Int. J. Electron. Commun. (AEÜ) 78 (2017) 213–219

Contents lists available at ScienceDirect

International Journal of Electronics and Communications (AEÜ)

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

Regular paper

Two-stage adaptive estimation of irrational linear systems

M.N. Kapetina *, M.R. Rapaić, Z.D. Jeličić

University of Novi Sad, Faculty of Technical Sciences, Trg Dositeja Obradovica 6, Novi Sad, Serbia

ARTICLE INFO

Article history:
Received 12 February 2017
Accepted 8 April 2017

Keywords:
Adaptive parameter estimation
Fractional order systems
Distributed parameter systems
Delayed systems
Recursive Least Square
Gradient algorithm

ABSTRACT

In the present paper, an adaptive parameter estimation algorithm applicable to linear systems with transfer functions of arbitrary structure is proposed. The approach can be applied to a wide class of linear processes, including non-linearly parameterized ones. The proposed method is applicable to fractional-order systems, distributed-parameter and delayed systems, and other classes of systems described by irrational transfer functions. In the first stage of the proposed procedure, values of the transfer function at specific frequencies are pinpointed by means of the Recursive Least Square algorithm with forgetting factor. In the second stage, the unknown parameters are found by numerically inverting complex non-linear relations linking them to the quantities estimated in the first stage. The inversion is performed by means of an iterative, gradient-based scheme. The method is illustrated by several detailedly explained numerical examples.

1. Introduction

System identification is concerned with development of mathematical models and estimation of the unknown parameters of those models. A special class of identification techniques has the remarkable ability to track parameter variations in time, i.e. to *learn* process behavior from data. Such techniques are essential part of adaptive control and signal processing, but are also becoming increasingly popular in other fields, such are process monitoring and fault detection.

Solution of any estimation problem consists of three steps. First, one needs to select a family of potential models $\mathcal{M}$. In typical engineering applications, $\mathcal{M}$ is the set of all rational transfer functions of certain order, or a set of state-space models in a particular form. There are many other possibilities; for example, one can choose to consider a set of specific fractional order transfer functions, or to allow for even more complex model structure, as we do in the sequel. In the second step, an experiment is performed in which a particular input $u(t)$ is applied to the system, and the response $y(t)$ is recorded. Finally, in the third step, a search is performed in order to find a model in $\mathcal{M}$ whose response to the selected input $u(t)$ is in the highest agreement with the recorded data. Typically, all models in $\mathcal{M}$ are parameterized (indexed) by a vector of unknown quantities $\hat{\theta}$. The aim of identification is to find the "*best*" value of $\hat{\theta}$, i.e. the one that minimizes the discrepancy measure

$$\left\| y(t) - \hat{y}(t, \hat{\theta}) \right\|, \tag{1}$$

where $\hat{y}$ is output of the model when excited with the same input signal u as the process, and $\|\cdot\|$ is some suitably selected norm. Consequences in choosing different types of norms in (1) are considered in [1]. Applicability of different methods for optimizing (1) and estimating the vector of unknown parameters θ depends strongly on the structure of the model. When considering dynamical systems described by rational transfer functions, it is always possible to express $\hat{y}$ as a linear mapping of the unknown parameters $\hat{\theta}$. When, in addition, one selects the norm in (1) to be Euclidean 2-norm, i.e. sum of squared errors, the underlying optimization problem is known as **Least Squares** (**LS**). The solution is obtained by solving the well-known system of "*normal equations*", and it can even be found in closed form by means of the *Penrose-Moore pseudo inverse*. The solution can also be obtained recursively, by the so called **Recursive Least Squares** (**RLS**) algorithm, and its various modification. Perhaps the best-known one is the celebrated **Kalman Filter** (**KF**) [2], which is arguably the most widely used adaptive estimation algorithm today. In depth discussion regarding adaptive estimation techniques, their properties, implementation and modifications can be found in [3,4].

Unfortunately, LS, RLS and KF are not, in general, applicable when one wishes to identify various parameters of irrational transfer functions. Many mathematical models found in recent literature are inherently non-rational. **Fractional-order models**, for example, are found to be superior to integer-order ones in a variety of applications [5–8]. **Distributed parameter models** are another

* Corresponding author.
E-mail address: mirna.kapetina@uns.ac.rs (M.N. Kapetina).

http://dx.doi.org/10.1016/j.aeue.2017.04.008
1434-8411/

prototypical example. They are typically described using *partial differential equations* with a variety of associated boundary conditions [9,10], leading to a plethora of highly non-rational transfer functions, [10,11]. Most industrial processes are also characterized by *transport and communication delays* [12,13], which can sometimes fluctuate. Processes exhibiting long "*dead time*", especially if it is changing or uncertain, are notoriously difficult to control [14]. By interconnecting several delayed processes, complex irrational **models of retarded type** are obtained [15].

If a process model under consideration is described by a non-rational transfer function, so that its response cannot be linearly parameterized, it is not straightforward to solve the underlying optimization problem. This is mainly because the criteria (1) is no longer convex. Indeed, as it will be demonstrated in the sequel, the criteria can exhibit several *global* optima, leading to non-uniqueness of the solution. In fact, a fundamental result established by Ljung [16] is that *global optimal solution to the identification problem is possible only when the underlying parameterization is intrinsically linear, i.e. when the transfer functions under consideration are completely rational*. In the case of off-line, non-adaptive identification it is possible to solve the underlying optimization problem by a suitable non-linear programming algorithm [17,1], or even some modern, global optimization technique such as Genetic Algorithm (GA), or Particle Swarm Optimization (PSO) [18,19]. The above-mentioned procedures are, however, computationally expensive and are not well suited for on-line applications. Several specific methods have been reported for parameter estimation in dynamical systems with particular model structure. For example, various identification techniques targeting fractional-order models have recently emerged. The majority of existing identification methods consider a known commensurate system order and then applies conventional techniques for estimation of the remaining parameters. In this context, gradient method has been utilized in [20,21], Least Square in [22–24], while [25] reports utilization of modulating functions. Complete identification of fractional order models is discussed in [26,27]. A detailed discussion of parameter estimation techniques targeting distributed parameter systems can be found in [28,29], while an example of delay estimation technique was investigated in [30]. However, none of the proposed schemes are general and all focus on models of a particular structure. To the best knowledge of the authors *there is no reported procedure capable of adaptively estimating parameters of linear models regardless their structure, and on the manner in which they are influenced by the unknown parameters*.

Motivated by these observations, the aim of this paper is to propose a solution to adaptive on-line parameter estimation problem applicable to linear systems with transfer functions of arbitrary structure: rational, fractional, distributed, etc. A novel method based on a two-stage identification (optimization) process is proposed. In the first stage, values of the transfer function (real and imaginary part, or equivalently amplitude and phase) at specific frequencies are pinpointed. This is a convex optimization problem which can be effectively solved by conventional adaptive techniques, including the Recursive Least Square algorithm with forgetting factor (FF-RLS). In the second stage, the unknown parameters are found by inverting complex relation linking them to the previously identified transfer function values. In rare occasions, this inversion can be made explicitly. However, in most cases the solution must be found numerically by means of a particularly chosen iterative procedure. Within the present work we propose the gradient algorithm [17], since it is simple and computationally cheap, which makes it particularly suitable for on-line applications and implementations in real-time environments. In spite of being simple, we have found that gradient algorithm converges with acceptable rate in all considered cases.

The paper is organized as follows: Section 2 contains problem formulation and also explains the main idea underlying the proposed approach. Detailed explanation of the proposed two-stage identification process is given in Section 3. Numerical studies of several carefully selected examples are presented in Section 4. Concluding remarks are given in the final Section 5.

2. Problem formulation and the main idea

We consider a linear time-invariant (LTI) single-input single-output (SISO) system described by transfer function $G(s;\theta)$, where s is the Laplace variable and $\theta \in \mathbb{R}^q$ $(q \in \mathbb{N})$ is the vector of unknown parameters. Let $u(t)$ and $y(t)$ be the input and output signals, respectively, both of which are available. With slight abuse of notation, we will use

$$y(t) = G(s;\theta)u(t) \tag{2}$$

to denote that $y(t)$ is the response of $G(s;\theta)$ to the input signal $u(t)$. More precisely (2) should be interpreted as

$$y(t) = g(t;\theta) * u(t) = \int_0^t g(\tau;\theta)u(t-\tau)d\tau \tag{3}$$

where $g(t;\theta) = \mathcal{L}^{-1}\{G(s;\theta)\}$, $\mathcal{L}^{-1}$ is the inverse Laplace Transform, and $*$ denotes convolution. Let us also introduce

$$\hat{y}(t,\hat{\theta}(t)) = G(s,\hat{\theta}(t))u(t) = \int_0^t g(\tau;\hat{\theta}(t))u(t-\tau)d\tau; \tag{4}$$

as the output of the process model assuming that the input is u, and that the actual parameters were fixed to $\hat{\theta}(t)$ throughout the process history. In the sequel, "hat" symbol $(\hat{\cdot})$ will always be used to denoted estimated values, while the "tilde" symbol $(\tilde{\cdot})$ will denote estimation errors, i.e. differences between actual and estimated values.

Consider the following identification problem: Assuming that the structure of $G(s,\theta)$ is known, but that θ is unknown, find $\hat{\theta}$ minimizing the estimation error

$$J(\hat{\theta}) = \sum_{i=1}^{N}\Big(y(t_i,\theta) - \hat{y}(t_i,\hat{\theta})\Big)^2 \tag{5}$$

where t_i are time instances in which outputs have been recorded. We assume, without loss of generality, that the output sampling has been uniform, i.e. that $t_i = t_0 + ih$, where t_0 is the initial time instant and $h > 0$ is the sampling interval.

As discussed in the Introduction, the problem of minimizing (5) with respect to $\hat{\theta}$ is usually not convex, and does not, in general, have a unique solution. Furthermore, this is a *principal difficulty*, since by [16] every identification problem having a unique global optimizer must reduce to a linearly parametrized one. Therefore, it is not that the proper algorithm is not yet found, such an algorithm cannot exist at all. Consequently, any adaptive technique may be usable only locally, provide that sufficiently good initial estimates of θ are available.

The main idea of our work is to extract as much information as possible regarding the process behavior globally, and then to use local techniques to pinpoint the exact values of the parameters. We consider a re-parametrization $\theta \leftrightarrow \beta$, where β consists of real and imaginary parts of the transfer function evaluated at *a priori* chosen set of input frequencies. The new parameters can be identified globally, and the mapping form θ to β is usually locally bijective and differentiable.

Therefore, we propose a hybrid, two-stage scheme. In the first stage, β is identified using Recursive Least Square algorithm with forgetting factor (FF-RLS). In the second stage, θ is obtained by inverting $\beta = \beta(\theta)$. Since in most cases this inversion cannot be