Automated Discovery of Continuous Dynamics

Kuang Huang[a,†,1], Dong Heon Cho[b,†,2], and Boyuan Chen[c,2,3,4]

[†] These authors contributed equally to this work.
[1]Department of Applied Physics and Applied Mathematics, Columbia University, New York, NY, USA
[2]Department of Computer Science, Duke University, Durham, NC, USA
[3]Department of Mechanical Engineering, Duke University, Durham, NC, USA
[4]Department of Electrical and Computer Engineering, Duke University, Durham, NC, USA

November 5, 2024

http://generalroboticslab.com/SmoothNSV

Abstract: Dynamical systems are predominantly described by physical laws, which involve a set of physical variables to represent the system's states and a set of equations to connect these variables to model the system's evolution over time. By identifying the appropriate physical variables and deriving equations from these variables to articulate the underlying physical principles, scientists can apply various mathematical tools to analyze these equations and gain a profound understanding of natural phenomena. This paradigm of scientific discovery, most well-known since the work of Tycho Brahe and Johannes Kepler more than 400 years ago, has been remarkably successful across almost all areas of modern science. However, applying this paradigm to new systems remains challenging as it requires significant effort to uncover the few hidden variables that can fully describe the system from noisy and high-dimensional raw observations. Thanks to advances in computational power and deep learning in recent years, using machines to process large amounts of high-dimensional observational data to aid human scientists in scientific discovery has become a promising direction, as evidenced by the rapidly growing body of literature in this area[1, 2, 3, 4].

[a]kh2862@columbia.edu
[b]dongheon.cho@duke.edu
[c]boyuan.chen@duke.edu

However a feasible framework for integrating the paradigm established among human scientists with highly expressive machine learning models is still lacking. Current data-driven machine learning methods often produce highly complex representations that are not suitable for the various powerful mathematical tools, such as calculus, to be applied for downstream scientific discovery tasks , or require strong prior knowledge for designing the model architecture or the training process , limiting the applicability of these methods to diverse systems. To address this challenge, first-principled constraints need to be applied to the machine learning models so that they can automatically discover state variables that 1) are non-redundant descriptions of the systems with their dimensionality matching the system's intrinsic dimension and 2) form smooth trajectories allowing calculus-based mathematical techniques to be applied for knowledge discovery from the systems.

Our proposed method discovers state variables of minimum dimensionality and their accompanying continuous vector field automatically from video streams without enforcing any physical priors during training. The discovered state variables, termed smooth neural state variables, preserve the smoothness of the system dynamics and are analogous to the position and velocity of a spring mass, and the discovered vector field, termed the neural state vector field, embeds the system's complete dynamic information and is analogous to Hooke's Law describing the motion of a spring mass. The prominence and effectiveness of the approach are demonstrated through quantitative and qualitative analyses of various dynamical systems. Without any prior physical knowledge, our method can identify each system's stable equilibrium, predict the natural frequencies of the spring-mass and single pendulum systems, identify chaotic behaviors in the motion of the double pendulum, and discover limit cycles when turbulence appears in a fluid flowing past a cylinder body. Moreover, our method allows for synthesizing physically plausible videos governed by physical laws that deviate from those in the currently observed systems. We believe that our approach opens up new opportunities for automated scientific discovery systems as a prominent and effective assistant to human scientists.

References

[1] Kathleen Champion, Bethany Lusch, J Nathan Kutz, and Steven L Brunton. Data-driven discovery of coordinates and governing equations. *Proceedings of the National Academy of Sciences*, 116(45):22445–22451, 2019.

[2] Boyuan Chen, Kuang Huang, Sunand Raghupathi, Ishaan Chandratreya, Qiang Du, and Hod Lipson. Automated discovery of fundamental variables hidden in experimental data. *Nature Computational Science*, 2(7):433–442, July 2022.

[3] Bethany Lusch, J Nathan Kutz, and Steven L Brunton. Deep learning for universal linear embeddings of nonlinear dynamics. *Nature communications*, 9(1):4950, 2018.

[4] Samuel A. Moore, Brian P. Mann, and Boyuan Chen. Automated global analysis of experimental dynamics through low-dimensional linear embeddings, 2024.