Identification of non-linear model equations based on data-science approaches

Simon Bäuerle*, Timo Baierl* and Hartmut Hetzler*
* *Institute for Mechanics, Engineering Dynamics, University of Kassel, Germany*

Abstract. This contribution presents two different data-science approaches for identifying non-linear ordinary differential equations from transient system state data. The theory, benefits and pitfalls of a differential approach (called SINDy) and a variational approach are discussed. These are used to identify the DUFFING oscillator equation purely from simulation data.

Introduction

Describing the behavior of dynamical systems with ordinary differential equations $\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = \boldsymbol{f}(\boldsymbol{x}, t)$ is at the core of many disciplines in science. The "classical" approach uses first principles (e.g., momentum balance) to derive a set of governing equations. Data-driven approaches are another way to describe system dynamics: here, a description is achieved purely based on system state data. One possibility is to use neural networks that are trained to replicate the system behavior. If a suitable architecture and hyper-parameters are found, the behavior can be predicted relatively fast. Another way is to identify the (non-linear) model equations directly from the data. This has the advantage that the model equations possibly allow for physical insight into the system and that – in general – less data is needed than for training a neural network. However, some basic understanding of the dynamics present in the system is required. Therefore, this approach tries to combine the two mentioned above. We explore this approach by using two different classes of methods: a differential approach called SINDy (Sparse Identification of Non-linear Dynamics) [1] and a variational approach [2]. Both methods are based on system state data obtained either through experiment or simulation (as done for this contribution). The SINDy approach is described by

$$\dot{\boldsymbol{X}} \overset{!}{=} \underbrace{\begin{bmatrix} \mathbf{1} & \boldsymbol{\Phi}_1(\boldsymbol{X}) & \boldsymbol{\Phi}_2(\boldsymbol{X}) & \dots \end{bmatrix}}_{=\boldsymbol{\Phi}} \underbrace{\begin{bmatrix} \boldsymbol{\xi}_1 & \boldsymbol{\xi}_2 & \boldsymbol{\xi}_3 & \dots \end{bmatrix}}_{=\boldsymbol{\xi}}, \qquad \underset{\boldsymbol{\xi}}{\operatorname{argmin}} \left(\|\boldsymbol{\Phi}\boldsymbol{\xi} - \dot{\boldsymbol{X}}\|_2^2 + \lambda \|\boldsymbol{\xi}\|_1 \right). \tag{1}$$

The matrix $\boldsymbol{X}$ contains the gathered data for each DoF (as column entries). The derivation w.r.t. time $\dot{\boldsymbol{X}}$ is carried out numerically. Then a set of k different possible candidate function terms $\boldsymbol{\Phi}_k$ (e.g., monomials, trigonometric functions, ...) is evaluated subject to $\boldsymbol{X}$. These function terms are weighted with k factors $\boldsymbol{\xi}$ and their sum is required to equal the differentiated data. This represents an over-determined equation system for a first order differential equation. Instead of using a classical least square approach, sparsity of the factors $\boldsymbol{\xi}$ is additionally promoted, since governing equations are mostly sparse in their functional terms. This is achieved by minimizing the Euclidean norm of the residuum *and* the summation norm of the factor vector $\boldsymbol{\xi}$. The variational approach uses a variational-integral formulation

$$\int_{T_1}^{T_2} \left(\delta U_0(\boldsymbol{x}, \dot{\boldsymbol{x}}) + \sum_k U_k(\boldsymbol{x}, \dot{\boldsymbol{x}}, t) \delta x_k \right) \mathrm{d}t = 0. \tag{2}$$

As for SINDy, U_0 and U_k are expressed by k possible candidate function terms $\boldsymbol{\Phi}$, which are evaluated with the gathered data $\boldsymbol{X}$ for arbitrary T_1, T_2 and δx . After (numerical) integration, the coefficients $\boldsymbol{\xi}$ can be calculated.

Results and discussion

We focus on the comparison between the differential and variational methods based on the application to the Duffing oscillator $\ddot{x} + 2D\dot{x} + x + \kappa x^3 = f\cos(\Omega t)$ with an overhanging resonance peak. This is an interesting example since three solutions co-exist. Therefore, we extend the results from [3] and [4] w.r.t. a higher degree of non-linearity. We show that the choice of data is essential for (re-)discovering the governing differential equation: transient data gathered away from attractors led here to better results. Additionally, numerical differentiation as a source of noise has a significant impact. Here, variational methods are less prone to noise, whereby differential methods tend to be more efficient (see also [2]). Finally, we study the influence of numerical parameters and the choice of candidate function terms.

References

[1] Brunton, S. L., Proctor, J. L., Kutz, J. N. (2016) Discovering governing equations from data by sparse identification of nonlinear dynamical systems *Proc. Natl. Acad. Sci. U.S.A.* **113(15)**: 3932–3937.

[2] Huang, Z., Tian, Y., Li, C., Lin, G.,Wu, L.,Wang, Y., Jiang, H. (2020) Data-driven automated discovery of variational laws hidden in physical systems *J. Mech. Phys. Solids* **137**: 103871.

[3] Li, C., Huang, Z., Wang, Y., Jiang, H. (2021). Rapid identification of switched systems: A data-driven method in variational framework. *Sci. China Technol. Sci.* , **64(1)**: 148–156.

[4] Baierl, T. P. (2022) Identification by machine learning of differential equations based on big data (in German). Bachelor Thesis, University of Kassel, Germany.