Discovering governing equations from data by sparse identification of nonlinear dynamical systems

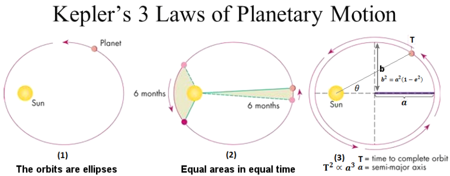
Steven L. Bruntona, Joshua L. Proctor , and J. Nathan Kutz

[1]Anurendra Kumar
Computer Science, UIUC

CS 598 DGDM, Class Presentation

A sneak peek into history: Modeling dynamics from data

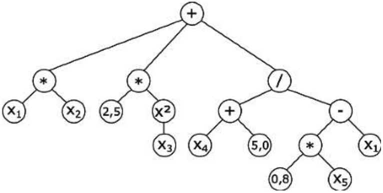
- **Copernicus**: Heliocentric theory.
- **Kepler's** three laws driven by data.
- Newton - Unified theory of motion across the universe.



Symbolic Regression: Modeling dynamics from data

- Generalization of regression.
- Search over the space of all possible mathematical formulas best predict the output variable, starting from a set of base functions like addition, trigonometric functions, and exponentials.
- **Problems:** Computationally expensive, does not clearly scale well to large-scale dynamical systems of interest, and may be prone to over-fitting.

$$f(x) = x_1 x_2 + 2.5\, x_3^2 + \frac{x_4 + 5.0}{0.8\, x_5 - x_1}$$



Background: Sparse Regression and Compressive Sensing

- Finding solutions to underdetermined linear systems. Sparsity helps us break the Nyquist–Shannon sampling theorem.
- $$x = \Psi s, \quad y = \Phi x = \Phi \Psi s$$
- x is K sparse.
- $l1$ norm regularization.
- Extensive research on Compressive Sensing. No need to perform a combinatorially intractable bruteforce search. Sparse solution is found with high probability using convex methods that scale to large problems. Terrence Tao.

[1] Baraniuk, R. G. (2007). Compressive sensing [lecture notes]. IEEE signal processing magazine, 24(4), 118-121.

Problem Setup: Dynamical Systems

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

- Change in notation.
- Vector $\mathbf{x}(t)$ denotes state time t, f denotes dynamics.

Dynamical Systems as sparse regression

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}^T(t_1) \\ \mathbf{x}^T(t_2) \\ \vdots \\ \mathbf{x}^T(t_m) \end{bmatrix} = \overset{\text{state}\ \longrightarrow}{\begin{bmatrix} x_1(t_1) & x_2(t_1) & \cdots & x_n(t_1) \\ x_1(t_2) & x_2(t_2) & \cdots & x_n(t_2) \\ \vdots & \vdots & \ddots & \vdots \\ x_1(t_m) & x_2(t_m) & \cdots & x_n(t_m) \end{bmatrix}} \downarrow \text{time}$$

$$\dot{\mathbf{X}} = \begin{bmatrix} \dot{\mathbf{x}}^T(t_1) \\ \dot{\mathbf{x}}^T(t_2) \\ \vdots \\ \dot{\mathbf{x}}^T(t_m) \end{bmatrix} = \begin{bmatrix} \dot{x}_1(t_1) & \dot{x}_2(t_1) & \cdots & \dot{x}_n(t_1) \\ \dot{x}_1(t_2) & \dot{x}_2(t_2) & \cdots & \dot{x}_n(t_2) \\ \vdots & \vdots & \ddots & \vdots \\ \dot{x}_1(t_m) & \dot{x}_2(t_m) & \cdots & \dot{x}_n(t_m) \end{bmatrix}.$$

Overrepresented Dictionary

$$\dot{\mathbf{X}} = \Theta W$$

- W has sparse coefficients.
- Each row has separate optimization.

$$\Theta(\mathbf{X}) = \begin{bmatrix} | & | & | & | & & | & | & \\ 1 & \mathbf{X} & \mathbf{X}^{P_2} & \mathbf{X}^{P_3} & \cdots & \sin(\mathbf{X}) & \cos(\mathbf{X}) & \cdots \\ | & | & | & | & & | & | & \end{bmatrix}.$$

$$\mathbf{X}^{P_2} = \begin{bmatrix} x_1^2(t_1) & x_1(t_1)x_2(t_1) & \cdots & x_2^2(t_1) & \cdots & x_n^2(t_1) \\ x_1^2(t_2) & x_1(t_2)x_2(t_2) & \cdots & x_2^2(t_2) & \cdots & x_n^2(t_2) \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ x_1^2(t_m) & x_1(t_m)x_2(t_m) & \cdots & x_2^2(t_m) & \cdots & x_n^2(t_m) \end{bmatrix}.$$

Learnt Dynamical Systmes

$$\dot{\mathbf{x}} = f(\mathbf{x}) = W^T \Theta(\mathbf{x}^T)^T$$

- Basis functions important.
- Test many different function bases and use the sparsity and accuracy of the resulting model as a diagnostic tool to determine the correct basis to represent the dynamics.

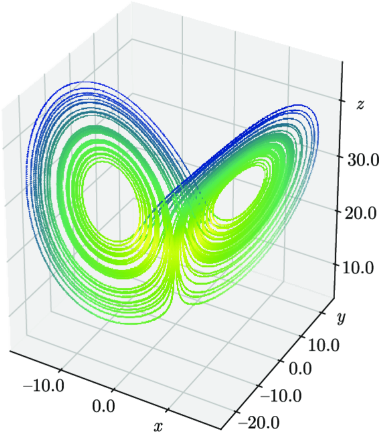
Approximating derivative

- $\dot{\mathbf{X}}$ is approximated and not known.
-
$$\dot{\mathbf{X}} = \Theta W + \sigma Z$$
- Z is iid normal distributed.
- Depending on the noise, it may be necessary to filter $\mathbf{X}$ and $\dot{\mathbf{X}}$. Total variation regularization to denoise derivative.

Dynamical Systems with PDE

- Numerical discretization on a spatial grid is exponentially large.
- Fluid Dynamics - Simple 2D and 3D flows may require tens of thousands up to billions of variables to represent the discretized system.
- Current formulation ill suited, since each row has separate optimization.
- Good news- Many high-dimensional systems of interest evolve on a low dimensional manifold or attractor that is well-approximated using a low-rank basis.

- Rich and chaotic dynamics that evolve on an attractor.
- Only a few terms in the right-hand side of the equation

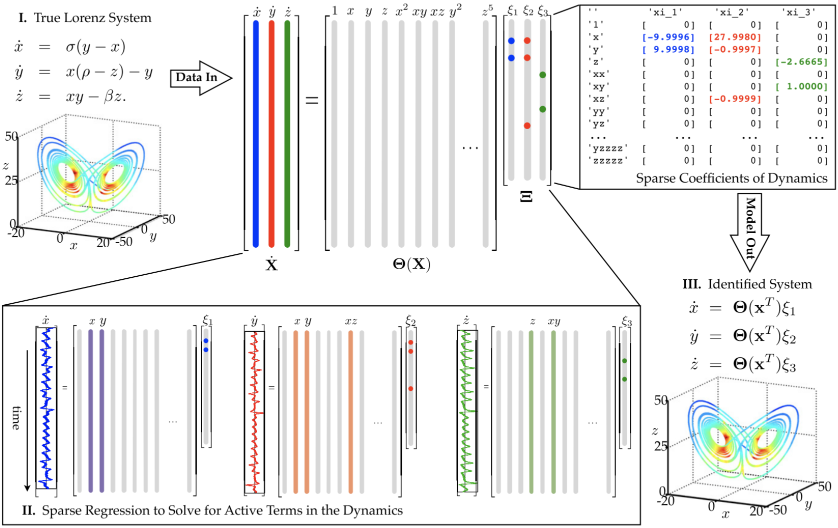


$$\dot{x} = \sigma(y - x),$$

$$\dot{y} = x(\rho - z) - y,$$

$$\dot{z} = xy - \beta z.$$

Results: Lorenz System



Results: Fluid Dynamics

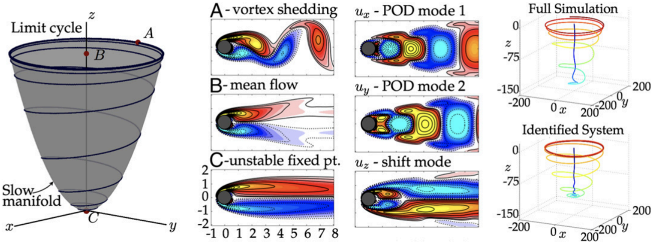
- Data are collected for the fluid flow past a cylinder using direct numerical simulations of the 2D Navier–Stokes equations.
- Hopf bifurcations, Reynolds number
- 15 years of research into mean field model - low rank basis.

$$\dot{x} = \mu x - \omega y + Axz,$$

$$\dot{y} = \omega x + \mu y + Ayz,$$

$$\dot{z} = -\lambda\left(z - x^2 - y^2\right).$$

Results: Lorenz System



Extensions

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}; \mu),$$

$$\dot{\mu} = 0.$$

Reference

- Brunton, S. L., Proctor, J. L., Kutz, J. N. (2016). Discovering governing equations from data by sparse identification of nonlinear dynamical systems. Proceedings of the national academy of sciences, 113(15), 3932-3937.