ARTICLE IN PRESS
JID:YDSPR AID:2236 /FLA [m5G; v1.225; Prn:22/11/2017; 16:14] P.1 (1-10)
Digital Signal Processing ••• (••••) •••–•••

Contents lists available at ScienceDirect

Digital Signal Processing

www.elsevier.com/locate/dsp

Expectation maximization estimation algorithm for Hammerstein models with non-Gaussian noise and random time delay from dual-rate sampled-data ☆

Junxia Ma, Jing Chen, Weili Xiong, Feng Ding

School of Internet of Things Engineering, Jiangnan University, Wuxi 214122, PR China

A R T I C L E I N F O

Article history:
Available online xxxx

Keywords:
Non-Gaussian noise
Time delay
Expectation maximization
Parameter estimation
Input nonlinear system

A B S T R A C T

This paper considers the robust identification for dual-rate input nonlinear equation-error systems with outliers and random time delay. To suppress the negative influence caused by the outliers to the accuracy of identification, the distribution of the noise is represented by a t-distribution rather than a Gaussian distribution. A random time delay is considered in the dual-rate input nonlinear systems. By treating the unknown time delay as the latent variable, the expectation maximization algorithm is derived for identifying the systems. Two numerical simulation examples demonstrate that the proposed algorithm can generate accurate identification results when the measurements are contaminated by the outliers.

1. Introduction

The identification for nonlinear systems has received a lot of attention [1–3]. Typical nonlinear models include Hammerstein models, Wiener models, Hammerstein–Wiener models and so on. Hammerstein model which has a static nonlinear block followed by a linear dynamic subsystem has been widely used to describe the system with the input nonlinearity [4]. The structure of the Hammerstein model is simple and flexible. The memoryless nonlinear block can be the continuous nonlinear function or the piecewise linear function. The dynamic subsystem may be the state space representation, the equation-error model or the transfer functions and so on [5]. The parameter estimation for the Hammerstein model with white noise has received great attention. However, few work has been focused on the robust identification for the Hammerstein model. Disturbances universally exist in industrial processes [7]. Especially in the areas of control system and signal processing [6], the observed outputs always contain disturbances from process environments [8,9]. The disturbances could be white noise or colored noise [10,11]. The noise may contain certain unexpected or uncertain large magnitude points, which are called outliers. The outliers are unavoidable in practical processes for some uncertainty elements, such as the measurement errors, signal interferences or disturbances from outside environments [12].

Using the measurements which are contaminated by outliers for system identification will bring negative influence to the accuracy of the parameter estimation [13,14]. Some screening methods have been studied to deal with the outliers. The main idea of this approach is to discard the unexpected points by trimming the measurement dataset [15]. But the simple removing of data will lead to biased estimation [16]. Traditionally, the Gaussian distribution is widely used to fit the random noise in the modeling process. However, one major shortcoming of such a Gaussian model is its sensitivity to outliers [17]. Another way to deal with the outliers or gross errors is to use the so-called contaminated Gaussian distribution. It is a two-component combined Gaussian distribution, the normal one representing regular noise with small variances and the non-normal one representing outliers with large variances [18]. For example, Jin and Huang modeled the noise with a contaminated Gaussian distribution and used the expectation maximization (EM) algorithm to identify the piecewise/switching autoregressive exogenous (ARX) process [19].

A general approach to cope with the potential outliers is to use t-distribution to describe the process contaminated by outliers. The t-distribution is employed since it has longer tails than a Gaussian distribution owing to its adjustable degree of freedom [20]. For example, Chamroukhi proposed a robust mixture of experts modeling using the t-distribution to deal with groups of observations with heavy tails or atypical observations [21]. Sammaknejad et al. proposed a robust approach to process monitoring and diagnosis based on a time-varying hidden Markov model [22]. In the discussed systems, the multivariate t-distributions are used to model

☆ This work was partially supported by the National Natural Science Foundation of China and the Natural Science Foundation of Jiangsu Province (No. BK20160162).
E-mail addresses: junxia.20@163.com (J. Ma), chenjing1981929@126.com (J. Chen), greenpre@163.com (W. Xiong), fding@jiangnan.edu.cn (F. Ding).

https://doi.org/10.1016/j.dsp.2017.11.009
1051-2004/

observations around different process operating modes with different likelihoods of outliers.

Because of the limitation of the measured sensors or the analysis of quality variables, the measured input–output data are impossibly always be sampled at the same rate [23]. For example, the pressure and temperature can be sampled at a fast rate by the sensor while the analysis of the components for variables can only be sampled at a slow rate. In process industries, the dual-rate sampled-data system and the multirate sampled-data system is very common [24–26]. Mo et al. presented a two-stage method to identify the dual-rate systems with fast input and slow output and applied the proposed method to an industrial distillation column to develop a composition observer [27]. Li et al. proposed a polynomial domain method to identify a fast single-rate linear model based on dual-rate input–output data and applied the proposed identification and control algorithm to a Shell Canada's continuous catalytic reforming reactor to improve the octane quality control [28].

Time delay is another issue that must be considered in dynamic systems [29]. Known or unknown delays in the process of signal processing and system identification have been discussed [30,31]. In these work, the time delay is considered as a fixed value, that means in every different sampling point, no matter it is known or unknown, the delay is assumed to remain the same. However, time delay may be time-varying [32,33] or randomly distributed [34]. For example, Zhao et al. derived a variational Bayesian approach to identify ARX models with time-varying time delays [35]. They considered the time-varying time delays following a random Markov chain and the transition of delay from one value to another is determined by a transition probability matrix. Moser et al. proposed an extend gradient-based least-squares algorithm to obtain the recursive parameter estimation for exhaust gas oxygen sensors, which are characterized by input-dependent time delays and linear parameters and the time delays are modeled as the functions of the input signals [36].

On the basis of the work in [34], this paper considers a robust identification problem, which not only means that the measurements are corrupted by larger variance of outliers but also means that there exists unknown random time delays in the process of signal transmission. Different from [34], this paper extends the parameter estimation problem to the nonlinear Hammerstein model with outliers rather than the linear FIR model with Gaussian noise. Moreover, for the dual-rate input nonlinear model, Chen et al. derived a recursive least squares (RLS) algorithm to estimate the unknown model parameters [37]. However, the RLS algorithm cannot get the estimates for the random time delay and is sensitive to the outliers in the noise. The main contributions of this paper are as follows.

- Random time delays rather than a fixed value one are considered in the process of signal transmission.
- Under the principle of the expectation maximization, modeling the disturbance as a t-distribution and treating the variance scale in the t-distribution and the unknown time delay as the hidden variables, a robust EM algorithm is derived to identify the dual-rate input nonlinear equation-error systems.
- By providing the full probability distribution of the time delay at each sampling instant, the accurate estimate for the time delay is presented.
- By means of the over-parameterization, the parameter product terms from the nonlinear block and the linear block are partially separated. The unknown parameters of the nonlinear block and the linear block are estimated simultaneously.

The remainder of this paper is organized as follows. Section 2 shows the identification model of the dual-rate input nonlinear Hammerstein models and gives the probability density function for the noise and the measurements. Section 3 provides the detailed procedures of deriving the EM algorithm for the input nonlinear dual-rate systems. Two simulation examples are provided to show the effectiveness of the proposed algorithm in Section 4. Finally, Section 5 offers some conclusions.

2. System description

In the dual-rate systems, the measured inputs and outputs are sampled by two different rates. Assume that the inputs have fast rate variables, while the outputs have the characteristic of slow change rates. Consider a dual-rate input nonlinear equation-error model (i.e., Hammerstein model):

$$A(z)y_{t_i} = B(z)f(u_{t_i-\lambda_i}) + e_{t_i}, \tag{1}$$

where $\{u_t, t = 1, 2, \dots, N\}$ is the fast rate input signal and available at each sampling time t; $\{y_{t_i}, i = 1, 2, \dots, L\}$ is the slow rate output and only available at time instant $t_i = i \cdot q$ ($q \geqslant 1$ is an integer), that means the slow rate sampling period is q times that of the fast rate ($q = N/L$); e_{t_i} is the measurement noise; $\{\lambda_i, i = 1, 2, \dots, L\}$ is the unknown random time delay; $A(z)$ and $B(z)$ are polynomials of unit backward shift operator ($z^{-1}u_t = u_{t-1}$, $z^{-1}y_{t_i} = y_{t_{i-1}}$):

$$A(z) := 1 + a_1z^{-1} + a_2z^{-2} + \cdots + a_{n_a}z^{n_a},$$
$$B(z) := b_1z^{-1} + b_2z^{-2} + \cdots + b_{n_b}z^{n_b}.$$

The input nonlinearity is fitted by a polynomial with known basis and unknown coefficients:

$$\bar{u}_t := f(u_t) = c_1\eta_1(u_t) + c_2\eta_2(u_t) + \cdots + c_{n_c}\eta_{n_c}(u_t), \tag{2}$$

where $\eta_j(u_t)$ is the known basis; c_j is the unknown coefficient.

Substituting the polynomials $A(z)$ and $B(z)$ and (2) into (1) gives

$$y_{t_i} = -\sum_{j=1}^{n_a} a_i y_{t_{i-j}} + \sum_{l=1}^{n_b} b_l \sum_{m=1}^{n_c} c_m\eta_m(u_{t_{i-l}-\lambda_i}) + e_{t_i}. \tag{3}$$

It can be easily find that there are product terms of parameters from nonlinear and linear blocks in the right-hand side of (3), it is impossible to obtain the unique parameter estimates. Any identification scheme cannot distinguish $(\boldsymbol{b}, \boldsymbol{c})$ from $(\alpha\boldsymbol{b}, \boldsymbol{c}/\alpha)$ for some nonzero and finite constant α, because any pair $(\alpha\boldsymbol{b}, \bar{u}_{t_i}/\alpha)$ would produce identical input and output measurements. Therefore, at least one parameter in two blocks should be fixed. For the sake of simplicity, the over-parameterization method [38] is adopted and the first non-zero coefficient in the polynomial $\bar{u}_t$ is assumed to be one, i.e., $c_1 = 1$.

Assume that the orders n_a, n_b and n_c are known and $y_t = 0$, $u_t = 0$ and $e_t = 0$ for $t \leqslant 0$. Let $n := n_a + n_b n_c$. Define the parameter vector $\boldsymbol{\theta}$ and the information vector $\boldsymbol{\phi}_{t_i}$ as

$$\begin{aligned}\boldsymbol{\theta} &:= [a_1, a_2, \dots, a_{n_a}, b_1, b_1c_2, \dots, b_1c_{n_c}, b_2, b_2c_2, \dots, b_2c_{n_c},\\ &\quad \dots, b_{n_b}, b_{n_b}c_2, \dots, b_{n_b}c_{n_c}]^{\mathrm{T}} \in \mathbb{R}^n,\\ \boldsymbol{\phi}_{t_i} &:= [-y_{t_{i-1}}, -y_{t_{i-2}}, \dots, -y_{t_{i-n_a}}, \eta_1(u_{t_{i-1}-\lambda_i}), \eta_2(u_{t_{i-1}-\lambda_i}),\\ &\quad \dots, \eta_{n_c}(u_{t_{i-1}-\lambda_i}), \eta_1(u_{t_{i-2}-\lambda_i}), \dots, \eta_{n_c}(u_{t_{i-2}-\lambda_i}), \dots,\\ &\quad \eta_1(u_{t_{i-n_b}-\lambda_i}), \dots, \eta_{n_c}(u_{t_{i-n_b}-\lambda_i})]^{\mathrm{T}} \in \mathbb{R}^n.\end{aligned} \tag{4}$$

Then Equation (3) can be rewritten as

$$y_{t_i} = \boldsymbol{\phi}_{t_i}^{\mathrm{T}}\boldsymbol{\theta} + e_{t_i}. \tag{5}$$