A whirlwind tour of SINDy

Jean-Christophe LOISEAU, DynFluid, Arts & Métiers - Paris

Identifying reduced-order models from data is a central challenge in mathematical physics, with a rich history of developments in fluid dynamics. The form of the dynamics is typically either constrained via prior knowledge, as in Galerkin projection, or a particular model structure is chosen heuristically and parameters are optimized to match the data. Simultaneous identification of the model structure and parameters from data is considerably more challenging as there are combinatorially many possible model structures. For such tasks, the *Sparse Identification of Nonlinear Dynamics* (SINDy) [1] methodology has gained a lot of traction over the past decade.

SINDy bypasses the intractable combinatorial search through all possible model structures, leveraging the fact that many systems may be modelled by dynamics $\mathbf{f}$ that is sparse in the space of possible right-hand side functions :

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = \mathbf{f}(\mathbf{x}),$$

where $\mathbf{x} \in R^n$ is the state vector of the system. Expanding the unknown function $\mathbf{f}$ in a suitable basis, i.e. $\mathbf{f}(x) = \sum_{i=0}^{m} \varphi_i(\mathbf{x})\alpha_i$, it is then possible to solve for the relevant terms that are active in the dynamics using various convex relaxations of the following optimization problem

$$\begin{aligned} &\text{minimize} && \|\boldsymbol{\alpha}\|_0 \\ &\text{subject to} && \int_0^T \left(\dot{\mathbf{x}} - \sum_{i=0}^{m} \varphi_i(\mathbf{x})\alpha_i \right)^2 \mathrm{d}t \leq \varepsilon, \end{aligned}$$

where $\|\boldsymbol{\alpha}\|_0$ denotes the ℓ_0 pseudo-norm (or cardinality), i.e. the number of non-zero elements in $\boldsymbol{\alpha}$.

In this talk, we will give a whirlwind tour of the SINDy ecosystem. Starting with the identification of simple nonlinear ordinary differential equations, we will illustrate how the basic SINDy methodology can be extended to enforce known symmetries or guaranteed stability, identify dynamics defined by rational functions or stochastic differential equations, and extend it with uncertainty quantification capabilities. We will also discuss the advantages and limitations of the various variants of SINDy as well as illustrate its use on increasingly complicated systems, from simple nonlinear oscillators to reduced-order models of quasi-periodic and chaotic two-dimensional flows. All the examples presented during the talk will be accompanied by `Jupuyter Notebooks` using the `pySINDy` library [3, 2], an open-source python package providing state-of-the-art implementations of the various SINDy flavours.

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

[2] A. A. Kaptanoglu, B. M. de Silva, U. Fasel, K. Kaheman, A. J. Goldschmidt, J. Callaham, C. B. Delahunt, Z. G. Nicolaou, K. Champion, J.-C. Loiseau, J. N. Kutz, S. L. Brunton. *PySINDy : A comprehensive Python package for robust sparse system identification.* Journal of Open Source Software, **7(69)**, 3994, 2022.

[3] B. M. de Silva, K. Champion, M. Quade, J.-C. Loiseau, J. N. Kutz, S. L. Brunton. *PySINDy : A Python package for the sparse identification of nonlinear dynamical systems from data.* Journal of Open Source Software, **5(49)**, 2104, 2020.

Contact : `jean-christophe.loiseau@ensam.eu`