Statistics and Probability Letters 132 (2018) 40–45

Contents lists available at ScienceDirect

Statistics and Probability Letters

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

Consistency of direct integral estimator for partially observed systems of ordinary differential equations

Ivan Vujačić [a], Itai Dattner [b,*]

[a] Department of Mathematics, VU University Amsterdam, Room S3.30, De Boelelaan 1081a, 1081HV Amsterdam, The Netherlands
[b] Department of Statistics, University of Haifa, 199 Aba Khoushy Ave. Mount Carmel, Haifa 3498838, Israel

ARTICLE INFO

Article history:
Received 5 March 2017
Received in revised form 9 August 2017
Accepted 22 August 2017
Available online 14 September 2017

Keywords:
Consistency
Ordinary differential equations
Nonparametric regression
Sieve extremum estimators

ABSTRACT

In this paper we use the sieve framework to prove consistency of the 'direct integral estimator' of parameters for partially observed systems of ordinary differential equations, which are commonly used for modeling dynamic processes.

1. Introduction

Dynamic systems are ubiquitous in nature and are used to model many processes in biology, chemistry, physics, medicine, and engineering. In particular, systems of ordinary differential equations (ODEs) are commonly used for the mathematical modeling of the rate of change of dynamic processes. In many practical applications, the process can only be partially measured, a fact that renders estimation of parameters of the system extremely challenging. Recently, a 'direct integral estimator' for partially observed ODEs was introduced (Dattner, 2015), and its finite sample performance was demonstrated using both synthetic and real data. In addition, the practical performance of the integral estimator was demonstrated in a challenging biological setup of partially observed systems of ODEs (Dattner et al., 2017). However, the theoretical properties of the direct integral estimator for partially observed systems of ODEs were not derived. In this paper we use the sieve framework to prove consistency of the direct integral estimator.

The process of interest is usually modeled by the ODE system

$$\begin{cases} \mathbf{x}'(t) = \boldsymbol{F}(\mathbf{x}(t);\boldsymbol{\theta}),\ \ t \in [0,T], \\ \mathbf{x}(0) = \boldsymbol{\xi}, \end{cases} \tag{1}$$

where $\mathbf{x}(t) \in \mathbb{R}^d$, $\boldsymbol{\xi} \in \Xi \subset \mathbb{R}^d$, $\boldsymbol{\theta} \in \boldsymbol{\Theta} \subset \mathbb{R}^p$, T is a known positive constant and $\boldsymbol{F} : \mathbb{R}^d \times \mathbb{R}^p \to \mathbb{R}^d$ is a known function of $\mathbf{x}(t)$ and $\boldsymbol{\theta}$. Given the values of $\boldsymbol{\xi}$ and $\boldsymbol{\theta}$, we denote the solution of (1) by $\mathbf{x}(t) = \mathbf{x}(t;\boldsymbol{\theta},\boldsymbol{\xi}),\ t \in [0,T]$. The aim is to estimate the unknown parameter $\boldsymbol{\theta}$ (and if necessary $\boldsymbol{\xi}$) from noisy observations

$$Y_j(t_i) = x_j(t_i;\boldsymbol{\theta},\boldsymbol{\xi}) + \varepsilon_{j,i},\ \ i = 1,\ldots,n, j = 1,\ldots,r, \tag{2}$$

* Corresponding author.
E-mail address: idattner@stat.haifa.ac.il (I. Dattner).

http://dx.doi.org/10.1016/j.spl.2017.08.013
0167-7152/

where $0 \le t_1 < \cdots < t_n = T < \infty$, $r \le d$, and $\varepsilon_{j,i}$ is the unobserved measurement error for x_j at time t_i. Consider systems linear in the parameters that have the form

$$\boldsymbol{F}(\mathbf{x}(t);\boldsymbol{\theta}) = \mathbf{g}(\mathbf{x}(t))\boldsymbol{\theta}, \tag{3}$$

where $\mathbf{g} : \mathbb{R}^d \to \mathbb{R}^{d\times p}$ maps the d-dimensional column vector $\mathbf{x}$ into a $d \times p$ matrix. Define the system of integral equations

$$\mathbf{x}(t) = \boldsymbol{\xi} + \int_0^t \mathbf{g}(\mathbf{x}(s))\mathrm{d}s\,\boldsymbol{\theta},\;\; t \in [0,T], \tag{4}$$

which follows from (1) and (3) by integration. The case of systems linear in (functions of) the parameters and $r = d$, was studied in Dattner and Klaassen (2015) where a 'direct integral estimator' was developed, and its $\sqrt{n}$-consistency was derived. Specifically, in view of (4), Dattner and Klaassen (2015) suggested to estimate the parameters $\boldsymbol{\theta}$ and $\boldsymbol{\xi}$ by minimizing

$$\int_0^T \|\widehat{\mathbf{x}}(t) - \boldsymbol{\xi} - \int_0^t \mathbf{g}(\widehat{\mathbf{x}}(s))\mathrm{d}s\,\boldsymbol{\theta}\|^2 \mathrm{d}t \tag{5}$$

with respect to $\boldsymbol{\theta}$ and $\boldsymbol{\xi}$, where $\widehat{\mathbf{x}}(t)$, $t \in [0,T]$, is a specific estimator of $\mathbf{x}(t;\boldsymbol{\theta},\boldsymbol{\xi})$. Here and subsequently, $\|\cdot\|$ denotes the Euclidean norm for a vector argument and the Frobenius norm for a matrix argument. We now describe the construction of the estimator for partially observed systems (when $r < d$) developed in Dattner (2015). Let $\mathcal{M}$ and $\mathcal{U}$ denote the sets of r-dimensional and $(d-r)$-dimensional vector functions on $[0,T]$ that correspond to $\mathbf{m}(\cdot) = \mathbf{m}(\cdot,\boldsymbol{\theta},\boldsymbol{\xi})$ and $\mathbf{u}^*(\cdot) = \mathbf{u}^*(\cdot,\boldsymbol{\theta},\boldsymbol{\xi})$, the measured and unmeasured 'true' solutions of (1), respectively. We first construct an estimator $\widehat{\mathbf{m}}_n(\cdot)$ of $\mathbf{m}(\cdot)$ using the observations (2) and for a given $\mathbf{u} \in \mathcal{U}$, we define

$$\begin{aligned}
&\widehat{\mathbf{x}}_{\mathbf{u}}(t) = (\widehat{\mathbf{m}}_n(t), \mathbf{u}(t)),\;\; \widehat{\mathbf{G}}_{\mathbf{u}}(t) = \int_0^t \mathbf{g}(\widehat{\mathbf{x}}_{\mathbf{u}}(s))\mathrm{d}s,\\
&\widehat{\mathbf{A}}_{\mathbf{u}} = \int_0^T \widehat{\mathbf{G}}_{\mathbf{u}}(t)\mathrm{d}t,\;\; \widehat{\mathbf{B}}_{\mathbf{u}} = \int_0^T \widehat{\mathbf{G}}_{\mathbf{u}}^\top(t)\widehat{\mathbf{G}}_{\mathbf{u}}(t)\mathrm{d}t.
\end{aligned} \tag{6}$$

Plugging $\widehat{\mathbf{x}}_{\mathbf{u}}(t)$ from (6) into (5), and minimizing with respect to $(\boldsymbol{\theta},\boldsymbol{\xi})$, the direct integral estimator is

$$\begin{aligned}
&\widehat{\boldsymbol{\xi}}_{\mathbf{u}} = \left(T\mathbf{I}_d - \widehat{\mathbf{A}}_{\mathbf{u}}\widehat{\mathbf{B}}_{\mathbf{u}}^{-1}\widehat{\mathbf{A}}_{\mathbf{u}}^\top\right)^{-1} \int_0^T \left\{\mathbf{I}_d - \widehat{\mathbf{A}}_{\mathbf{u}}\widehat{\mathbf{B}}_{\mathbf{u}}^{-1}\widehat{\mathbf{G}}_{\mathbf{u}}^\top(t)\right\}\widehat{\mathbf{x}}_{\mathbf{u}}(t)\mathrm{d}t,\\
&\widehat{\boldsymbol{\theta}}_{\mathbf{u}} = \widehat{\mathbf{B}}_{\mathbf{u}}^{-1} \int_0^T \widehat{\mathbf{G}}_{\mathbf{u}}^\top(t)\{\widehat{\mathbf{x}}_{\mathbf{u}}(t) - \widehat{\boldsymbol{\xi}}_{\mathbf{u}}\}\mathrm{d}t.
\end{aligned} \tag{7}$$

The estimators of the parameter and initial value are well-defined only if the inverse matrices in (7) exist. For a subset $\mathcal{U}_n \subset \mathcal{U}$ let

$$\widehat{\mathbf{u}}_n := \operatorname*{argmin}_{\mathbf{u}\in\mathcal{U}_n} M_n(\mathbf{u}), \tag{8}$$

where

$$M_n(\mathbf{u}) = \int_0^T \|\widehat{\mathbf{x}}_{\mathbf{u}}(t) - \widehat{\boldsymbol{\xi}}_{\mathbf{u}} - \widehat{\mathbf{G}}_{\mathbf{u}}(t)\widehat{\boldsymbol{\theta}}_{\mathbf{u}}\|^2 \mathrm{d}t. \tag{9}$$

The estimators of initial value $\boldsymbol{\xi}$ and the parameter $\boldsymbol{\theta}$ are

$$(\widehat{\boldsymbol{\theta}}_n, \widehat{\boldsymbol{\xi}}_n) := (\widehat{\boldsymbol{\theta}}_{\widehat{\mathbf{u}}_n}, \widehat{\boldsymbol{\xi}}_{\widehat{\mathbf{u}}_n}). \tag{10}$$

The aim of this paper is to establish consistency of $(\widehat{\boldsymbol{\theta}}_n, \widehat{\boldsymbol{\xi}}_n)$ given by (10) under suitable assumptions. In the case of fully observed systems Dattner and Klaassen (2015) showed that if the estimator $\widehat{\mathbf{x}}_n(\cdot) = (\widehat{\mathbf{m}}_n(\cdot), \widehat{\mathbf{u}}_n(\cdot))$ of $\mathbf{x}(\cdot)$ is consistent in the supnorm then $(\widehat{\boldsymbol{\theta}}_n, \widehat{\boldsymbol{\xi}}_n)$ are consistent estimators of $(\boldsymbol{\theta},\boldsymbol{\xi})$. However, in partially observed systems only $\mathbf{m}(\cdot)$ is observed. Therefore, based on the data we are only able to construct consistent estimator $\widehat{\mathbf{m}}_n(\cdot)$ of $\mathbf{m}(\cdot)$. In the next section we show that given an appropriate identifiability requirement, consistent $\widehat{\mathbf{m}}_n(\cdot)$ gives rise to consistent estimator $\widehat{\mathbf{u}}_n(\cdot)$ of $\mathbf{u}^*(\cdot)$, defined in (8).

2. Consistency of direct integral estimator

In this section we prove consistency of $\widehat{\mathbf{u}}_n$ given in (8), which in turn, will imply consistency of $(\widehat{\boldsymbol{\theta}}_n, \widehat{\boldsymbol{\xi}}_n)$. In order to achieve parameters consistency, we first require a consistent estimator $\widehat{\mathbf{m}}_n$. One can take $\widehat{\mathbf{m}}_n$ to be a local polynomial estimator as in Dattner and Klaassen (2015).