IOPscience iopscience.iop.org

Home Search Collections Journals About Contact us My IOPscience

LSSVM based initialization approach for parameter estimation of dynamical systems

This content has been downloaded from IOPscience. Please scroll down to see the full text.

2014 J. Phys.: Conf. Ser. 490 012004

(http://iopscience.iop.org/1742-6596/490/1/012004)

View the table of contents for this issue, or go to the journal homepage for more

Download details:

IP Address: 134.58.253.57
This content was downloaded on 03/04/2014 at 11:59

Please note that terms and conditions apply.

Journal of Physics: Conference Series **490** (2014) 012004 doi:10.1088/1742-6596/490/1/012004

LSSVM based initialization approach for parameter estimation of dynamical systems

Siamak Mehrkanoon, Rien Quirynen, Moritz Diehl and Johan A.K. Suykens

KU Leuven, ESAT-SCD, Kasteelpark Arenberg 10, B-3001 Leuven (Heverlee), Belgium

E-mail:
{Siamak.Mehrkanoon,Rien.Quirynen,Moritz.Diehl,Johan.Suykens}@esat.kuleuven.be

Abstract. In this study the estimation of parameters in dynamical systems governed by parameter-affine ordinary differential equations is explored. The described method by Mehrkanoon et al.∼ in [1] is utilized as an initialization of the nonlinear optimization problem for parameter estimation. In contrast to existing convex initialization approaches [2] that use a first order Euler discretization, we do not require any integration method to simulate the dynamical system. Furthermore, a denoising scheme using LSSVM is proposed to first filter the measured data then proceed with the filtered signals for parameter estimation problem. Experimental results demonstrate the efficiency of the proposed method, compared to alternative approaches on different examples from the literature.

1. Introduction

Mathematical models are widely used in various fields of application to describe physical systems. These models involve some unknown parameters that require to be estimated and the performance of the model largely depends on the way it is parameterized. In this study the estimation of parameters in dynamical systems governed by ordinary differential equations (ODEs) is explored. In general, during the process of parameter estimation one tries to make the differences between simulation results and the observational data as small as possible using an optimization algorithm. However, depending on the nature of the model, the optimization problem can be non-convex, potentially leading to multiple local minima. In the case of parameter-affine models, a convex formulation approach for initialization of the parameter estimation problem is described in [2]. In this approach in order to keep the approximate problem convex, one relies on a simple Euler discretization of the system. Recently Mehrkanoon et al.∼ [1] proposed a different approach based on Least Squares Support Vector Machines (LSSVM) for estimating the unknown parameters in ODEs. As opposed to the approach described in [2], the method introduced in [1] does not need to use any integration method to simulate the dynamical system. Therefore the drawbacks of using the Euler method concerning its stability region are removed. The aim of this paper is to first employ the method described in [1] to obtain an initial guess for the parameters and then to solve the original nonconvex problem. The latter is done using a multiple shooting discretization and constrained Gauss-Newton to solve the nonlinear programming problem (NLP). Moreover a scheme based on LSSVM is proposed to pre-process the data and use the filtered data in the convex approximate problem.


Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
Published under licence by IOP Publishing Ltd

2. Problem statement

Consider a dynamical system of the form: $\frac{dX}{dt} = F(t, X, \theta),\ X(0) = X_0$. In what follows, $X = [x_1, ..., x_m]^T$ is the state vector of the system, $\theta = [\theta_1, ..., \theta_p]^T$ are the unknown parameters of the system and X_0 are initial values. In order to estimate the unknown parameters θ, the state variable $X(t)$ is observed at N time instants $\{t_1, ..., t_N\}$, i.e.$\sim Y(t_i) = X(t_i) + E_i,\ i = 1, ..., N$, where $\{E_i\}_{i=1}^N$ are independent measurement errors with zero mean. The objective is to determine appropriate parameter values so that errors between the model prediction and the measured data are minimized.

3. Brief overview of the existing initialization approaches

Parameter estimation is typically formulated as following non-convex optimization problem:

$$\begin{aligned} \underset{X(t_0),...,X(t_N),\theta}{\text{minimize}} \quad & \frac{1}{2}\sum_{i=0}^{N} \|Y(t_i) - X(t_i)\|_2^2 \\ \text{subject to} \quad & X(t_{k+1}) = X(t_k) + \int_{t_k}^{t_{k+1}} F(\tau, X(\tau), \theta)\, d\tau, \quad k = 0, ..., N-1, \end{aligned} \tag{1}$$

Due to the nonconvexity coming from the nonlinear model, a Newton type method can only find locally optimal solutions. Depending on the initialization, one can obtain a different local solution. In case of parameter-affine models attempts have been made to provide a good initial guess through a convex optimization approach. A Least Squares Prediction Error Method (PEM) proposed in [3] is formulated as a convex problem to provide such an initial guess for (1), with parameter-affine function F, as follows:

$$\begin{aligned} \underset{X(t_0),...,X(t_N),\theta}{\text{minimize}} \quad & \frac{1}{2}\sum_{i=0}^{N} \|Y(t_i) - X(t_i)\|_2^2 \\ \text{subject to} \quad & X(t_{k+1}) = Y(t_k) + T_s F(t_k, Y(t_k), \theta), \quad k = 0, ..., N-1, \end{aligned} \tag{2}$$

where T_s is the sampling time. However the solution obtained by the PEM approach can be biased if the process and measurement noise are not modeled appropriately. Therefore the method does not perform well in the presence of noisy data and one needs to filter the residual errors. On the other hand, the authors in [2] proposed a so-called Least Squares Convex Approach (CA). In contrast to the PEM approach, there is no need for filtering the residual error in CA formulation [2]:

$$\begin{aligned} \underset{X(t_0),...,X(t_N),\theta}{\text{minimize}} \quad & \frac{1}{2}\sum_{i=0}^{N} \|Y(t_i) - X(t_i)\|_2^2 \\ \text{subject to} \quad & X(t_{k+1}) = X(t_k) + T_s F(t_k, Y(t_k), \theta), \quad k = 0, ..., N-1. \end{aligned} \tag{3}$$

However in this approach in order to keep the optimization problem (3) convex, one still has to rely on a simple Euler discretization of the system. Recently, Mehrkanoon et al.$\sim$ [1] proposed an approach based on Least Squares Support Vector Machines (LSSVM) [4] for parameter estimation of ODEs. Closed-form approximate models for the state and its derivative are first derived from the observed data by means of LSSVM. The time-derivative information is then substituted into the given dynamical system, reducing the parameter estimation problem into an algebraic optimization problem. The problem is formulated as the following convex optimization problem:

$$\underset{\theta}{\text{minimize}} \quad \frac{1}{2}\sum_{i} \left\| \frac{d}{dt}\hat{X}(t_i) - F(t_i, \hat{X}(t_i), \theta) \right\|_2^2 \tag{4}$$

where $\hat{X}(t_i)$ and $\frac{d}{dt}\hat{X}(t_i)$ are obtained using the LSSVM model (see [1, 5] for more details). As opposed to the previous approaches, one does not need to use any integration method to simulate the dynamical system. Therefore the drawbacks of using the Euler method, concerning its stability region, are removed. We refer further to this approach as LSSVM.

4. Pre-processing using LSSVM

Data pre-processing plays a very important role in many applications. We will make use of the LSSVM regression ability to reduce the effect of noise and as a result having a smoother signal to proceed with. Given observational data $Y(t) = [y_1(t), \ldots, y_m(t)]^T$, and assuming a model of the form $\hat{y}_k(t) = w_k^T\varphi(t) + b_k$ for $y_k(t)$, $k = 1, \ldots, m$, the LSSVM regression is formulated as the following optimization problem [4]:

$$\begin{aligned} \underset{w_k, b_k, e_k(t_1), \ldots, e_k(t_N)}{\text{minimize}} \quad & \frac{1}{2}w_k^T w_k + \frac{\gamma_k}{2}\sum_{i=1}^{N} e_k(t_i)^2 \\ \text{subject to} \quad & y_k(t_i) = w_k^T\varphi(t_i) + b_k + e_k(t_i), \quad i = 1, ..., N, \end{aligned} \tag{5}$$

where $\gamma_k \in \mathbb{R}^+, b_k \in \mathbb{R}$, $w_k \in \mathbb{R}^h$, $\varphi(\cdot) : \mathbb{R} \to \mathbb{R}^h$ is the feature map and h is the dimension of the feature space. The dual solution is then given by

$$\left[\begin{array}{c|c} \Omega + \gamma^{-1}I_N & 1_N \\ \hline 1_N^T & 0 \end{array}\right] \left[\begin{array}{c} \alpha^k \\ \hline b_k \end{array}\right] = \left[\begin{array}{c} y^k \\ \hline 0 \end{array}\right] \tag{6}$$

where $\Omega_{ij} = K(t_i, t_j) = \varphi(t_i)^T\varphi(t_j)$ is the (i, j)-th entry of the positive definite kernel matrix. $1_N = [1, ..., 1]^T \in \mathbb{R}^N$, $\alpha^k = [\alpha_1^k, ..., \alpha_N^k]^T$, $y^k = [y_k(t_1), ..., y_k(t_N)]^T$ and I_N is the identity matrix. The model in dual form becomes: $\hat{y}_k(t) = w_k^T\varphi(t) + b_k = \sum_{i=1}^{N}\alpha_i^k K(t_i, t) + b_k$ where K is the kernel function. The signal $\hat{y}_k(t)$ can be considered as a denoised version of the $y_k(t)$ measurements. The idea now would be to replace the measurements $Y(t)$ in PEM and CA formulations (2) and (3) by $\widehat{Y}(t)$ where $\widehat{Y}(t) = [\hat{y}_1(t), \ldots, \hat{y}_m(t)]^T$. These schemes will be referred to as PEM+LSSVM and CA+LSSVM respectively.

5. Numerical Results

Five initialization approaches i.e.∼ PEM, PEM+LSSVM, LSSVM, CA and CA+LSSVM are considered for providing a good starting point for solving the original nonlinear least squares parameter estimation problem (1). Then the constrained Gauss-Newton method (CGN) is applied for solving the multiple shooting formulation in (1).

Problem 5.1: Consider the Lorenz equation

$$\frac{dx_1}{dt} = \theta_1(x_2 - x_1), \quad \frac{dx_2}{dt} = x_1(\theta_2 - x_3) - x_2, \quad \frac{dx_3}{dt} = x_1x_2 - \theta_3x_3 \tag{7}$$

where $\theta = (\theta_1, \theta_2, \theta_3)$ are the unknown parameters of the system. The initial values at $t = 0$ are $(-9.42, -9.34, 28.3)$ and the correct parameter values to be estimated are $\theta = (10, 28, 8/3)$.

Problem 5.2 Consider Barne's problem [1]:

$$\frac{dx_1}{dt} = \theta_1x_1 - \theta_2x_1x_2, \quad \frac{dx_2}{dt} = \theta_2x_1x_2 - \theta_3x_2, \tag{8}$$

with initial state values $(1.00, 0.3)$ and $\theta = (\theta_1, \theta_2, \theta_3) = (0.86, 2.07, 1.81)$ as the true unknown parameters of the system.

Numerical results illustrating the performance of different initialization methods in terms of Mean Squared Error (MSE) can be found in Figure 1(a) and (b), respectively using 300 measurements at 0.04s (Lorenz) and 200 measurements at 1s sampling time (Barne). In Figure 1,

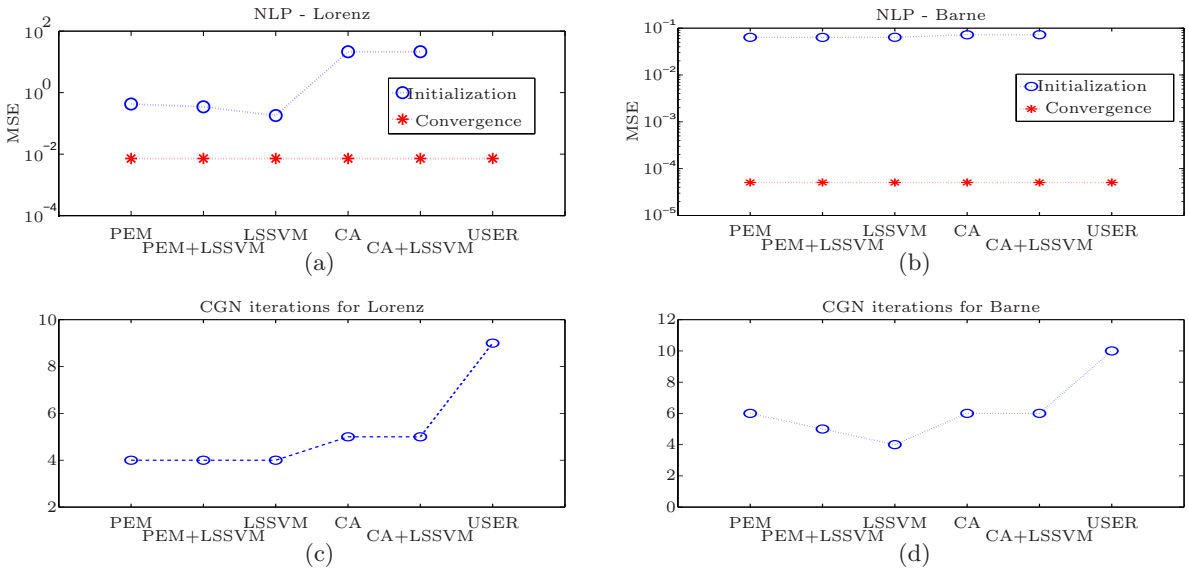

Figure 1. (a) and (b) The obtained Mean Squared Error for Problem 5.1 and 5.2 respectively. (c) and (d) Number of required Newton iterations needed to satisfy the given tolerance.

the case where the user is providing some starting point, based upon available prior knowledge, is referred to as USER initialization approach. The implementations and simulations were carried out in MATLAB and for the discretization of (1), the ACADO integrators were used as presented in [6]. Figure 1 shows that LSSVM based initialization is comparable to that of conventional convex initialization approaches and for Barne system it requires the least number of iterations to converge.

6. Conclusion

In this paper we presented an alternative method based on LSSVM for the initialization of nonlinear least squares parameter estimation problems. As opposed to conventional approaches the proposed method does not need to simulate the given dynamical system in order to provide a good approximate solution.

Acknowledgments

This work was supported by: • Research Council KUL: GOA/10/09 MaNet, PFV/10/002 (OPTEC), several PhD/postdoc & fellow grants • Flemish Government: ○ IOF: IOF/KP/SCORES4CHEM; ○ FWO: PhD/postdoc grants, projects: G.0320.08 (convex MPC), G.0558.08 (Robust MHE), G.0557.08 (Glycemia2), G.0588.09 (Brain-machine), G.0377.09 (Mechatronics MPC); G.0377.12 (Structured systems) research community (WOG: MLDM); ○ IWT: PhD Grants, projects: Eureka-Flite+, SBO LeCoPro, SBO Climaqs, SBO POM, O&O-Dsquare • Belgian Federal Science Policy Office: IUAP P7/ (DYSCO, Dynamical systems, control and optimization, 2012-2017) • IBBT • EU: ERNSI, FP7-EMBOCON (ICT-248940), FP7-SADCO (MC ITN-264735), ERC ST HIGHWIND (259 166), ERC AdG A-DATADRIVE-B • COST: Action ICO806: IntelliCIS • Contract Research: AMINAL • Other: ACCM. Johan Suykens is a professor at the KU Leuven, Belgium.

References

[1] Mehrkanoon S, Falck T, Suykens J A K 2012 Parameter Estimation for Time Varying Dynamical Systems using Least Squares Support Vector Machines, *In Proc. of the 16th IFAC Symposium on System Identification (SYSID 2012), Brussels, Belgium*, pp. 1300-1305.

[2] Bonilla J, Diehl M, De Moor B and Van Impe J 2008 A Nonlinear Least Squares Estimation Procedure without Initial Parameter Guesses, *Proceedings of the 47th IEEE Conference on Decision and Control* Cancun, Mexico, pp. 5519-5524.

[3] Ljung L 1999 *System Identification, Theory for the Users*, Prentice Hall, NJ.

[4] Suykens J A K, Van Gestel T, De Brabanter J, De Moor B and Vandewalle J 2002 *Least Squares Support Vector Machines* (World Scientific, Singapore.)

[5] Mehrkanoon S and Suykens J A K 2012 LS-SVM approximate solution to linear time varying descriptor systems *Automatica*, **48**, 2502-2511.

[6] Quirynen R, Gros S, Diehl M 2013 Fast auto generated ACADO integrators and application to MHE with multi-rate measurements, *Proceedings of the European Control Conference*, accepted.