System Identification of Partial Differential Equations

Anthony Corso[1] – acorso@stanford.edu

[1] SISL Lab, Aeronautics and Astronautics, Stanford University

Stanford
CS 229

Objective

The goal of this project was to implement a system that takes in observed data and outputs a partial differential equation that describes the data. The system should

- Report results in terms interpretable by a human
- Be robust to noisy data
- Operate on small amounts of data

Data

Synthetic noisy data was computed for two model processes from their exact solutions, with varying resolution and amount of noise

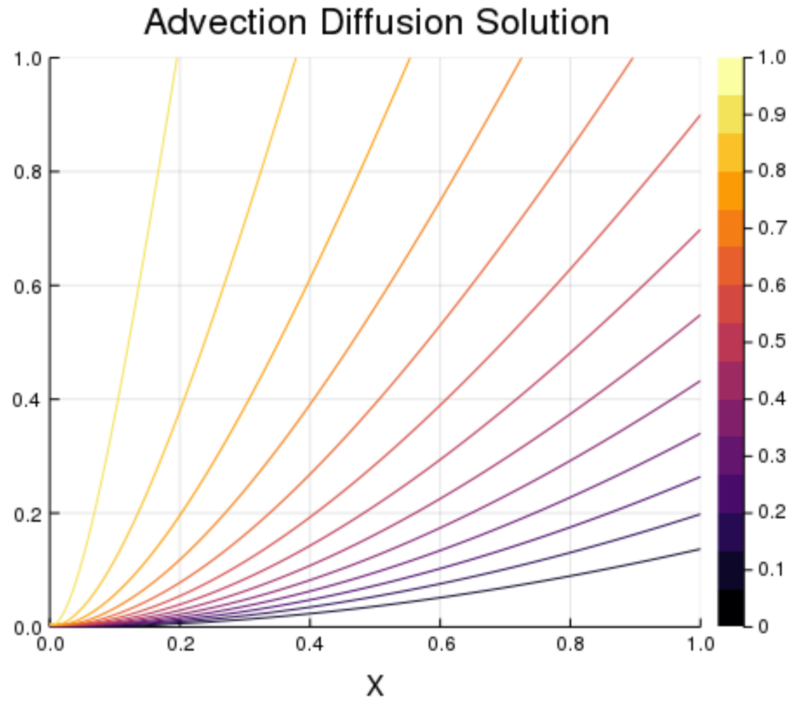
Advection-Diffusion Equation:

$$\frac{\partial u}{\partial t} = D\,\frac{\partial^2 u}{\partial x^2} - v\,\frac{\partial u}{\partial x}$$

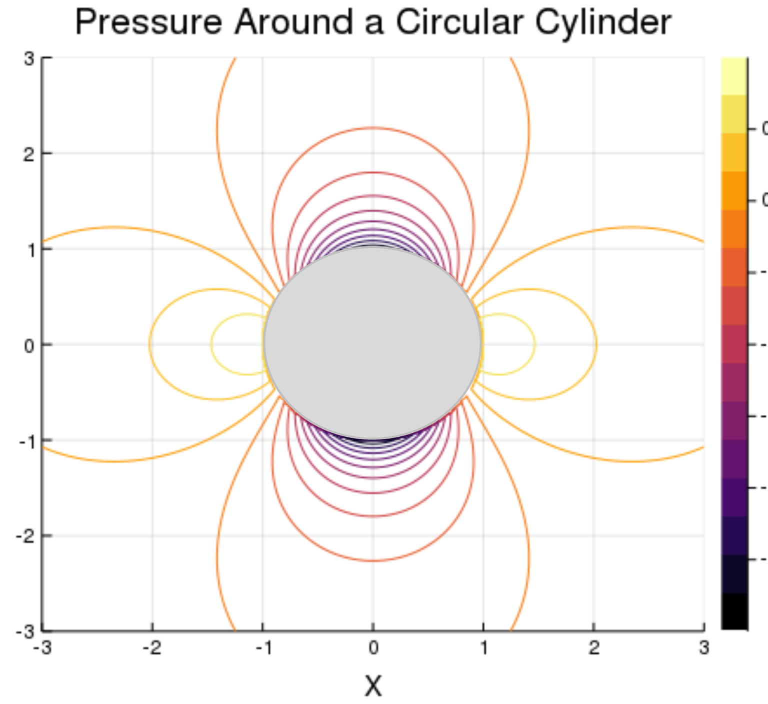
2D Euler Equations:

$$u_j\,\frac{\partial u_i}{\partial x_j} = -\frac{1}{\rho}\frac{\partial p}{\partial x_i}$$

Advection Diffusion Solution



Pressure Around a Circular Cylinder



Adding Noise

To add noise at a level η, add AGWN at each point

$$u_n = u + \epsilon \qquad \epsilon \sim N(0, \eta\,\mathrm{std}(u))$$

Noise was filtered using total variation denoising

Model and Feature Selection

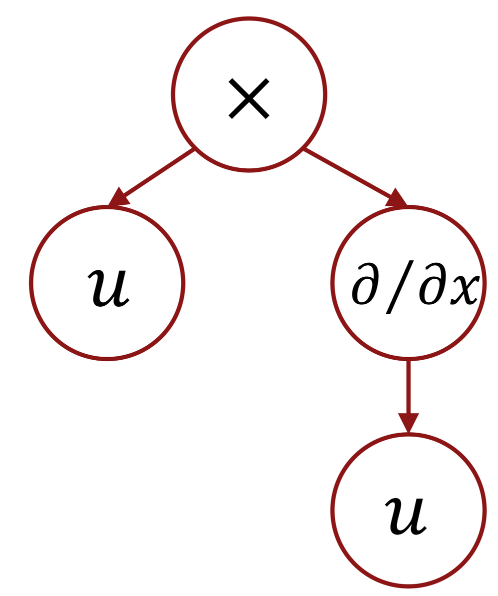
Every PDE can be represented as a linear combination of nonlinear features: $\partial u/\partial t = \theta_i\, f_i(u)$

The candidate features are constructed using a domain specific grammar. Expressions are sampled with a preference for concise statements

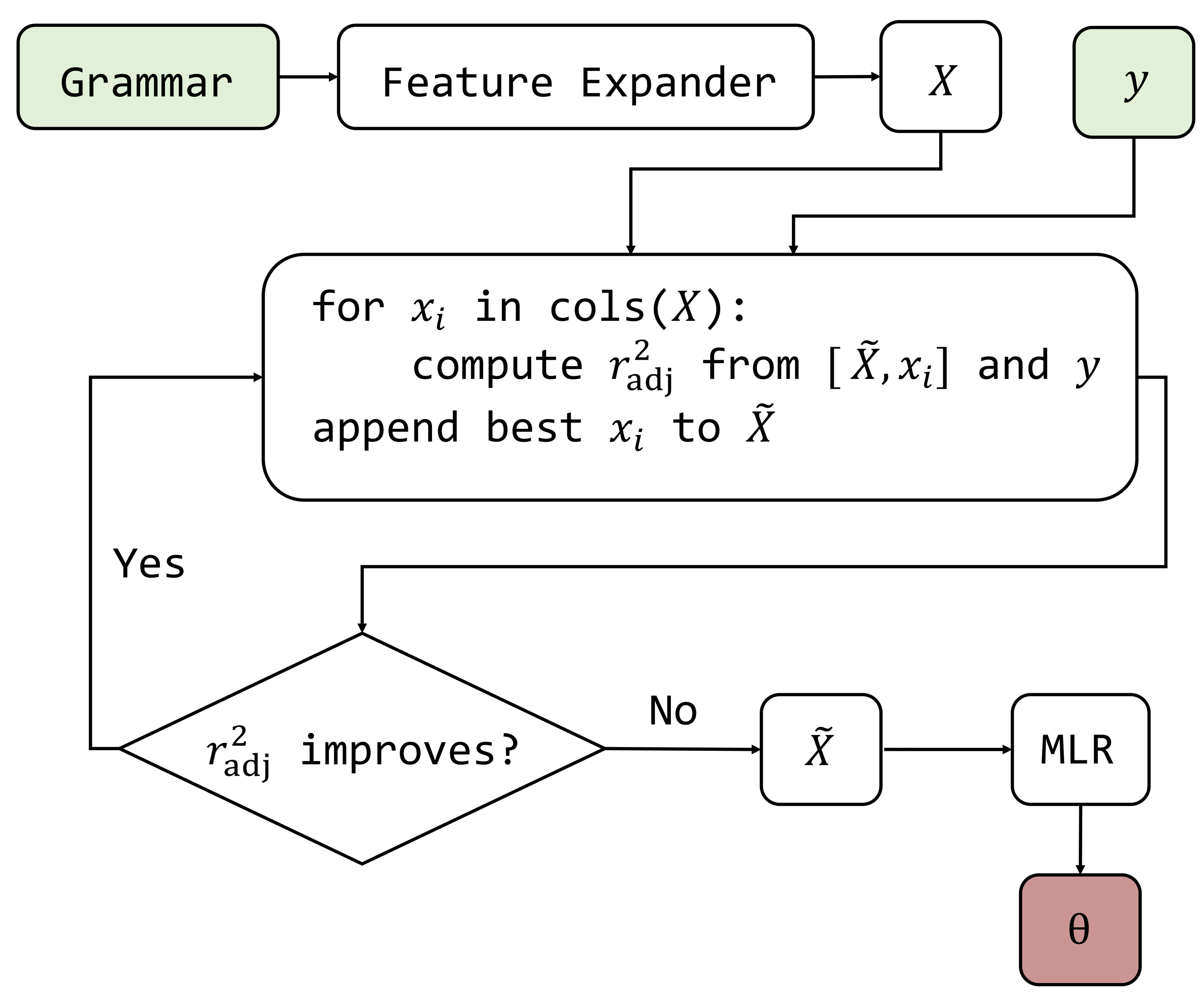
Grammar:

$$\mathbb{R} \mapsto \partial\mathbb{R}/\partial x_i$$
$$\mathbb{R} \mapsto \mathbb{R}\times\mathbb{R}$$
$$\mathbb{R} \mapsto \mathbb{R}/\mathbb{R}$$
$$\mathbb{R} \mapsto \sqrt{|\mathbb{R}|}$$
$$\mathbb{R} \mapsto u_i \mid p$$

Sample Expression: $u \times \partial u/\partial x$



Feature Search Algorithm



Results

Points (1D)	100	50	15	5
Adv-Diff	$D_{err} = 0.03\%$, $v_{err} = 0.03\%$	$D_{err} = 0.08\%$, $v_{err} = 0.05\%$	$D_{err} = 0.7\%$, $v_{err} = 0.3\%$	$D_{err} = 7\%$, $v_{err} = 3\%$
Euler	$\rho_{err} = 0.001\%$, $R_{err} = 0.002\%$	$\rho_{err} = 0.005\%$, $R_{err} = 0.006\%$	$\rho_{err} = 0.06\%$, $R_{err} = 0.06\%$	$\rho_{err} = 0.6\%$, $R_{err} = 0.6\%$
Noise	**1%**	**5%**	**15%**	**50%**
Adv-Diff	$D_{err} = 13\%$, $v_{err} = 7\%$	$D_{err} = 14\%$, $v_{err} = 7\%$	$D_{err} = 24\%$, $v_{err} = 5\%$	$D_{err} = 70\%$, $v_{err} = .5\%$
Euler	$\rho_{err} = 0.3\%$, $R_{err} = 2\%$	$\rho_{err} = 0.1\%$, $R_{err} = 7.4\%$	N/A	N/A

Correct Identification   Correct process didn't have highest r^2_{adj} Not Identified

Discussion

The system successfully identifies the model processes with sufficient data. The system is robust to noise for the 1D process while the 2D process performed worse. Both processes were identified in the low-data limit with only moderate parameter error.

Future Work

- Improve feature search algorithm to include stochasticity through genetic algorithms
- Check robustness to different types of noise
- Apply to real world data of fluid flows

References

[1] Bridewell, W., Langley, P., Todorovski, L., \& DÅŸeroski, S. (2008). Inductive process modeling. Machine learning, 71(1), 1-32.
[2] 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*, 201517384.
[3] Kochenderfer, M. J., & Wheeler, T. A. (2018). *Algorithms for Optimization*. MIT Press.
[4] Rudin, L. I., Osher, S., & Fatemi, E. (1992). Nonlinear total variation based noise removal algorithms. Physica D: nonlinear phenomena, 60(1-4), 259-268.
[5] Van Genuchten, M. T., & Alves, W. J. (1982). Analytical solutions of the one-dimensional convective-dispersive solute transport equation (No. 157268). United States Department of Agriculture, Economic Research Service.