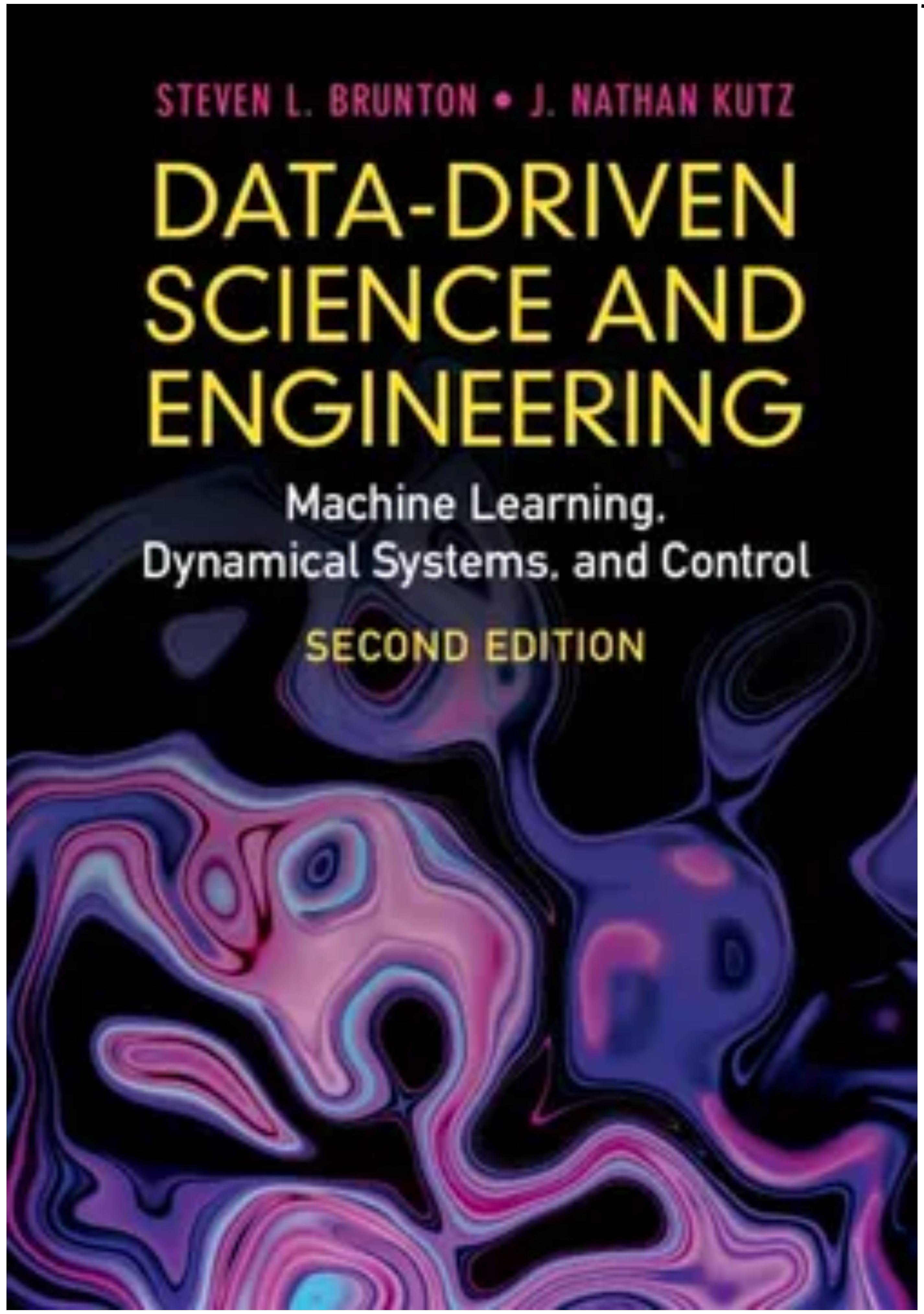
Introduction to Scientific ML : Day 2

By Juan Felipe Osorio Ramirez

MindLab
Department of Computer Systems and Industrial Engineering, UN
Summer 2023

UNIVERSITY *of* WASHINGTON

Data-driven science and engineering



Dimensionality Reduction and Transforms

- Singular value decomposition
- Fourier and Wavelet Transform
- Sparsity and Compressed Sensing
- …

Machine learning and Data Analysis

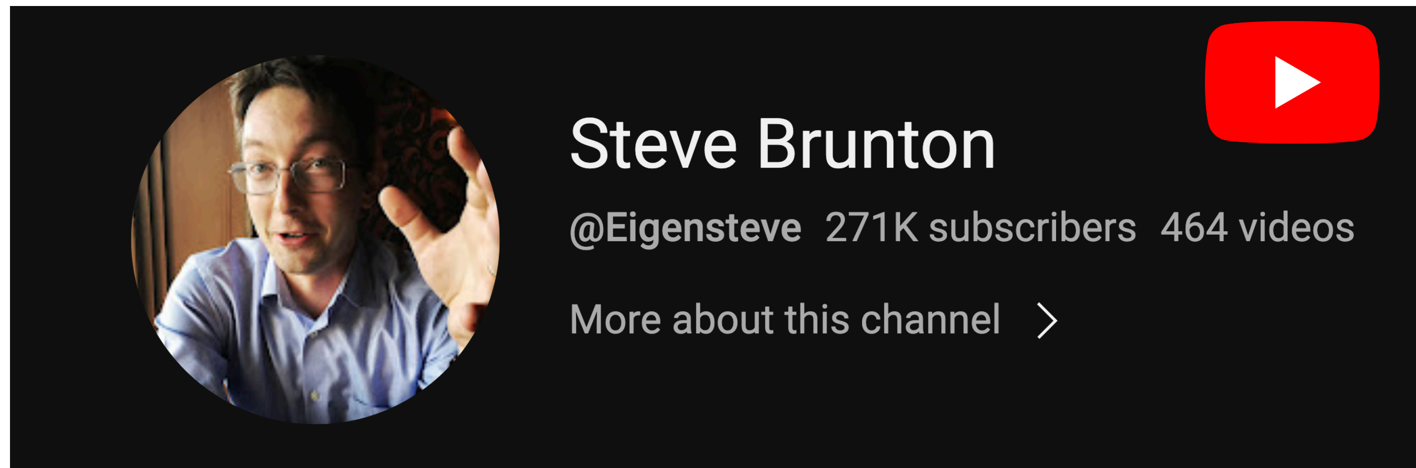
Dynamics and Control

- Data-driven Dynamical Systems
- Linead Control Theory
- ….

Advanced Dara-Driven Modeling and Control

- Data-driven Control
- Reinforcement learning
- Physics informed Machine Learning

Videos for every section !!!



Recently published

+ https://github.com/dynamicslab/databook_python

What is a dynamical system?

We will agree to represent it as a **Nonlinear Ordinary Differential Equation**

$$\frac{d}{dt}\mathbf{x}(t) = \mathbf{f}(\mathbf{x}(t), t; \boldsymbol{\beta})$$

The vector $\mathbf{x} \in \mathbb{R}^n$ is the state of the system evolving in time t .

β are parameters.

$\mathbf{f}$ is the vector field. Generally assumed to be Lipschitz continuous to guarantee existence and uniqueness.

Picard-Lindelöf theorem

Katok, Anatole, and Boris Hasselblatt. *Introduction to the modern theory of dynamical systems*. No. 54. Cambridge university press, 1995.

Applications

Linear algebra

Numerical analysis

Geometry

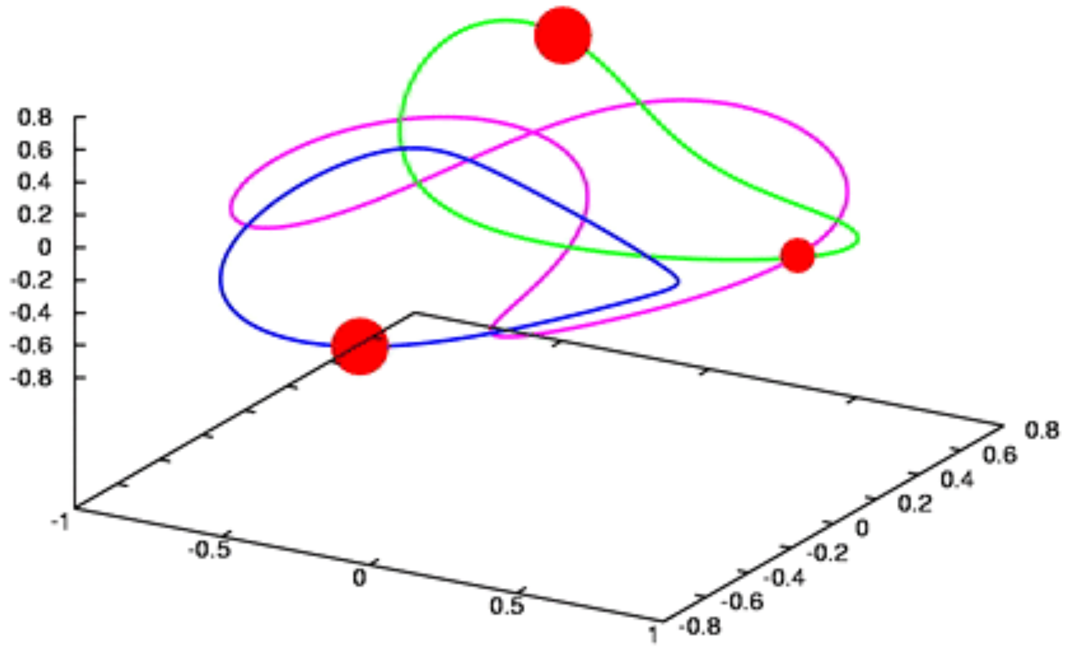
Topology

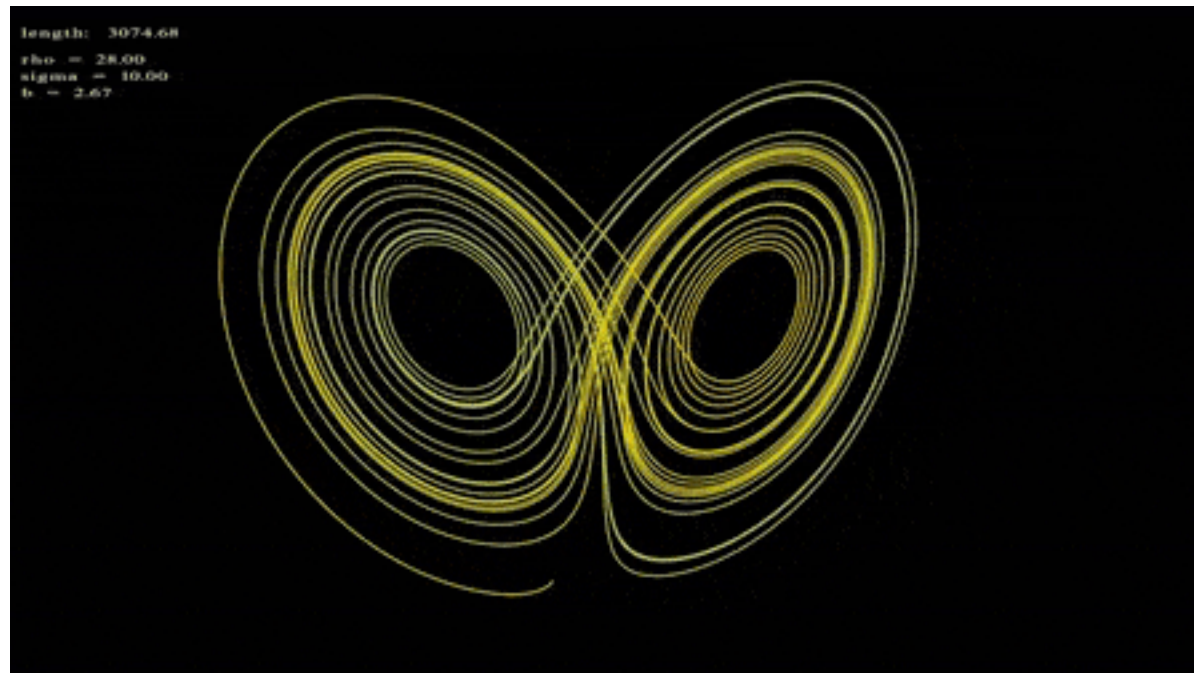
Dynamical systems provide a **mathematical framework** to model interactions between **quantities** that **co-evolve in time**

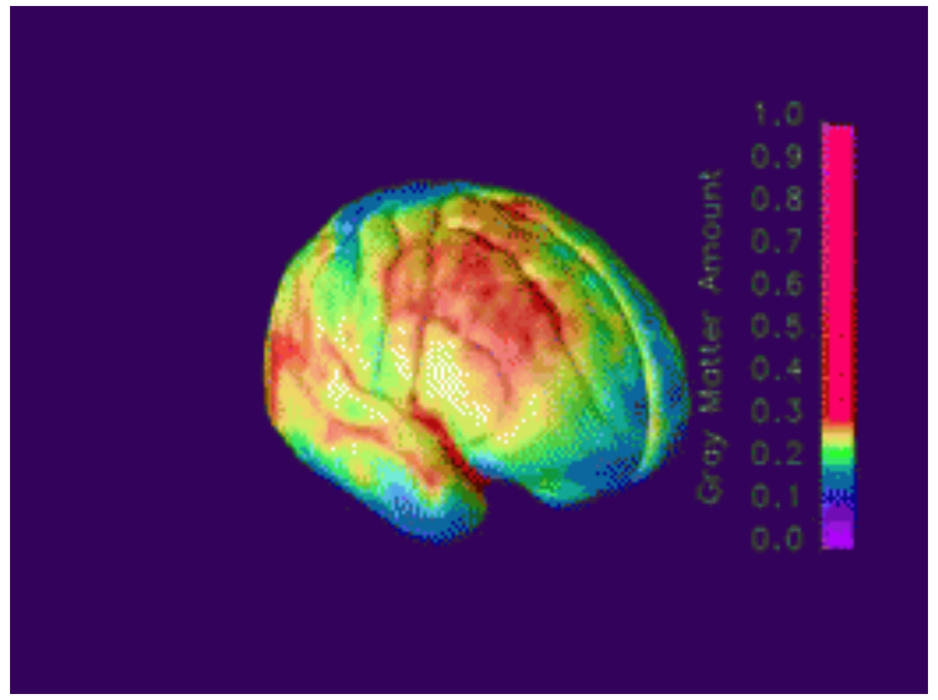
Classical mechanics   Electrical circuits   Turbulent fluids   Climate science   Finance

Ecology   Social systems   Neuroscience   Epidemiology

Seminal work by Poincaré on the 3-body problem







Brunton, Steven L., and J. Nathan Kutz. *Data-driven science and engineering: Machine learning, dynamical systems, and control*. Cambridge University Press, 2022.

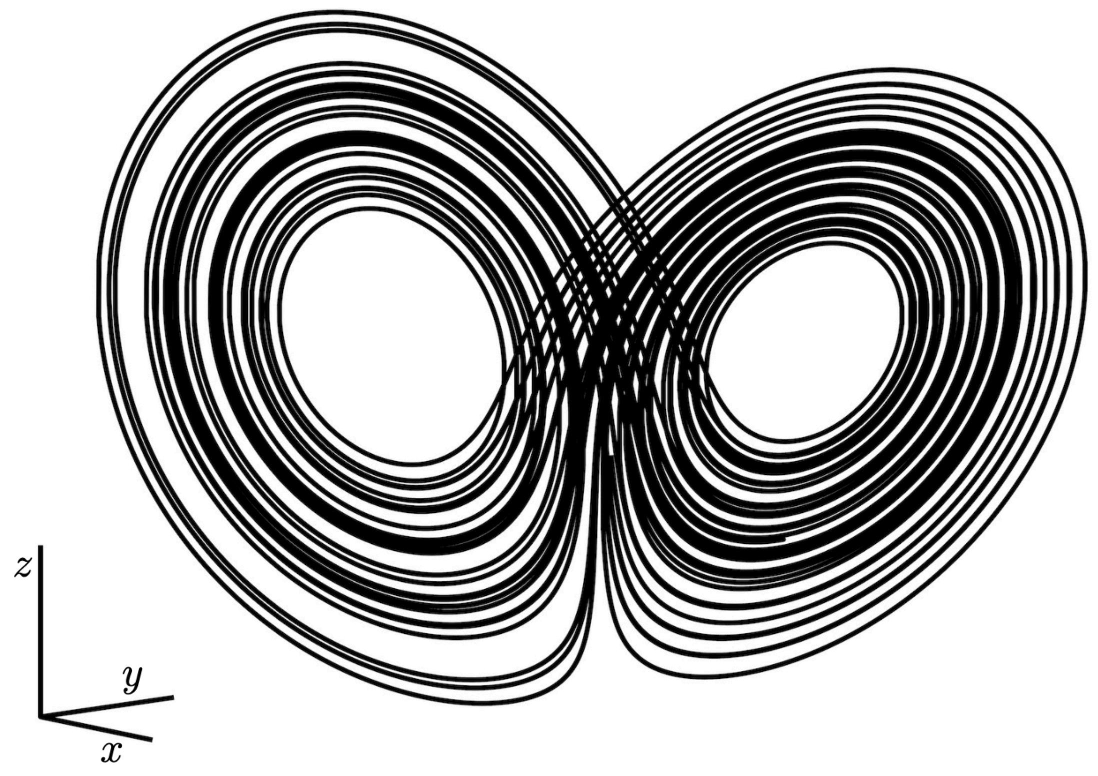
$$\frac{d}{dt}\mathbf{x}(t) = \mathbf{f}(\mathbf{x}(t), t; \boldsymbol{\beta})$$

Example: Lorenz equations

Consider the equations

$$\dot{x} = \sigma(y - x)$$
$$\dot{y} = x(\rho - z) - y$$
$$\dot{z} = xy - \beta z$$

With parameters $\sigma = 10, \quad \rho = 28$, and $\beta = 8/3.$ A trajectory of the Lorenz system



Simple system that exhibit chaos (sensitive dependence of initial conditions).

Edward N. Lorentz. Deterministic nonperiodic flow. Journal of the Atmospheric Sciences, 1963

Goals and Challenges in Modern Dynamical Systems

Future state prediction

Long-time predictions

Design and optimization

Tune parameters of a system

Estimation and control

Control a DS through feedback

Interpretability and physical understanding

Provide insight trough analyzing trajectories

More on CH. 7

Brunton, Steven L., and J. Nathan Kutz. *Data-driven science and engineering: Machine learning, dynamical systems, and control*. Cambridge University Press, 2022.

Open source libraries

For data-driven dynamical systems

PyDMD

Dynamic Mode Decomposition for a data-driven model simplification based on spatiotemporal coherent structures

PyKoopman

Data-driven approximations to the Koopman operator.

PySINDy

TODAY

Sparse regression package with several implementations for the Sparse Identification of Nonlinear Dynamical systems

Data-driven dynamical systems toolbox

Implementation of the algorithms used in several papers. Includes optimizations in C++

Deeptime

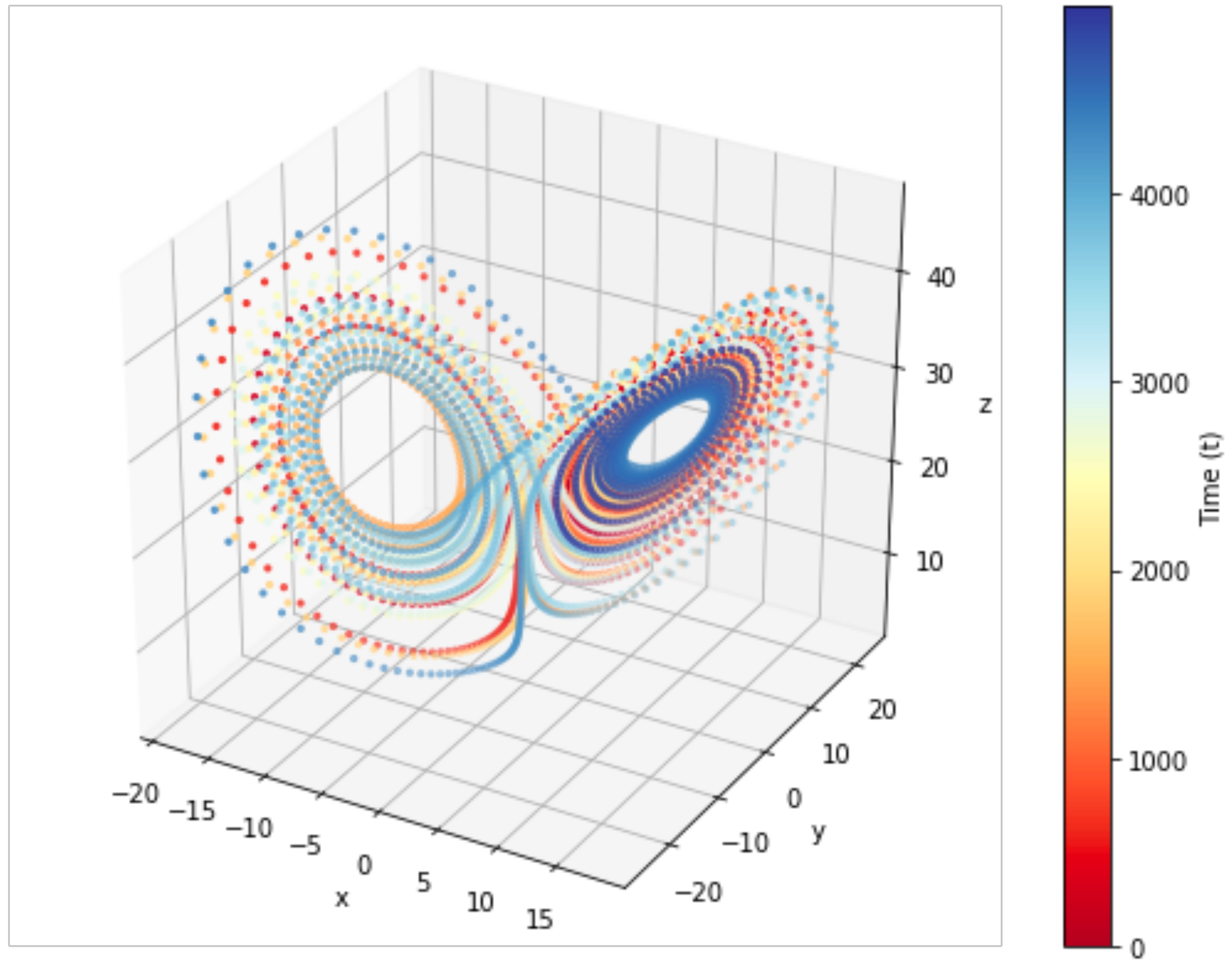
Estimate dynamical models based on time-series data including conventional linear learning methods, such as Markov State Models (MSMs), Hidden Markov Models (HMMs) and Koopman models, as well as kernel and deep learning approaches such as VAMPnets and deep MSMs

Brunton, Steven L., and J. Nathan Kutz. *Data-driven science and engineering: Machine learning, dynamical systems, and control*. Cambridge University Press, 2022.

Sparse identification of Nonlinear Dynamical Systems

Brunton, Proctor, Kutz (2016) PNAS

Data

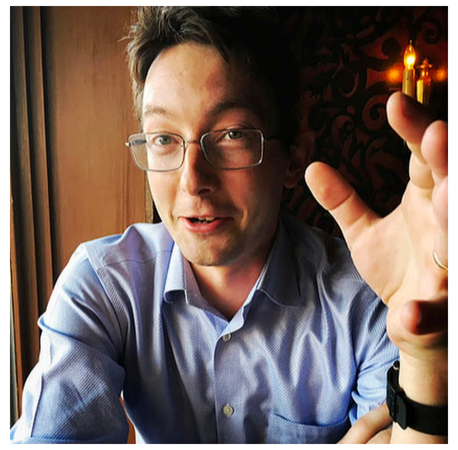


SINDy $\longrightarrow$

Dynamics (ODE)

$$\dot{x} = \sigma(y - x)$$
$$\dot{y} = x(\rho - z) - y$$
$$\dot{z} = xy - \beta z$$

Goal: We wish to identify the governing equations that generated the observed data.


Nathan Kutz, Steve Brunton and Joshua Proctor (from left to right).

Main ideas

Suppose we have a set of measurements $\mathbf{x}(t) \in \mathbb{R}^n$ from some physical system at different points in time t.

SINDy seeks to represent the time evolution of $\mathbf{x}(t)$ in terms of a nonlinear function $\mathbf{f}$:

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

How the system evolves in time.

This equation constitutes a dynamical system for the measurements $\mathbf{x}(t)$.

The vector $\mathbf{x}(t) = [x_1(t), x_2(t), \ldots x_n(t)]^\top$ gives the state of the physical system at time t .

Main ideas

The key idea behind SINDy is that the function f is often **sparse** in the space of an appropriate set of basis functions. For example, the function

$$\frac{d}{dt}\mathbf{x} = \mathbf{x}(\mathbf{x}) = \begin{bmatrix} f_1(\mathbf{x}) \\ f_2(\mathbf{x}) \end{bmatrix} = \begin{bmatrix} 1 - x_1 + 3x_1x_2 \\ x_2^2 - 5x_1^3 \end{bmatrix}$$

is sparse with respect to the set of polynomials of two variables in the sense that if we were to write an expansion of the component functions of f in this basis, e.g.

$$f_1(x) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} a_{i,j} x_1^i x_2^j$$

only a small number of coefficients $a_{i,j}$ would be nonzero.

SINDy employs **sparse regression** to find a linear combination of basis functions that best capture the dynamic behavior of the physical system.

Main ideas

The key idea behind SINDy is that the function f is often **sparse** in the space of an appropriate set of basis functions.

SINDy seeks to approximate f be a generalized linear model

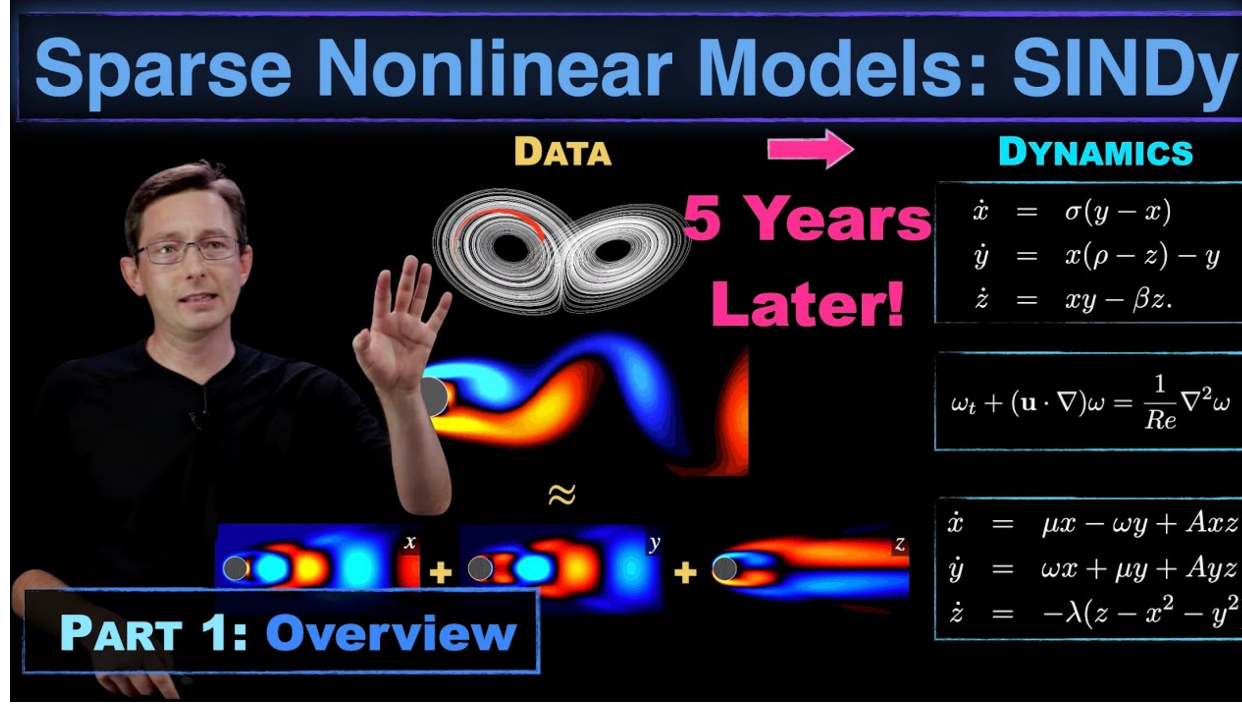
$$\mathbf{f}(\mathbf{x}) \approx \sum_{k=1}^{p} \theta_k(\mathbf{x})\xi_k = \boldsymbol{\Theta}(\mathbf{x})\boldsymbol{\xi}$$

With the fewest non-zero terms in $\boldsymbol{\xi}$ as possible.

Could you name a sparse regression algorithm ?
What type of penalization on the vector of parameters do we choose ?

SINDy employs **sparse regression** to find a linear combination of basis functions that best capture the dynamic behavior of the physical system.

Method



1. **Input measurements**. Time-series data is collected

$$\mathbf{X} = \begin{pmatrix} \mathbf{x}(t_1) & \mathbf{x}(t_2) & \cdots & \mathbf{x}(t_m) \end{pmatrix}^T$$

2. **Differentiation.** Construct a similar matrix of derivatives

Be careful with noisy data

$$\dot{\mathbf{X}} = \begin{pmatrix} \dot{\mathbf{x}}(t_1) & \dot{\mathbf{x}}(t_2) & \cdots & \dot{\mathbf{x}}(t_m) \end{pmatrix}^T$$

3. **Feature library.** Construct a library of candidate nonlinear functions from $\mathbf{X}$

All possible d-th degree polynomials

$$\boldsymbol{\Theta}(\mathbf{X}) = \begin{pmatrix} 1 & \mathbf{X} & \mathbf{X}^2 & \cdots & \mathbf{X}^d & \cdots & \sin(\mathbf{X}) & \cdots \end{pmatrix}$$

Domain specific knowledge

Method

A **parsimonious model** will provide an accurate model fit in

$$\dot{\mathbf{X}} = \mathbf{\Theta}(\mathbf{X})\mathbf{\Xi}$$

with as few terms as possible in $\mathbf{\Xi}$.

Identify the model via convex ℓ_1-regularized sparse regression:

$$\boldsymbol{\xi}_k = \underset{\boldsymbol{\xi}'_k}{\operatorname{argmin}} \left\| \dot{\mathbf{X}}_k - \mathbf{\Theta}(\mathbf{X})\boldsymbol{\xi}'_k \right\|_2 + \lambda \left\| \boldsymbol{\xi}'_k \right\|_1$$

- LASSO
- Sequential thresholded least-squares

k-th column of $\dot{\mathbf{X}}$

sparsity-promoting knob

See section 7.3 in Brunton, Steven L., and J. Nathan Kutz. Data-driven science and engineering: Machine learning, dynamical systems, and control. Cambridge University Press, 2022.

Method

A **parsimonious model** will provide an accurate model fit in

data matrix

$$\dot{\mathbf{X}} = \boldsymbol{\Theta}(\mathbf{X})\boldsymbol{\Xi}$$

with as few terms as possible in $\boldsymbol{\Xi}$.

The sparse vectors $\boldsymbol{\xi}_k$ may be synthesized into a dynamical system

$$\dot{x}_k = \boldsymbol{\Theta}(\mathbf{x})\boldsymbol{\xi}_k$$

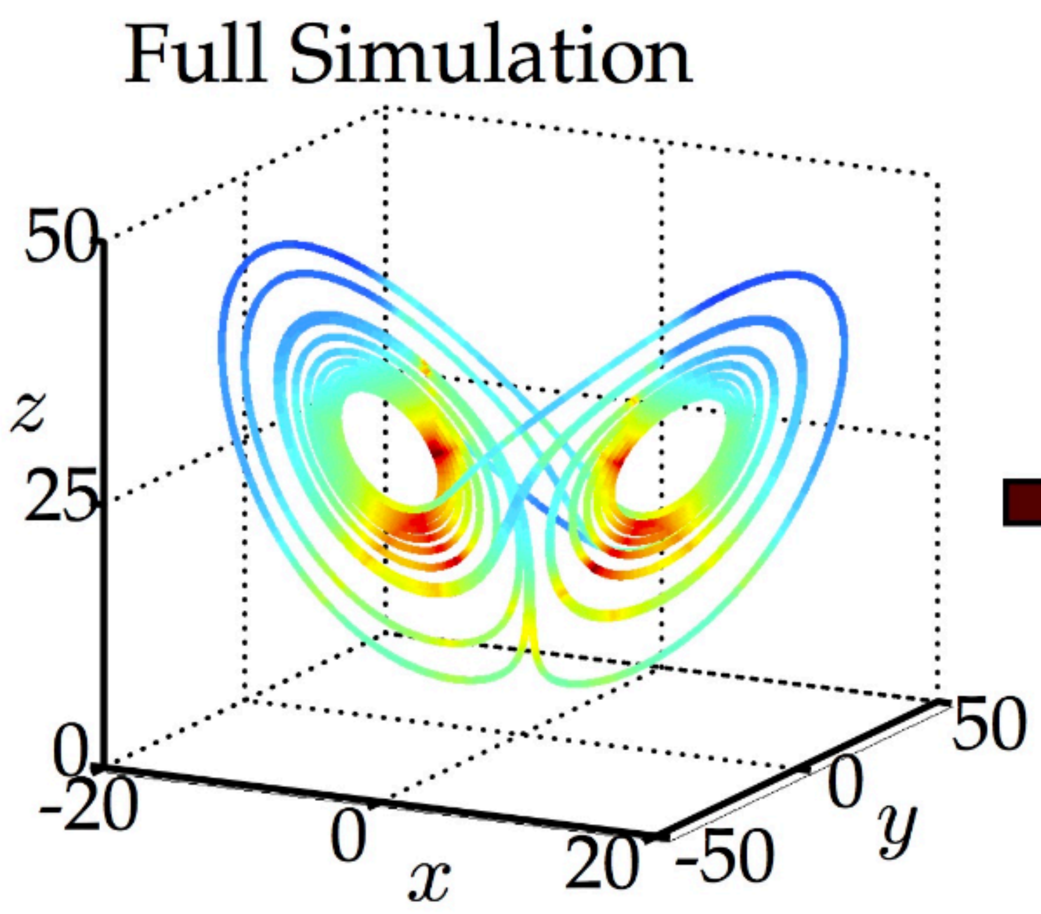
k-th element of $\dot{\mathbf{x}}$

row vector of symbolic functions of $\dot{\mathbf{x}}$

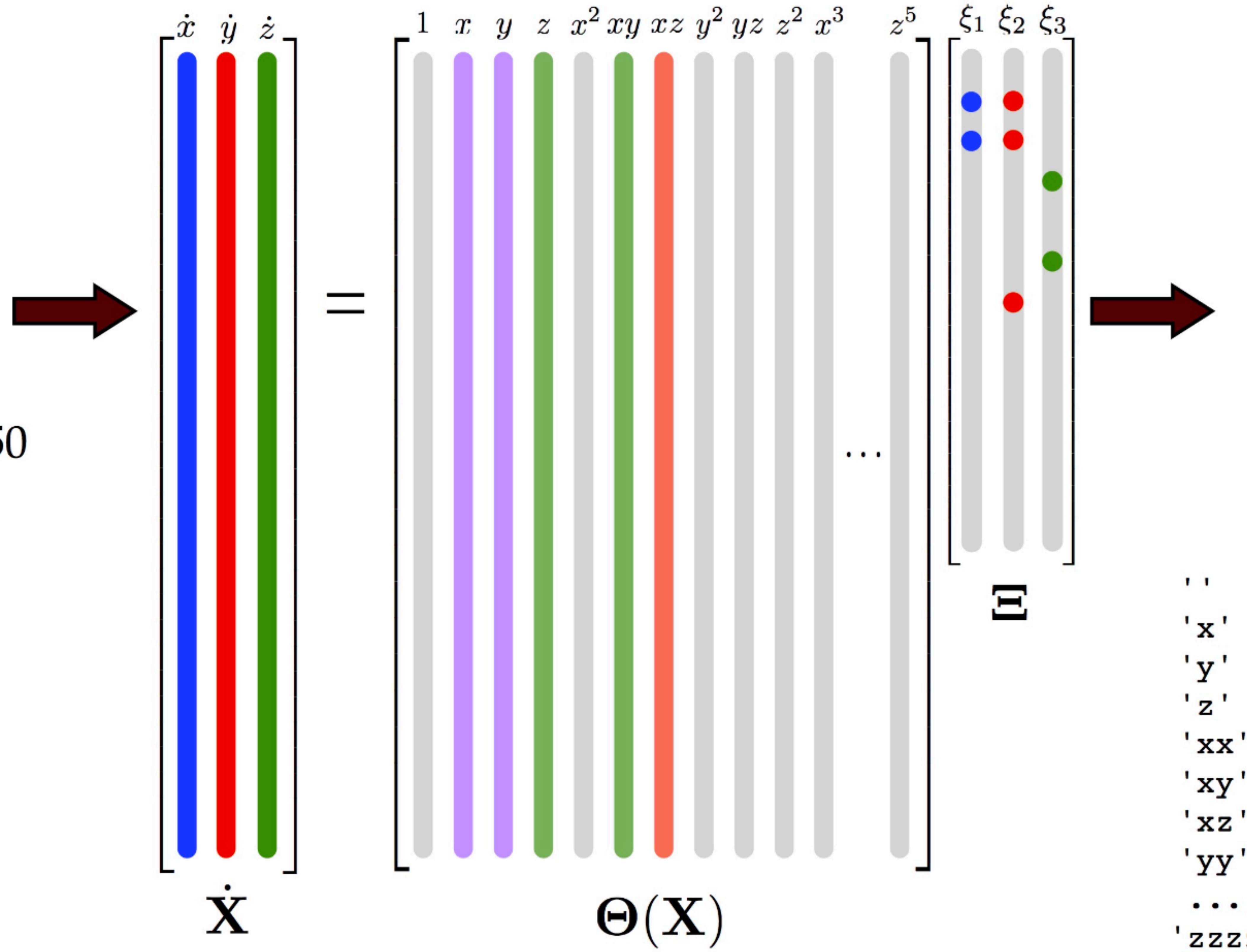
See section 7.3 in Brunton, Steven L., and J. Nathan Kutz. Data-driven science and engineering: Machine learning, dynamical systems, and control. Cambridge University Press, 2022.

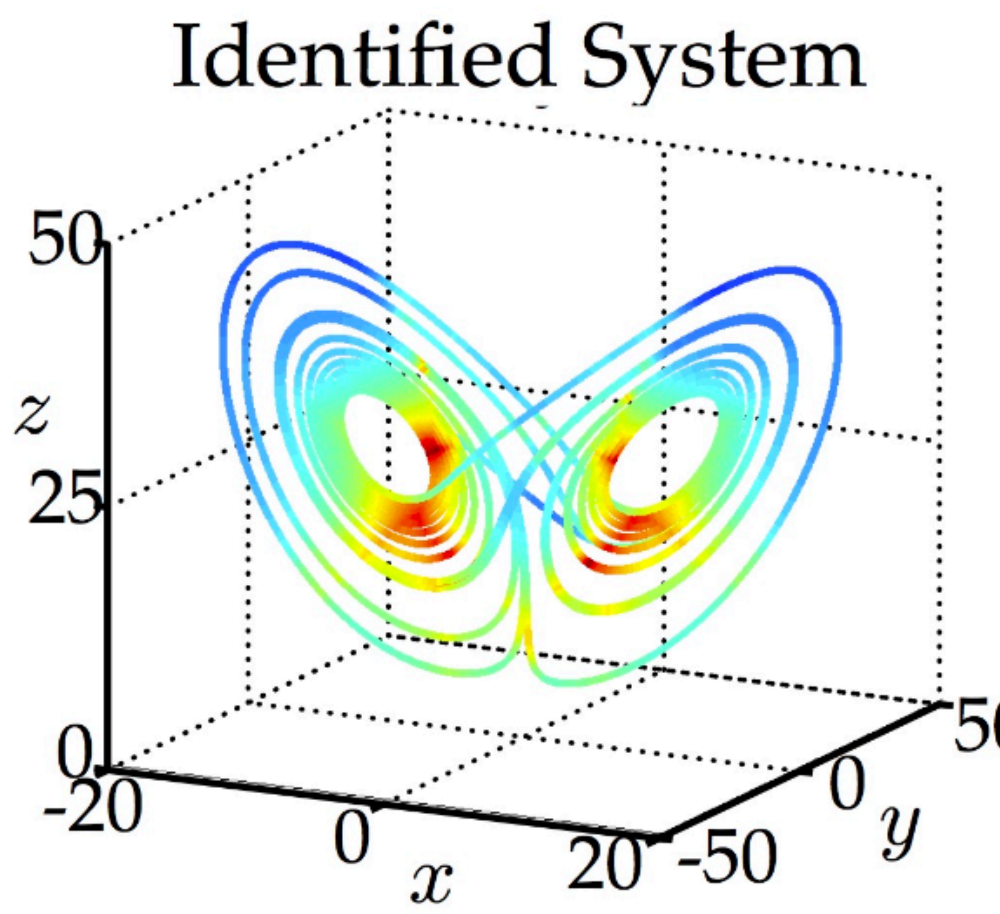
Schematic overall view

Parsimonious model is selected



$$
\begin{aligned}
\dot{x} &= \sigma(y - x) \\
\dot{y} &= x(\rho - z) - y \\
\dot{z} &= xy - \beta z.
\end{aligned}
$$





''	'xi_1'	'xi_2'	'xi_3'
'x'	[-9.9996]	[27.9980]	[0]
'y'	[9.9998]	[-0.9997]	[0]
'z'	[0]	[0]	[-2.6665]
'xx'	[0]	[0]	[0]
'xy'	[0]	[0]	[1.0000]
'xz'	[0]	[-0.9999]	[0]
'yy'	[0]	[0]	[0]
...	...	...	...
'zzzzz'	[0]	[0]	[0]

See section 7.3 in Brunton, Steven L., and J. Nathan Kutz. Data-driven science and engineering: Machine learning, dynamical systems, and control. Cambridge University Press, 2022.

Let's run the demo

See section 7.3 in Brunton, Steven L., and J. Nathan Kutz. Data-driven science and engineering: Machine learning, dynamical systems, and control. Cambridge University Press, 2022.