrix}^T, \]
\[ \tilde{\phi}_n(t) = \begin{bmatrix} -\bar{w}(t-1), -\bar{w}(t-2), \ldots, -\bar{w}(t-n_c) \end{bmatrix}^T, \]
\[ \bar{u}(t) = f(u(t)) \tilde{y}(t), \]
\[ \tilde{y}(t) = y(t) - \tilde{\phi}_n^T(t) \hat{\theta}_n(t), \]
\[ f(u(t)) = [\varphi_1(u(t)), \varphi_2(u(t)), \ldots, \varphi_n(u(t))]^T, \]
\[ \hat{\theta}(t) = \begin{bmatrix} \hat{\theta}_n(t) \\ \hat{\theta}_1(t) \end{bmatrix}, \]
\[ \hat{\theta}_1(t) = \begin{bmatrix} \bar{a}^T(t), \bar{b}^T(t), \bar{v}^T(t) \end{bmatrix}^T, \]
\[ \bar{a}(t) = \begin{bmatrix} \bar{a}_1(t), \bar{a}_2(t), \ldots, \bar{a}_n(t) \end{bmatrix}^T, \]
\[ \bar{b}(t) = \begin{bmatrix} \bar{b}_1(t), \bar{b}_2(t), \ldots, \bar{b}_n(t) \end{bmatrix}^T, \]
\[ \bar{v}(t) = \begin{bmatrix} \bar{v}_1(t), \bar{v}_2(t), \ldots, \bar{v}_n(t) \end{bmatrix}^T, \]
\[ \bar{\theta}_n(t) = \begin{bmatrix} \bar{c}_1(t), \bar{c}_2(t), \ldots, \bar{c}_n(t) \end{bmatrix}^T, \]
(12)
where \( \bar{X}(t) \) represents the estimate of \( X \) at time \( t \); for example,
\[ \hat{\theta}(t) = \begin{bmatrix} \hat{\theta}_1(t) \\ \hat{\theta}_n(t) \end{bmatrix} \in \mathbb{R}^{n_1+n_2+n_3} \] is the estimate of \( \theta = [\varphi_n] \) at time \( t \).

4. The Multi-Innovation Stochastic Gradient Algorithm

This section deduces the multi-innovation stochastic gradient identification algorithm for the IN-CARAR system using the decomposition technique [1].

Define two intermediate variables,
\[ y_1(t) := y(t) - \varphi_n^T(t) \theta_n, \]
\[ y_2(t) := y(t) - \varphi_n^T(t) \theta_n, \]
(13)
(14)
(15)
These two subsystems include the parameter vectors \( \theta_1 \) and \( \theta_n \), respectively; \( \theta_1 \) contains the parameters of the system model and \( \theta_n \) contains the parameters of the noise model.

Define the stacked information matrices and the stacked output vectors:
\[ Y(p,t) := [y(t), y(t-1), \ldots, y(t-p+1)]^T \in \mathbb{R}^p, \]
\[ Y_1(p,t) := [y_1(t), y_1(t-1), \ldots, y_1(t-p+1)]^T \in \mathbb{R}^p, \]
\[ Y_2(p,t) := [y_2(t), y_2(t-1), \ldots, y_2(t-p+1)]^T \in \mathbb{R}^p, \]
\[ \Phi_n(p,t) := \begin{bmatrix} \varphi_1(t), \varphi_2(t), \ldots, \varphi_n(t-p+1) \end{bmatrix} \in \mathbb{R}^{p \times (n_1+n_2+n_3)}, \]
\[ \Phi_n(p,t) := \begin{bmatrix} \varphi_1(t), \varphi_2(t), \ldots, \varphi_n(t-p+1) \end{bmatrix} \in \mathbb{R}^{p \times (n_1+n_2+n_3)}. \]
(16)
(17)
According to the multi-innovation identification theory [29–41], we expand the scalar innovations:
\[ e_1(t) = y_1(t) - \varphi_1^T(t) \hat{\theta}_1(t-1), \]
\[ e_n(t) = y_2(t) - \varphi_n^T(t) \hat{\theta}_n(t-1), \]
(18)
to the innovation vectors,
\[ E_1(p,t) = \bar{Y}_1(p,t) - \Phi_n^T(p,t) \hat{\theta}_1(t-1), \]
\[ E_n(p,t) = \bar{Y}_2(p,t) - \Phi_n^T(p,t) \hat{\theta}_n(t-1). \]
(19)
Define two criterion functions,
\[ J_1(\theta_1) := \|Y_1(p,t) - \Phi_n^T(p,t) \theta_1\|^2, \]
\[ J_2(\theta_n) := \|Y_2(p,t) - \Phi_n^T(p,t) \theta_n\|^2. \]
The gradients of $J_1$ and $J_2$ with respect to $\theta_n$ and $\theta_n$, respectively, are
\[
\text{grad } [J_1 (\theta_n)] = \frac{\partial f_1 (\theta_n)}{\partial \theta_n},
\]
\[
= -2\Phi_2 (p, t) \left[ Y_1 (p, t) - \Phi_1^T (p, t) \theta_1 \right],
\]
\[
\text{grad } [J_2 (\theta_n)] = \frac{\partial f_2 (\theta_n)}{\partial \theta_n},
\]
\[
= -2\Phi_3 (p, t) \left[ Y_2 (p, t) - \Phi_3^T (p, t) \theta_2 \right].
\]
(19)

Minimizing $J_1 (\theta_1)$ and $J_2 (\theta_2)$ using the negative gradient search, we can obtain the multi-innovation stochastic gradient algorithm (MISG) for the IN-CARAR system:
\[
\bar{\hat{\theta}}_1 (t) = \hat{\theta}_1 (t-1) + \frac{\Phi_1 (p, t)}{r_1 (t)} E_1 (p, t),
\]
\[
E_1 (p, t) = Y_1 (p, t) - \Phi_1^T (p, t) \hat{\theta}_1 (t-1)
\]
\[
= Y (p, t) - \Phi_1^T (p, t) \hat{\theta}_1 (t-1) - \Phi_1^T (p, t) \hat{\theta}_1 (t-1),
\]
\[
r_1 (t) = r_1 (t-1) + \left\| \Phi_1 (p, t) \right\|^2, \quad r_1 (0) = 1,
\]
\[
Y (p, t) = \left[ y(t), y(t-1), \ldots, y(t-p+1) \right]^T,
\]
\[
\Phi_1 (p, t) = \left[ \Phi_1 (t), \Phi_1 (t-1), \ldots, \Phi_1 (t-p+1) \right]^T,
\]
\[
\hat{\theta}_1 (t) = \hat{\theta}_1 (t-1) + \frac{\Phi_1 (p, t)}{r_2 (t)} E_1 (p, t),
\]
\[
E_2 (p, t) = Y_2 (p, t) - \Phi_2^T (p, t) \hat{\theta}_2 (t-1)
\]
\[
= Y (p, t) - \Phi_2^T (p, t) \hat{\theta}_2 (t-1) - \Phi_2^T (p, t) \hat{\theta}_2 (t-1),
\]
\[
r_2 (t) = r_2 (t-1) + \left\| \Phi_2 (p, t) \right\|^2, \quad r_2 (0) = 1,
\]
\[
\Phi_2 (p, t) = \left[ \Phi_2 (t), \Phi_2 (t-1), \ldots, \Phi_2 (t-p+1) \right]^T,
\]
\[
\hat{\theta}_2 (t) = \hat{\theta}_2 (t-1) + \frac{\Phi_2 (p, t)}{r_2 (t)} E_2 (p, t),
\]
\[
E_3 (p, t) = Y_3 (p, t) - \Phi_3^T (p, t) \hat{\theta}_3 (t-1)
\]
\[
= Y (p, t) - \Phi_3^T (p, t) \hat{\theta}_3 (t-1) - \Phi_3^T (p, t) \hat{\theta}_3 (t-1),
\]
\[
r_3 (t) = r_3 (t-1) + \left\| \Phi_3 (p, t) \right\|^2, \quad r_3 (0) = 1,
\]
\[
\Phi_3 (p, t) = \left[ \Phi_3 (t), \Phi_3 (t-1), \ldots, \Phi_3 (t-p+1) \right]^T,
\]
\[
\hat{\theta}_3 (t) = \hat{\theta}_3 (t-1) + \frac{\Phi_3 (p, t)}{r_3 (t)} E_3 (p, t),
\]
\[
E_n (p, t) = Y_n (p, t) - \Phi_n^T (p, t) \hat{\theta}_n (t-1)
\]
\[
= Y (p, t) - \Phi_n^T (p, t) \hat{\theta}_n (t-1) - \Phi_n^T (p, t) \hat{\theta}_n (t-1),
\]
\[
r_n (t) = r_n (t-1) + \left\| \Phi_n (p, t) \right\|^2, \quad r_n (0) = 1,
\]
\[
\Phi_n (p, t) = \left[ \Phi_n (t), \Phi_n (t-1), \ldots, \Phi_n (t-p+1) \right]^T,
\]
\[
\hat{\theta}_n (t) = \hat{\theta}_n (t-1) + \frac{\Phi_n (p, t)}{r_n (t)} E_n (p, t),
\]
(20)
The initial values can be taken to be $\hat{\theta}(0) = 1_{n_1,n_2,n_3,n_4}/p_0$, $\tilde{w}(i) = 1/p_0$, $i \leq 0$, and $p_0 = 10^6$.

5. Numerical Examples

Consider the following IN-CARAR system:

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

$$A(z) = 1 + a_1 z^{-1} + a_2 z^{-2} = 1 + 1.80 z^{-1} + 0.80 z^{-2},$$

$$B(z) = 1 + b_1 z^{-1} + b_2 z^{-2} = 1 + 0.50 z^{-1} + 0.65 z^{-2},$$

$$C(z) = 1 + c_1 z^{-1} + c_2 z^{-2} = 1 + 0.30 z^{-1} + 0.20 z^{-2},$$

$$\bar{u}(t) = f(u(t)) y = y_1 f_1(u(t)) + y_2 f_2(u(t)) + y_3 f_3(u(t))$$

$$= 1.00 u(t) + 0.50 u^2(t) + 0.25 u^3(t),$$

$$\theta = [a_1, a_2, b_1, b_2, y_1, y_2, y_3, c_1, c_2]^T$$

$$= [1.80, 0.80, 0.5, 0.65, 1.00, 0.50, 0.25, 0.30, 0.20]^T. \quad (21)$$

In this example, the input $\{u(t)\}$ is taken as a persistent excitation signal sequence with zero mean and unit variance and $\{v(t)\}$ as a white noise sequence with zero mean and variance $\sigma^2 = 0.50^2$. Applying the SG algorithm and the MISG algorithm to estimate the parameters of this IN-CARAR system, the parameter estimates and their estimation errors are shown in Tables 1 and 2 with the data length $L = 3000$ and the estimation error $\delta := \|\hat{\theta}(t) - \theta\|/\|\theta\|$ versus $t$ being shown in Figures 2 and 3.

From Tables 1 and 2 and Figures 2 and 3, we can draw the conclusions. The parameters estimation errors become smaller with the data length $t$ increasing. The estimation errors given by the MISG algorithm are much smaller than that of the SG algorithm. The convergence speed of the multi-innovation SG algorithm is faster than those of the SG algorithm. These indicate that the MISG algorithm has better performance than the SG algorithm.

6. Conclusions

The gradient and least squares algorithms are two different kinds of important identification methods. It is well known that the gradient algorithm has poor convergence rates. This paper studies the multi-innovation SG identification methods for IN-CARAR systems. The numerical examples show that the proposed MISG algorithm can estimate effectively the parameters of input nonlinear systems and indicate that increasing the innovation length can improve parameter estimation accuracy of the multi-innovation identification algorithm because the algorithm uses more information in each recursion for a large innovation length. The proposed method can be applied to nonlinear output error systems. Although the algorithm is presented for the IN-CARAR systems, the basic idea can be extended to other linear or nonlinear systems with colored noises [42–61].

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China (nos. 61273194, 61203111), the University Graduate Scientific Research Innovation Program of Jiangsu Province (CXLX13-738), the Fundamental Research Funds for the Central Universities (JUDCF13035), the PAPD of Jiangsu Higher Education Institutions.

References


Submit your manuscripts at http://www.hindawi.com