Automatica 50 (2014) 3276–3280

Contents lists available at ScienceDirect

Automatica

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

Brief paper

A continuous-time framework for least squares parameter estimation[☆]

Samer Riachy[1]

ECS-Lab, EA 3649, Ecole Nationale Supérieure de l'Electronique et ses Applications, 6 avenue du Ponceau, 95014, Cergy, France
Non-A team, INRIA Lille-Nord-Europe, Lille, France

ARTICLE INFO

Article history:
Received 30 September 2013
Received in revised form
14 June 2014
Accepted 3 August 2014
Available online 27 October 2014

Keywords:
Evolution equations
Parameter estimation
Least squares
Sobolev spaces
Estimation under noise

ABSTRACT

This paper proposes a continuous-time framework for the least-squares parameter estimation method through evolution equations. Nonlinear systems in the standard state space representation that are linear in the unknown, constant parameters are investigated. Two estimators are studied. The first one consists of a linear evolution equation while the second one consists of an impulsive linear evolution equation. The paper discusses some theoretical aspects related to the proposed estimators: uniqueness of a solution and an attractive equilibrium point which solves for the unknown parameters. A deterministic framework for the estimation under noisy measurements is proposed using a Sobolev space with negative index to model the noise. The noise can be of large magnitude. Concrete signals issued from an electronic device are used to discuss numerical aspects.

1. Introduction

Least squares (LS) is by far the most popular method for parameter estimation. It has been developed under different guises. Discrete and continuous-time approaches were proposed for discrete and continuous-time, linear and nonlinear systems (Ioannou & Fidan, 2006; Krstic, 2009; Krstic, Kanellakopoulos, & Kokotović, 1995; Krstic & Kokotović, 1995; Ljung, 1999; Sastry & Bodson, 1989; Sastry & Isidori, 1989). The present work is concerned with continuous-time frameworks for LS. It seems that continuous-time LS estimators have been mainly developed in the context of adaptive control (Ioannou & Fidan, 2006; Krstic, 2009; Krstic et al., 1995; Krstic & Kokotović, 1995; Sastry & Bodson, 1989; Sastry & Isidori, 1989). In general, such estimators consist of a set of ordinary differential equations fed by the system input and output data. Major advantages of such estimators reside in the real-time implementability and suitability to an adaptive control loop. Nonetheless, during its convergence, the estimator should be continuously fed by the system data. Consequently, depending on the convergence rate, the system data can be needed on a big interval of time[2] which may not be possible in some applications.[3] Typical examples are unstable plants or plants with restrictions on the state vector (linear actuators for example) when a certainty equivalence controller or a primary stabilizing controller is not available. Therefore, it is quite interesting to develop a continuous-time theoretical basis to account for LS estimation when the data is available on a bounded interval of time. The present work proposes such a continuous-time framework. The approach relies on evolution equations (infinite dimension) and thus cannot be integrated in an adaptive control scheme. In order to highlight the practical utility of the approach, a discrete implementation is given. The computation cost is evaluated and shown to be carried out by low-cost real-time microcontrollers in order to accomplish fast parameter estimation for online plants.

An underlying problem to the design of a continuous-time estimator resides in the unavailability of the derivative of the state vector.[4] The standard solution (Ioannou & Fidan, 2006; Krstic et al., 1995; Sastry & Bodson, 1989) uses a filtered version of the

☆ The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Andrea Serrani under the direction of Editor Miroslav Krstic.

E-mail address: riachy@ensea.fr.

1 Tel.: +33 130736638; fax: +33 130736641.

2 This time interval is the whole positive real line in theory due to asymptotic convergence while in practice it can be considerably reduced by an efficient tuning of the estimator.

3 The same reasoning applies also to recursive discrete-time least squares (Ljung, 1999) which are not addressed in this work.

4 Assuming that the state vector is measured.

http://dx.doi.org/10.1016/j.automatica.2014.10.055

system equations. The basics of the approach can be seen on the scalar example $y'(\tau) = \theta y(\tau)$ where θ denotes the unknown parameter. With $\check{y}(\lambda)$ to denote the Laplace transform of $y(\tau)$, one has $\frac{1}{\lambda+1}(\lambda\check{y} - y_0) = \theta\frac{1}{\lambda+1}\check{y}$. Then the algebraic equation $Y_1 - e^{-\tau}y_0 = Y_2\theta$ is used in the LS estimator with $\check{Y}_1 = \frac{\lambda}{\lambda+1}\check{y}$ and $\check{Y}_2 = \frac{1}{\lambda+1}\check{y}$. A typical estimator is given by $\dot{\hat{\theta}} = \frac{Y2(Y_1-e^{-\tau}y_0-Y_2\hat{\theta})}{\sqrt{\kappa+Y_2^2}}$ where κ is a positive constant and the normalization $(\sqrt{\kappa + Y_2^2})^{-1}$ enhances the convergence rate of the estimator. A quite interesting approach that does not rely on estimating the state derivative is developed in Krstic (2009). An ingenious method (Fliess, Join, & Sira-Ramirez, 2008; Fliess & Sira-Ramirez, 2003) based on the algebraic derivative concept permits to obtain an estimator that does not depend neither on the state derivative nor the initial condition. Let us illustrate the basics of this method on our scalar example. Differentiating $\lambda\check{y} - y_0 = \theta\check{y}$ once with respect to λ gives $\lambda^{-2}\left[\check{y} + \lambda\frac{d}{d\lambda}\check{y} - \theta\frac{d\check{y}}{d\lambda}\right] = 0$. Notice that y_0 has disappeared. In addition, the multiplication by λ^{-2} permits us to obtain integrals in the time domain $\int_0^\tau(\tau - 2\tau_1)yd\tau_1 = \theta\int_0^\tau(\tau - \tau_1)(-\tau_1)yd\tau_1$. The integrals offer a low-pass filtering effect. Thus one can solve for θ whenever the right-hand side is different from zero. The approach of Fliess et al. (2008) and Fliess and Sira-Ramirez (2003) has been specialized to noisy signal derivation in Mboup, Join, and Fliess (2009) and shown to admit a least squares interpretation. The method offers a systematic approach to annihilate initial conditions as well as structured perturbations. In order to deal with the state derivative in the present work, the computation technique summarized in the scalar example (Fliess et al., 2008; Fliess & Sira-Ramirez, 2003) is applied prior to the estimator design.

The contribution of this work is twofold. First, LS estimators are developed which are suitable when the systems data are available on a bounded interval of time. Second, the proposed framework provides a simple deterministic description for the LS estimation under noisy measurements where the noise can be of large magnitude. Note that bounded unknown functions are usually used to model the noise in a deterministic context.

The paper is organized as follows. Two LS estimators, the first one based on a linear evolution equation, and the second one based on an impulsive evolution equation, are proposed and discussed in Section 3. The deterministic framework for the estimation under noisy measurements is proposed in Section 4 for linear input–output systems. Section 5 proposes a numerical implementation of the estimators. Section 6 is dedicated to an experimental validation in order to discuss the numerical aspect. Let us start in Section 2 with a brief recall about evolution equations in order to clarify subsequent developments.

2. Elementary notions about evolution equations

The content of this section can be found in any introductory textbook to the theory of evolution equations (see for example Engel & Nagel, 2000, Evans, 1997, Pazy, 1983 and Zheng, 2004). Let $\{S(t); t \geq 0\}$ be a family of linear operators defined on a Banach space $\mathcal{B}$. $S(t)$ is said to be a linear semigroup on $\mathcal{B}$ if $S(0) = I$ (I denotes the identity operator on $\mathcal{B}$) and $S(t_1 + t_2) = S(t_2)S(t_1) = S(t_1)S(t_2)$. It is said to be a semigroup of contractions if, moreover, $\|S(t)\| \leq 1$ where $\|\cdot\|$ is an operator norm on $\mathcal{B}$. In addition, the semigroup is strongly continuous if $\lim_{h\to+0} S(h)\xi = \xi$. The operator A defined by $A := \lim_{h\to+0}\frac{S(h)-I}{h}$ is called the infinitesimal generator of $S(t)$. The Hille–Yosida theorem provides necessary and sufficient conditions an operator A should satisfy in order to be a generator of a semigroup of contractions. Let us denote by $\mathcal{D}(A) \subset \mathcal{B}$ the domain of definition of an operator A.

Theorem 1 (*Hille–Yosida, Zheng, 2004*). *The linear operator $A : \mathcal{D}(A) \subset \mathcal{B} \mapsto \mathcal{B}$ is the infinitesimal generator of a linear semigroup of contractions if, and only if,*

(1) *A is a densely defined ($\mathcal{D}(A)$ is dense in $\mathcal{B}$) and closed operator in $\mathcal{B}$,*
(2) *$\forall\lambda > 0,\ \lambda I - A$ is a one-to-one and onto mapping,*
(3) *$\|(\lambda I - A)^{-1}\| \leq \frac{1}{\lambda}$.*

A linear homogeneous evolution equation is a system given by ($\dot{\tilde{\Theta}} := \frac{d\tilde{\Theta}}{dt}$):

$$\dot{\tilde{\Theta}} = A\tilde{\Theta}; \qquad \tilde{\Theta}_{t=0} = \tilde{\xi}, \quad \tilde{\xi} \in \mathcal{D}(A) \tag{1}$$

where A is the infinitesimal generator of a strongly continuous semigroup. A (Banach space valued) function $\tilde{\Theta}(t) : [0, +\infty) \mapsto \mathcal{B}$ is said to be a solution of (1) if $\tilde{\Theta}(t) \in \mathfrak{C}^1([0,\infty); \mathcal{D}(A))$ such that (1) is satisfied. $\mathfrak{C}^1$ represents the space of continuously differentiable functions defined on $\mathcal{B}$. The existence of a unique solution for (1) is ensured by the following theorem.[5]

Theorem 2 (*Existence and Uniqueness*). *If $A : \mathcal{D}(A) \subset \mathcal{B} \mapsto \mathcal{B}$ generates a strongly continuous semigroup of contractions $S(t)$ on $\mathcal{B}$ then $\forall\tilde{\xi} \in \mathcal{D}(A)$ the system (1) admits a unique solution given by $\tilde{\Theta}(t) = S(t)\tilde{\xi}$.*

3. The estimators

3.1. Problem description and basic assumptions

Consider the nonlinear system which is linear in the parameter Θ:

$$x'(\tau) = \phi(x,u) + \varphi(x,u)\Theta \tag{2}$$

where x and u are scalar real variables while $\Theta = [\theta_1, \dots, \theta_p]^T \in \mathbb{R}^p$ is the vector of unknown constant parameters. The scalar field $\phi(x,u) : \mathbb{R}\times\mathbb{R} \mapsto \mathbb{R}$ and the vector field $\varphi = [\varphi_1, \dots, \varphi_p]$ with $\varphi_i(x,u) : \mathbb{R}\times\mathbb{R} \mapsto \mathbb{R},\ i = 1, \dots, p$ are known and satisfy standard assumptions about the existence of a unique classical solution for (2) for a given initial condition. A scalar system is considered in order to simplify the presentation, and the extension to multidimensional systems can be done straightforwardly if the state vector is accessible for measurements. Let $\mathbb{T}$ be a bounded and connected subset of $\mathbb{R}$. Given a solution of (2) on $\mathbb{T}$, it is legitimate to identify $\phi(x(\tau), u(\tau))$ and $\varphi(x(\tau), u(\tau))$ with functions depending on τ only. Thus the notations $\phi(\tau) := \phi(x(\tau), u(\tau))$ and $\varphi_i(\tau) := \varphi_i(x(\tau), u(\tau)),\ \tau \in \mathbb{T}$ are adopted. Moreover, the following is assumed.

Assumption 3.1. $u(\tau), \phi(\tau)$ and $\varphi_i(\tau) \in \mathcal{L}^2(\mathbb{T})$. $\mathcal{L}^2(\mathbb{T})$ is the space of square summable functions on $\mathbb{T}$.

Let ϵ be a positive constant and denote by $\mathbb{D}$ the set $\mathbb{D} = [-\epsilon, \epsilon]$. In order to deal with the derivative x', let $v(s),\ s \in \mathbb{D}$, be a continuously differentiable function, supported on $\mathbb{D}$ such that $v(-\epsilon) = v(\epsilon) = 0$. Introduce $\mathbb{T}_\epsilon$, an ϵ-neighborhood of $\mathbb{T}$ such that $\mathbb{T} \subseteq \mathbb{T}_\epsilon \subset \mathbb{R}$ and $T_\epsilon = T_1 + 2\epsilon$. T_ϵ, T_1 and 2ϵ are the Lebesgue measures of $\mathbb{T}_\epsilon, \mathbb{T}$ and $\mathbb{D}$ respectively. Define the convolutions $x'_\epsilon = v\star x'$, $\phi_\epsilon = v\star\phi$ and $\varphi_{\epsilon,i} = v\star\varphi_i$. They are supported on $\mathbb{T}_\epsilon$ since x', ϕ and φ are supported on $\mathbb{T}$. Moreover, notice that $-\frac{dv}{ds}\star x = v\star x'$. Such $v(s)$ can be given by

$$v(s) = (\epsilon^2 - s^2)\chi_{\mathbb{D}} \tag{3}$$

where $\chi_{\mathbb{D}}$ is the indicator function of $\mathbb{D}$. Convolving (2) with (3) leads to:

$$x'_\epsilon(\tau) = \phi_\epsilon(\tau) + \varphi_\epsilon(\tau)\Theta, \quad \tau \in \mathbb{T}_\epsilon. \tag{4}$$

[5] Theorem 2 corresponds to Proposition 6.2 p. 145 of Engel and Nagel (2000).