Discover governing reactions from concentration data

October 26, 2017

Introduction

When presented with a time series of possibly noisy concentrations as output of, e.g., measurements or simulations that were parameterized by microscopic rates (cite ReaDDy?), one can ask for the corresponding macroscopic rates and generating reaction network. In this paper we present an application of the shallow learning method SINDy [BPK15] which is able to identify generating parsimonious nonlinear dynamics in data that stems from dynamical systems. In our application we, opposed to the original method, do not only look for macroscopic rates of net species change but investigate the specific reactions that might have lead to the observations. We demonstrate the algorithm on a toy problem demonstrating the identification abilities.

The method

The underlying model that we want to fit the data to is a law of mass action type dynamical system. To this end, let S be number of species and T be the number observations. The concentration data at a time $t \in [1, T]$ can be represented by a vector

$$\mathbf{x}(t) = \begin{pmatrix} x_1(t) \\ \vdots \\ x_S(t) \end{pmatrix} \in \mathbb{R}^S.$$

Further, one can choose R possible ansatz reactions with their respective reaction function

$$\theta_r(\mathbf{x}(t)) = \begin{pmatrix} \theta_{r,1}(\mathbf{x}(t)) \\ \vdots \\ \theta_{r,S}(\mathbf{x}(t)) \end{pmatrix}$$

so that the change of concentration for species i at time t is represented by the dynamical system

$$\dot{\mathbf{x}}_i(t) = \sum_{r=1}^{R} \theta_{r,i}(\mathbf{x}(t))\xi_r,$$

where ξ_r are the to-be estimated macroscopic reaction rates.

Examples

Regression without regularization

Regression with regularization

References

[BPK15] Steven L Brunton, Joshua L Proctor, and J Nathan Kutz. Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *PNAS*, 1(609):1–26, 2015.