Data-driven delay identification with SINDy

Ákos Tamás Köpeczi-Bócz*, Henrik Sykora ** and Dénes Takács*
*Department of Applied Mechanics, Budapest University of Technology and Economics, HU
**Institute of Sound and Vibration Research, University of Southampton, UK

Abstract. In this work, we investigate the capabilities of the *Sparse Identification of Nonlinear Dynamics* method for time-delay identification. We test the robustness and effectiveness of the method through data generated using reference systems with known time-delay.

Introduction

Model construction based on data-driven techniques has gained considerable ground over the past years due to the more versatile measurement toolset (such as image recognition, smartphone sensors, etc.) and a large amount of available data. The development of computer sciences, statistical and machine learning tools enable more efficient data processing and model discovery. The appearance of Scientific Machine Learning [1] (SciML) made the model construction more sophisticated by opening a physics-informed toolset for us to discover phenomena described by dynamical systems. Throughout this work, we investigate one of these techniques, the *Sparse Identification of Nonlinear Dynamics* (SINDy) in presence of time delay. The core idea of the method was first presented by Brunton et al. in 2016 [2], that we modify to find the underlying physics of time-delayed systems. We investigate the limits and the prerequisites of this technique through numerical experiments using a dataset constructed by simulating dynamic systems with known parameters and time-delays. We also include a stochastic effect to test out the robustness of the method. We apply the SINDy with the sequentially thresholded least squares algorithm (STLSQ) [3].

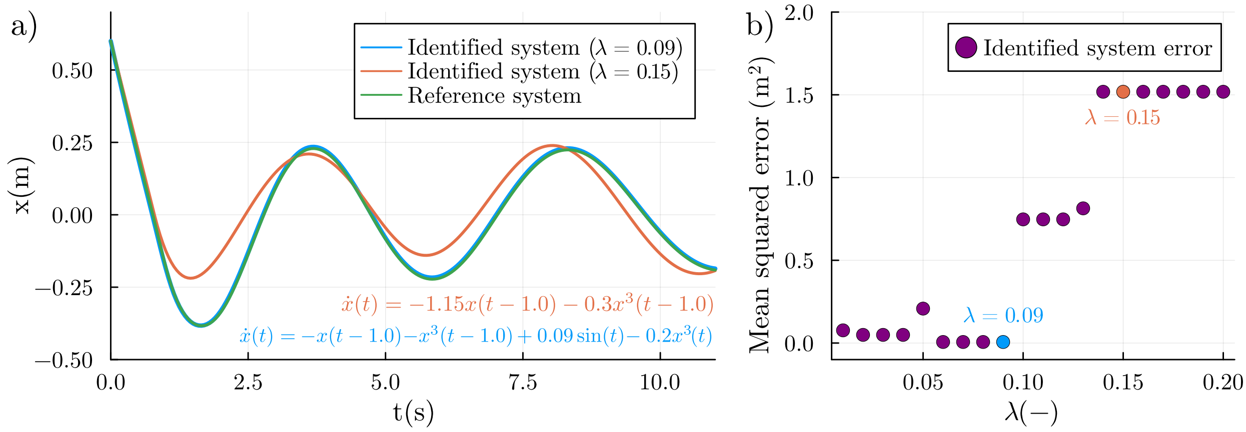

Figure 1: (a) Two identified systems for different λ threshold values of the STLSQ algorithm. The reference system for this case was $\dot{x}(t) = -x^3(t-\tau) - x(t-\tau) + 0.09\sin(t)$ with $\tau = 1.0$ s. By increasing the value of λ we get a system with increased error on behalf of having less terms to describe it. (b) The errors of the identified systems are shown with respect to the λ values. Identification of the best fitting solution is possible this way. Two cases are highlighted with $\lambda = 0.09$ and $\lambda = 0.15$ thresholds in correspondence with panel a. Note, that even in cases when the error was larger the identified time-delay was correct.

Results and discussion

We have found that on simulated non-linear systems the SINDy algorithm worked well with the proposed method, see example in Fig.1(a). To successfully achieve a sufficient fit we identified the proper threshold value λ for the STLSQ algorithm by sweeping through a range of different values and finding the one with the lowest error as in Fig.1(b). Despite of falsely identified terms, the found time-delay was correct. In our study, the physically meaningful range of delay has to be determined for the algorithm and the candidate terms have to be introduced with a sparse delay distribution. This method was useful in cases when the behavior of the system could be described by expected terms. The developed method is also capable to identify multiple time delays in the same system which is also the subject of our research.

References

[1] Christopher Rackauckas, Yingbo Ma, Julius Martensen, Collin Warner, Kirill Zubov, Rohit Supekar, Dominic Skinner, Ali Ramadhan, Alan Edelman (2020) Universal Differential Equations for Scientific Machine Learning DOI: 10.48550/ARXIV.2001.04385

[2] Steven L. Brunton, Joshua L. Proctor, and J. Nathan Kutz. (2016) Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proceedings of the National Academy of Sciences* **113(15)**: 3932–3937.

[3] Steven L. Brunton, J. Nathan Kutz (2019) Data-Driven Science and Engineering: Machine Learning, Dynamical Systems, and Control, Cambridge University Press, ISBN: 9781108422093