Available online at www.sciencedirect.com
ScienceDirect
Systems & Control Letters 56 (2007) 373 – 380
SYSTEMS & CONTROL LETTERS
www.elsevier.com/locate/sysconle

Auxiliary model-based least-squares identification methods for Hammerstein output-error systems ☆

Feng Ding[a,*,1], Yang Shi[b], Tongwen Chen[c]

[a] Control Science and Engineering Research Center, Southern Yangtze University, Wuxi 214122, P.R. China
[b] Department of Mechanical Engineering, University of Saskatchewan, Saskatoon, Canada S7N 5A9
[c] Department of Electrical and Computer Engineering, University of Alberta, Edmonton, Canada T6G 2V4

Received 28 August 2004; received in revised form 26 October 2006; accepted 30 October 2006
Available online 2 January 2007

Abstract

The difficulty in identification of a Hammerstein (a linear dynamical block following a memoryless nonlinear block) nonlinear output-error model is that the information vector in the identification model contains unknown variables—the noise-free (true) outputs of the system. In this paper, an auxiliary model-based least-squares identification algorithm is developed. The basic idea is to replace the unknown variables by the output of an auxiliary model. Convergence analysis of the algorithm indicates that the parameter estimation error consistently converges to zero under a generalized persistent excitation condition. The simulation results show the effectiveness of the proposed algorithms.

Keywords: Recursive identification; Parameter estimation; Least squares; Multi-innovation identification; Hierarchical identification; Auxiliary model; Convergence properties; Stochastic gradient; Hammerstein models; Wiener models; Martingale convergence theorem

1. Introduction

Two typical classes of nonlinear systems—linear time-invariant blocks following (or followed by) static nonlinear blocks—are Hammerstein and Wiener (H–W) nonlinear systems, which are common in industry, e.g., the valve saturation nonlinearities, dead-zone nonlinearities and linear systems equipped with nonlinear sensors [35]. In general, existing identification approaches for H–W models can be roughly divided into two categories: the iterative and the recursive algorithms. In order to distinguish on-line from off-line calculation, we use *iterative* for off-line algorithms, and *recursive* for on-line ones. We imply that a recursive algorithm can be on-line implemented, but an iterative one cannot. For a recursive algorithm, new information (input and/or output data) is always used in the algorithm which recursively computes the parameter estimates every step as time increases.

Some iterative and/or off-line algorithms of H–W models were discussed in [2,1,4,6,7,22,23,25,28,30,36,11,20] and other recursive and/or on-line algorithms were studied in, e.g., [35,11,20,32,33,3,5,34]. In the identification area of nonlinear systems, Bai reported a two-stage identification algorithm for Hammerstein–Wiener nonlinear systems based on singular value decomposition (SVD) and least squares [1] and studied identification problem of systems with hard input nonlinearities of known structure [2]; Vörös presented a half-substitution algorithm to identify Hammerstein systems with two-segment nonlinearities and with multisegment piecewise-linear characteristics, but no convergence analysis was carried out [32,33]; Cerone and Regruto analyzed parameter error bounds in the Hammerstein models by using the output measurement error bound [6]. Also, Pawlak used the series expansion approach to study the identification of Hammerstein nonlinear output-error (state-space) models [29]. Recently, an iterative least-squares and a recursive least-squares identification methods were reported in [11], and an iterative gradient and a recursive

☆ This research was supported by the Natural Sciences and Engineering Research Council of Canada and the National Natural Science Foundation of China (No. 60574051 and 60528007).
* Corresponding author. Tel.: 86510 88637783; fax: +86510 85910652.
E-mail addresses: dingf@tsinghua.edu.cn, fding@sytu.edu.cn (F. Ding), tchen@ece.ualberta.ca (T. Chen).
[1] F. Ding is also an Adjunct Professor in the College of Automation at the Nanjing University of Technology, Nanjing 210009, P.R. China.

0167-6911/$ - see front matter
doi:10.1016/j.sysconle.2006.10.026

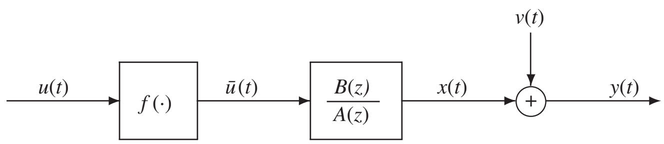

Fig. 1. The Hammerstein nonlinear output-error system.

stochastic gradient algorithms were developed in [20] for nonlinear ARMAX models based only on the available input–output data, and the convergence properties of the recursive algorithms involved were proved.

However, most of the identification approaches in the literature assume that the systems under consideration are nonlinear ARX models, or equation-error-like models [35,2,1,7,3]. That is, each element of the information vector consisting of input–output data is measured. In this paper, we focus on the identification problem of a class of Hammerstein output-error-type nonlinear systems and present an auxiliary-model least-squares (AMLS) algorithm, which is different from the ones mentioned above in that the information vector in our identification model contains *unknown variables* (namely, unavailable noise-free outputs), and we adopt an auxiliary model or reference model to estimate these unknown variables and further use the outputs of the auxiliary model instead of the unknown noise-free outputs to identify the system parameters. The basic idea is to extend the Landau's output-error method to study identification problem of nonlinear systems [21,31]. To the best of our knowledge, few publications addressed identification methods of Hammerstein nonlinear output-error systems, especially the convergence problem of the algorithms involved, which are the focus of this work.

The objective of this paper is, by means of the auxiliary model identification principle, to derive an algorithm to estimate the system parameters of the nonlinear output-error models based on the available input–output data $\{u(t), y(t)\}$, and to study the properties of the algorithm involved.

Briefly, the paper is organized as follows. Section 2 describes the identification algorithms related to the Hammerstein systems. Section 3 analyzes the properties of the proposed stochastic algorithm. Section 4 provides an illustrative example to show the effectiveness of the algorithm proposed. Finally, we offer some concluding remarks in Section 5.

2. The algorithm description

Consider the Hammerstein output-error system shown in Fig. 1 that consists of a nonlinear memoryless element followed by a linear output-error model [6,29], where the true output (namely, the noise-free output) $x(t)$ and the inner variable $\bar{u}(t)$ (namely, the output of the nonlinear block) are unmeasurable, $u(t)$ is the system input, $y(t)$ is the measurement of $x(t)$, $v(t)$ is an additive noise with zero mean. The nonlinear part in the Hammerstein model is a polynomial of a known order in the input [7,22,28], or, more generally, a nonlinear function of a known basis $(\gamma_1, \gamma_2, \dots, \gamma_m)$ as follows [6,11,20]:

$$\bar{u}(t) = f(u(t)) = c_1\gamma_1(u(t)) + c_2\gamma_2(u(t)) + \cdots + c_m\gamma_m(u(t)) = \sum_{j=1}^{m} c_j\gamma_j(u(t)). \tag{1}$$

Then the Hammerstein nonlinear output-error model in Fig. 1 may be expressed as

$$x(t) = \frac{B(z)}{A(z)}\bar{u}(t) = \frac{B(z)}{A(z)}[c_1\gamma_1(u(t)) + c_2\gamma_2(u(t)) + \cdots + c_m\gamma_m(u(t))],$$

$$y(t) = x(t) + v(t).$$

Here, $A(z)$ and $B(z)$ are polynomials in the shift operator z^{-1} $[z^{-1}y(t) = y(t-1)]$ with

$$A(z) = 1 + a_1z^{-1} + a_2z^{-2} + \cdots + a_nz^{-n},$$
$$B(z) = b_1z^{-1} + b_2z^{-2} + b_3z^{-3} + \cdots + b_nz^{-n}.$$

Notice that for the Hammerstein model shown in Fig. 1, $f(u)$ and $G(z) := B(z)/A(z)$ are actually not unique. Any pair $(\alpha f(u), G(z)/\alpha)$ for some nonzero and finite constant α would produce identical input and output measurements. In other words, any identification scheme cannot distinguish between $(f(u), G(z))$ and $(\alpha f(u), G(z)/\alpha)$. Therefore, to get a unique parameterization, without loss of generality, one of the gains of $f(u)$ and $G(z)$ has to be fixed. There are several ways to normalize the gains [1,6,22]. Here, we adopt the assumption [22,3]: the first coefficient of the function $f(\cdot)$ equals 1; i.e., $c_1 = 1$ [11,20].

Eq. (2) can be rewritten as a recursive form

$$x(t) = -\sum_{i=1}^{n} a_ix(t-i) + \sum_{i=1}^{n} b_i\bar{u}(t-i) = -\sum_{i=1}^{n} a_ix(t-i) + \sum_{i=1}^{n} b_i\sum_{j=1}^{m} c_j\gamma_j(u(t-i)).$$

Define the parameter vector $\boldsymbol{\theta}$ and information vector $\boldsymbol{\varphi}_0(t)$ as

$$\boldsymbol{\theta} = \begin{bmatrix} \boldsymbol{a} \\ c_1\boldsymbol{b} \\ c_2\boldsymbol{b} \\ \vdots \\ c_m\boldsymbol{b} \end{bmatrix} \in \mathbb{R}^{n_0},$$

$$\boldsymbol{\varphi}_0(t) = \begin{bmatrix} -x(t-1) \\ -x(t-2) \\ \vdots \\ -x(t-n) \\ \boldsymbol{\psi}(t) \end{bmatrix} \in \mathbb{R}^{n_0}, \quad n_0 := (m+1)n, \tag{2}$$

$$\boldsymbol{a} = [a_1, a_2, \dots, a_n]^{\mathrm{T}} \in \mathbb{R}^n, \quad \boldsymbol{b} = [b_1, b_2, \dots, b_n]^{\mathrm{T}} \in \mathbb{R}^n,$$

$$\boldsymbol{c} = [c_2, c_3, \dots, c_m]^{\mathrm{T}} \in \mathbb{R}^{m-1},$$

$$\boldsymbol{\psi}(t) = [\boldsymbol{\psi}_1^{\mathrm{T}}(t), \boldsymbol{\psi}_2^{\mathrm{T}}(t), \dots, \boldsymbol{\psi}_m^{\mathrm{T}}(t)]^{\mathrm{T}} \in \mathbb{R}^{mn}, \tag{3}$$