Automatica 60 (2015) 48–56

Contents lists available at ScienceDirect

Automatica

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

Brief paper

Parameter estimation for nonlinear time-delay systems with noisy output measurements[☆]

Qun Lin[a], Ryan Loxton[a,b], Chao Xu[b,1], Kok Lay Teo[a,b]

[a] *Department of Mathematics and Statistics, Curtin University, Perth, Australia*

[b] *Institute of Cyber-Systems and Control, Zhejiang University, Hangzhou, China*

ARTICLE INFO

Article history:
Received 17 October 2013
Received in revised form
2 March 2015
Accepted 9 June 2015
Available online 13 July 2015

Keywords:
Time-delay
Nonlinear system
Parameter estimation
Dynamic optimization
Nonlinear optimization

ABSTRACT

This paper considers the problem of using noisy output data to estimate unknown time-delays and unknown system parameters in a general nonlinear time-delay system. We formulate the problem as a dynamic optimization problem in which the unknown quantities are decision variables to be chosen optimally, with the cost function penalizing the mean and variance of the least-squares error between actual and predicted system output. Since the time-delays and system parameters influence the cost function implicitly through the governing time-delay system, the cost function's gradient – which is required to solve the problem using gradient-based optimization techniques – cannot be computed analytically using standard differentiation rules. We instead develop two computational methods for evaluating this gradient: one involves solving an auxiliary time-delay system forward in time; the other involves solving an auxiliary time-advance system backward in time. On this basis, we propose an efficient optimization algorithm for determining optimal estimates for the time-delays and system parameters. We conclude the paper by examining the performance of this algorithm on a dynamic model of a continuously-stirred tank reactor.

1. Introduction

Developing a mathematical model is a two-step process: first, the general structure of the model is derived based on fundamental physical principles; then, the model is matched to a particular system of interest by tuning various model parameters. This second step, known as *parameter estimation* or *parameter identification*, usually involves comparing the system output predicted by the model with the real system output measured during an experiment (or a series of experiments).

This paper is concerned with parameter estimation for nonlinear time-delay systems. We consider a general dynamic model consisting of nonlinear delay-differential equations with multiple time-delays and multiple system parameters, each of which is unknown and needs to be estimated appropriately. The problem that we investigate – called the *parameter estimation problem* – is to determine optimal estimates for the time-delays and system parameters so that the dynamic model best fits the real system under consideration. Such problems are commonly referred to as *inverse problems*.

Parameter estimation for time-delay systems has attracted considerable research interest over the past two decades (see Belkoura, Richard, & Fliess, 2009, Drakunov, Perruquetti, Richard, & Belkoura, 2006, Lunel, 2001, Orlov, Belkoura, Richard, & Dambrine, 2002, 2003, Park, Han, & Kwon, 2013, Tuch, Feuer, & Palmor, 1994 and Zheng, Barbot, & Boutat, 2013). Popular approaches for solving the parameter estimation problem include swarm intelligence algorithms such as particle swarm optimization (Gao, Qi, Yin, & Xiao, 2010; Tang & Guan, 2009), or finite-dimensional approximation schemes for the original infinite-dimensional time-delay model (Banks, Rehm, & Sutton, 2010). Recently, a new gradient-based optimization approach has been proposed by Chai, Loxton, Teo, and Yang (2013a,b) and Loxton, Teo, and Rehbock (2010). In this approach, the parameter estimates are chosen as the solution of a dynamic optimization problem in which the cost

☆ The authors acknowledge support from the National Natural Science Foundation of China (Grant Numbers 61473253 and 11350110208), the State Key Laboratory for Industrial Control Technology at Zhejiang University, China (Open Research Project ICT1301) and the National High Technology Research and Development Program of China (Grant Number 2012AA041701). The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Michael V. Basin under the direction of Editor Ian R. Petersen.

E-mail addresses: q.lin@curtin.edu.au (Q. Lin), r.loxton@curtin.edu.au (R. Loxton), cxu@zju.edu.cn (C. Xu), k.l.teo@curtin.edu.au (K.L. Teo).

[1] Tel.: +86 571 87952233; fax: +86 571 87952279.

http://dx.doi.org/10.1016/j.automatica.2015.06.028
0005-1098/

function penalizes the deviation between predicted and measured system output. Special dynamic optimization techniques can then be deployed to solve this problem and obtain accurate estimates for the model parameters. This approach was introduced by Loxton et al. (2010) for nonlinear time-delay systems in which each nonlinear term contains a single delay and no other model parameters, and then extended by Chai et al. (2013a) to more general nonlinear systems with multiple delays and multiple system parameters. In Chai et al. (2013b), the approach was applied to a more difficult parameter estimation problem in which the dynamic system contains both state- and input-delays, and the input function is discontinuous.

The two main advantages of the parameter estimation methods proposed by Chai et al. (2013a,b) and Loxton et al. (2010) are: (i) these methods can readily handle system nonlinearities; and (ii) these methods can simultaneously compute optimal estimates for the time-delays and system parameters in a unified fashion. One limitation, however, is that these methods do not take into account the possibility of noise in the output data. Thus, the output measurements used in the cost function (recall that the cost function penalizes the discrepancy between predicted and measured system output) are assumed to be exact. This is, of course, an idealistic assumption, as it is impossible to guarantee perfect precision when measuring the output of a real system.

The purpose of this paper is to address this limitation. Building on the results in Chai et al. (2013a,b) and Loxton et al. (2010), we will devise a new method for parameter estimation that explicitly takes output measurement noise into account. The main idea is to consider the output data points as random variables, rather than fixed constants. This allows for possible discrepancies between the actual and observed system output due to measurement errors. With the output data as random variables, our parameter estimation problem is formulated as a stochastic dynamic optimization problem in which the aim is to choose the time-delays and system parameters to minimize a weighted sum of the expectation and variance of the least-squares error between actual and predicted system output. We will develop a computational approach for solving this problem based on novel dynamic optimization techniques. The result is a unified parameter estimation method for nonlinear time-delay systems that is fast, versatile, and capable of handling uncertainties in the measured output data.

2. Problem statement

Consider the following nonlinear time-delay system:

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{f}(\boldsymbol{x}(t), \boldsymbol{x}(t-\tau_1), \ldots, \boldsymbol{x}(t-\tau_m), \boldsymbol{\zeta}), \quad t \geq 0, \tag{1}$$

$$\boldsymbol{x}(t) = \boldsymbol{\phi}(t, \boldsymbol{\zeta}), \quad t \leq 0, \tag{2}$$

where $\boldsymbol{x}(t) \in \mathbb{R}^n$ is the *state vector*; $\boldsymbol{\zeta} \in \mathbb{R}^r$ is the *parameter vector*; τ_i, $i = 1, \ldots, m$, are *time-delays*; and $\boldsymbol{f} : \mathbb{R}^{(m+1)n} \times \mathbb{R}^r \to \mathbb{R}^n$ and $\boldsymbol{\phi} : \mathbb{R} \times \mathbb{R}^r \to \mathbb{R}^n$ are given continuously differentiable functions.

The output $\boldsymbol{y}(t) \in \mathbb{R}^q$ of system (1)–(2) is given by the following equation:

$$\boldsymbol{y}(t) = \boldsymbol{g}(\boldsymbol{x}(t), \boldsymbol{\zeta}), \quad t \geq 0, \tag{3}$$

where $\boldsymbol{g} : \mathbb{R}^n \times \mathbb{R}^r \to \mathbb{R}^q$ is a given continuously differentiable function.

Both the time-delays τ_i, $i = 1, \ldots, m$, and the parameter vector $\boldsymbol{\zeta}$ are unknown and need to be estimated. Let a_i and b_i denote the lower and upper bounds of the ith time-delay. Then

$$a_i \leq \tau_i \leq b_i, \quad i = 1, \ldots, m. \tag{4}$$

Any vector $\boldsymbol{\tau} \in \mathbb{R}^m$ with components satisfying (4) is called a *candidate time-delay vector* for system (1)–(3). Let $\mathcal{T}$ denote the set of all candidate time-delay vectors.

Similarly, let c_j and d_j denote the lower and upper bounds of the jth system parameter in $\boldsymbol{\zeta}$. Then

$$c_j \leq \zeta_j \leq d_j, \quad j = 1, \ldots, r. \tag{5}$$

Any vector $\boldsymbol{\zeta} \in \mathbb{R}^r$ with components satisfying (5) is called a *candidate parameter vector* for system (1)–(3). Let $\mathcal{Z}$ denote the set of all candidate parameter vectors.

For each candidate pair $(\boldsymbol{\tau}, \boldsymbol{\zeta}) \in \mathcal{T} \times \mathcal{Z}$, let $\boldsymbol{x}(\cdot|\boldsymbol{\tau}, \boldsymbol{\zeta})$ denote the state trajectory obtained by solving Eqs. (1)–(2) with the components of $\boldsymbol{\tau}$ and $\boldsymbol{\zeta}$ used as the time-delays and system parameters, respectively. Furthermore, let $\boldsymbol{y}(\cdot|\boldsymbol{\tau}, \boldsymbol{\zeta})$ denote the corresponding output function obtained by substituting $\boldsymbol{x}(\cdot|\boldsymbol{\tau}, \boldsymbol{\zeta})$ into (3).

Our goal is to estimate the unknown time-delays and system parameters by comparing the predicted system output (obtained by solving the model (1)–(3)) with the actual system output (measured during a series of experiments) at a set of sample times $\{t_k\}_{k=1}^{p}$, where

$$0 = t_0 < t_1 < t_2 < \cdots < t_{p-1} < t_p.$$

Let $\hat{\boldsymbol{y}}^k$ denote the actual system output at time $t = t_k$. In Chai et al. (2013a,b) and Loxton et al. (2010), we assumed that the output vectors $\hat{\boldsymbol{y}}^k$, $k = 1, \ldots, p$, can be measured exactly. However, this assumption is unrealistic; due to system noise and measurement errors, the true system output will often differ slightly from the measured output. Thus, in this paper, we view $\hat{\boldsymbol{y}}^k$, $k = 1, \ldots, p$, as random vectors of known distribution.

We assume that the following matrices can be obtained from the distribution of $\hat{\boldsymbol{y}}^k$, $k = 1, \ldots, p$:

$$\boldsymbol{\Xi}^{k,l} = \mathrm{Cov}\{\hat{\boldsymbol{y}}^k, \hat{\boldsymbol{y}}^l\}, \qquad \boldsymbol{\Upsilon}^{k,l} = \mathrm{Cov}\{(\hat{\boldsymbol{y}}^k)^2, \hat{\boldsymbol{y}}^l\}, \tag{6}$$

where $(\hat{\boldsymbol{y}}^k)^2$ denotes the vector obtained by squaring each element of $\hat{\boldsymbol{y}}^k$. The issue of computing these matrices is discussed in Section 4.

Any $\boldsymbol{\tau} \in \mathcal{T}$ is a candidate for the real time-delay vector. Similarly, any $\boldsymbol{\zeta} \in \mathcal{Z}$ is a candidate for the real parameter vector. To measure estimation accuracy, we use the following least-squares error function:

$$J(\boldsymbol{\tau}, \boldsymbol{\zeta}) = \sum_{k=1}^{p} \left\| \boldsymbol{y}(t_k|\boldsymbol{\tau}, \boldsymbol{\zeta}) - \hat{\boldsymbol{y}}^k \right\|^2.$$

Our parameter estimation problem is stated as follows.

Problem P. Choose $\boldsymbol{\tau} \in \mathcal{T}$ and $\boldsymbol{\zeta} \in \mathcal{Z}$ to minimize

$$G(\boldsymbol{\tau}, \boldsymbol{\zeta}) = \gamma \mathrm{E}\{J(\boldsymbol{\tau}, \boldsymbol{\zeta})\} + (1-\gamma)\mathrm{Var}\{J(\boldsymbol{\tau}, \boldsymbol{\zeta})\},$$

where $\mathrm{E}\{\cdot\}$ denotes expectation, $\mathrm{Var}\{\cdot\}$ denotes variance, and $\gamma \in [0, 1]$ is a given weight.

The aim in Problem P is to minimize *both* the average error and the error variance. The weight γ controls the relative importance between these two objectives. If γ is close to one, then the priority is to minimize average error; if γ is close to zero, then the priority is to minimize error variance. When the output distribution is known exactly, $\gamma = 1$ is the best option for minimizing the expected error. However, as we show in the numerical simulations in Section 6, when there are errors and/or uncertainties in the output distribution, it is essential to choose $\gamma < 1$ to ensure solution robustness.