European Journal of Control 20 (2014) 14–23

Contents lists available at ScienceDirect

European Journal of Control

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

Recursive identification of errors-in-variables Wiener–Hammerstein systems

Bi-Qiang Mu, Han-Fu Chen *

Key Laboratory of Systems and Control, Institute of Systems Science, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, PR China

ARTICLE INFO

Article history:
Received 5 July 2013
Received in revised form
8 October 2013
Accepted 16 October 2013
Recommended by Astolfi Alessandro
Available online 26 October 2013

Keywords:
Wiener–Hammerstein systems
Errors-in-variables
Stochastic approximation
Recursive estimation
Strong consistency

ABSTRACT

This paper considers the recursive identification of errors-in-variables Wiener–Hammerstein system, which is composed of a static nonlinearity sandwiched by two linear dynamic subsystems. Both the system input and output are observed with additive noises being ARMA processes with unknown coefficients. By the stochastic approximation algorithms incorporated with the deconvolution kernel functions, the coefficients of the linear subsystems and the values of the nonlinear function are recursively estimated. All the estimates are proved to converge to the true values with probability one. A simulation example is given to verify the theoretical analysis.

1. Introduction

The block-oriented systems [14] are widely applied to model the practical nonlinear systems owing to their simple structure and excellent modeling ability. The Wiener–Hammerstein system composed of two dynamic linear subsystems with a static nonlinear function in between has a great flexibility for modeling practical systems, for example, sensor systems, electromechanical systems in robotics, mechatronics, biological and chemical systems, and others. The well-studied Hammerstein and Wiener systems can be thought of the special cases of the Wiener–Hammerstein system. Thus, the identification issue of Wiener–Hammerstein systems has received a considerable attention from both theoretical researchers and engineers.

In the early literature [5,3,16] on identification of the Wiener–Hammerstein system, the impulse responses of the two linear subsystems are connected with the correlation functions between the system input and output under the Gaussian input. Based on the maximum likelihood method, a time domain identification algorithm is proposed in [6], and a simple recursive identification technique for multi-input single-output Wiener–Hammerstein system is presented in [4] with the help of a weighted extended least squares method. Some recent work can be found in [25,29,12,18], and among others.

To identify the nonlinear function in a Wiener–Hammerstein system there are parametric [2–4,6,25] and nonparametric approaches [18,15], according to the different descriptions of the nonlinear function. The parametric approach is applied when the nonlinear function is expressed as a linear combination of basis functions such as polynomials, cubic splines functions, piecewise linear functions, neural networks with unknown coefficients, etc. In this case identification turns out to be a parameter estimation problem that can be solved by the standard optimization method such as the gradient method, Newton-Raphson method, and others. The nonparametric approach is used to estimate the values of the nonlinear function at any given point with the help of the kernel functions, requiring no structural information about the nonlinearity. For this there have been some literature [23,22] dealing with nonparametric regression by stochastic approximation involving the kernel functions. Likewise, we adopt the nonparametric method in the paper. To be specific, the stochastic approximation and the deconvolution kernel functions are together used to achieve this. Here we consider the case where the input and output of the system are not accurately available, but they are observed with additive noises, i.e., we intend to identify the errors-in-variables (EIV) Wiener–Hammerstein systems.

There exist some papers on identifiability [1] and identification [24] of the linear EIV systems. Various estimation methods for identifying linear EIV systems, for example, the instrumental variables based methods, the bias-compensation approaches, the Frisch scheme, the frequency domain methods, the prediction error and the ML methods, are well summarized in the survey paper [24], but the methods mentioned there are nonrecursive. The recursive identification for the linear EIV systems is considered under different assumptions on the system input and on the observation noise in [26,8,30,18]. There are also a few papers [28,17,19] on the identification of nonlinear EIV systems.

* Corresponding author. Tel.: +86 10 62579540.
E-mail address: hfchen@iss.ac.cn (H.-F. Chen).

0947-3580/$ - see front matter
http://dx.doi.org/10.1016/j.ejcon.2013.10.005

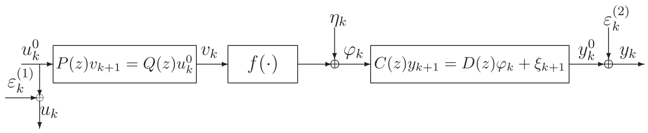

Fig. 1. EIV Wiener–Hammerstein system.

In the paper we consider the SISO EIV Wiener–Hammerstein system (see Fig. 1) described as follows:

$$P(z)v_{k+1} = Q(z)u_k^0, \tag{1}$$

$$\varphi_k = f(v_k) + \eta_k, \tag{2}$$

$$C(z)y_{k+1}^0 = D(z)\varphi_k + \xi_{k+1}, \tag{3}$$

where z denotes the backward-shift operator $zy_{k+1}^0 = y_k^0$, while $f(\cdot)$ is the unknown nonlinear function, and

$$P(z) = 1 + p_1 z + p_2 z^2 + \cdots + p_{n_p} z^{n_p}, \tag{4}$$

$$Q(z) = 1 + q_1 z + q_2 z^2 + \cdots + q_{n_q} z^{n_q}, \tag{5}$$

$$C(z) = 1 + c_1 z + c_2 z^2 + \cdots + c_{n_c} z^{n_c}, \tag{6}$$

$$D(z) = 1 + d_1 z + d_2 z^2 + \cdots + d_{n_d} z^{n_d} \tag{7}$$

are polynomials with unknown coefficients but with known orders n_p, n_q, n_c, n_d. The noise-free input ${u_k}^0$ and output ${y_k}^0$ are observed with additive noises $\varepsilon_k^{(1)}$ and $\varepsilon_k^{(2)}$:

$$u_k = u_k^0 + \varepsilon_k^{(1)}, \quad y_k = y_k^0 + \varepsilon_k^{(2)}. \tag{8}$$

Identification of the EIV Wiener–Hammerstein system is more difficult in comparison with that for the EIV Wiener system discussed in [19]:

(1) The output of the EIV Wiener system is an α-mixing with mixing coefficients decaying exponentially to zero but this is no longer true for the EIV Wiener–Hammerstein system. In [19] it is seen that the mixing property plays an important role in convergence analysis.

(2) Because of the linear subsystem at the output end, more complicated relationships relating the impulse responses and the correlation functions should be taken into account (see Lemma 1).

The goal of this paper is to recursively estimate the unknown parameters of the two linear subsystems $\{p_1, \ldots, p_{n_p}, q_1, \ldots, q_{n_q}, c_1, \ldots, c_{n_c}, d_1, \ldots, d_{n_d}\}$ and the value of $f(x)$ at any given x at the real axis on the basis of the observed data $\{u_k, y_k\}$.

The rest of the paper is arranged as follows. The system assumptions and the recursive algorithms are given in Section 2. The strong consistency of the estimates for the linear and nonlinear parts is proved in Sections 3 and 4, respectively. A numerical example is presented in Section 5, and a brief conclusion is given in Section 6.

2. Assumptions and recursive identification algorithms

2.1. Assumptions

We first give the conditions for identifying the two linear subsystems.

H1: The noise-free input $\{u_k^0\}$ is a sequence of mutually independent, identically distributed (i.i.d.) Gaussian random variables: $u_k^0 \in \mathcal{N}(0, \vartheta^2)$ with unknown $\vartheta > 0$, and is independent of the internal noises $\{\eta_k\}$ and $\{\xi_k\}$ and the observation noises $\{\varepsilon_k^{(1)}\}$ and $\{\varepsilon_k^{(2)}\}$.

H2: $P(z)$ and $Q(z)$ are coprime and $P(z)$ is stable: $P(z) \neq 0$, $\forall |z| \leq 1$.

H3: $C(z)$ and $D(z)$ are coprime and both are stable: $C(z) \neq 0$ and $D(z) \neq 0$, $\forall |z| \leq 1$.

By the stability of $P(z)$ and $C(z)$, we have

$$L(z) \triangleq \frac{Q(z)}{P(z)} = \sum_{i=0}^{\infty} l_i z^i, \tag{9}$$

$$H(z) \triangleq \frac{D(z)}{C(z)} = \sum_{i=0}^{\infty} h_i z^i, \tag{10}$$

where $|l_i| = O(e^{-r_1 i})$, $r_1 > 0$, $i \geq 1$ and $|h_i| = O(e^{-r_2 i})$, $r_2 > 0$, $i \geq 1$, and $l_0 = 1$ and $h_0 = 1$ since all polynomials (4)–(7) are monic. The numbers $\{l_i, i \geq 0\}$ and $\{h_i, i \geq 0\}$ are called the impulse responses of the two linear subsystems, respectively.

H4: Both the measurement noises $\{\varepsilon_k^{(1)}\}$ and $\{\varepsilon_k^{(2)}\}$ belong to the ARMA process:

$$F_1(z)\varepsilon_k^{(1)} = G_1(z)\varsigma_k^{(1)}, \quad F_2(z)\varepsilon_k^{(2)} = G_2(z)\varsigma_k^{(2)}, \tag{11}$$

where

$$F_1(z) = 1 + f_{1,1} z + f_{1,2} z^2 + \cdots + f_{1,n_{f_1}} z^{n_{f_1}}, \tag{12}$$

$$G_1(z) = 1 + g_{1,1} z + g_{1,2} z^2 + \cdots + g_{1,n_{g_1}} z^{n_{g_1}}, \tag{13}$$

$$F_2(z) = 1 + f_{2,1} z + f_{2,2} z^2 + \cdots + f_{2,n_{f_2}} z^{n_{f_2}}, \tag{14}$$

$$G_2(z) = 1 + g_{2,1} z + g_{2,2} z^2 + \cdots + g_{2,n_{g_2}} z^{n_{g_2}}. \tag{15}$$

The polynomial $F_1(z)$ has no common factor with $G_1(z)G_1(z^{-1})z^{n_{g_1}}$, and $F_1(z)$ and $F_2(z)$ are both stable. The driven noises $\{\varsigma_k^{(1)}\}$ and $\{\varsigma_k^{(2)}\}$ and the internal noises $\{\eta_k\}$ and $\{\xi_k\}$ are mutually independent, and each of them is a sequence of i.i.d. zero mean random variables with probability density. Moreover, $E(|\eta_k|^{\Delta}) < \infty$, $E(|\xi_k|^{\Delta}) < \infty$, $E(|\varsigma_k^{(1)}|^{\Delta+3}) < \infty$, and $E(|\varsigma_k^{(2)}|^{\Delta}) < \infty$ for some $\Delta > 3$.

H5: The nonlinear function $f(\cdot)$ is measurable and has both the left limit $f(x^-)$ and the right limit $f(x^+)$ at any point x. The growth rate of $f(x)$ as $|x| \to \infty$ is not faster than a polynomial. Further, at least one of the parameters ρ and κ is nonzero, where

$$\rho \triangleq \frac{1}{\sqrt{2\pi}\sigma^5 \vartheta} \int_{\mathcal{R}} (x^2 - \sigma^2 \vartheta^2) f(x) e^{-x^2/2\sigma^2\vartheta^2} \, dx, \tag{16}$$

$$\kappa \triangleq \frac{1}{\sqrt{2\pi}\sigma^7 \vartheta} \int_{\mathcal{R}} (x^3 - 3\sigma^2 \vartheta^2 x) f(x) e^{-x^2/2\sigma^2\vartheta^2} \, dx, \tag{17}$$

where $\sigma^2 = \sum_{i=0}^{\infty} l_i^2$.

Remark 1. The growth rate restriction in H5 implies that there are constants $\alpha > 0$ and $\beta \geq 1$ such that

$$|f(x)| \leq \alpha(1 + |x|^{\beta}) \quad \forall x \in \mathcal{R}. \tag{18}$$

Therefore, the integrals (16) and (17) are finite.

Let us explain conditions imposed here. Conditions H1 and H4 and also H6 and H7 to be introduced later concern the signals in the system, while Conditions H2, H3, and H5 are on the structure of the system. The purpose of applying the Gaussian input in H1 is to derive the explicit relationships (25)–(28) connecting the impulse responses of the two linear subsystems and the correlation functions between the observed input and output. These relationships are the basis of the proposed algorithms for estimating the impulse responses by using the observed input and output. It is clear that H2 and H3 are the standard condition on the linear subsystems. Condition H4 concerns the measurement errors, which are not negligible in consideration of the present paper. Here we allow them to be correlated. In Condition H5, the function $f(\cdot)$ is allowed to be discontinuous: it is discontinuous at x if $f(x^-) \neq f(x^+)$. Further, the assumption that at least one of the constants ρ and κ is nonzero holds for many practical nonlinearities including polynomials,