SINDYc-Based Identification of Controlled SIR Epidemic Dynamics with Physically Consistent Model Selection

Ayman Abdelaziz
University of Szeged, Szeged, Hungary

Data-driven identification of controlled epidemic models is essential for understanding the impact of public health interventions without requiring prior knowledge of system parameters. In this work, we apply Sparse Identification of Nonlinear Dynamics with Control (SINDYc) to identify the governing equations of a controlled SIR epidemic model that incorporates vaccination and treatment as external control inputs. A candidate function library is constructed from polynomial and interaction terms of the state variables and control signals, and sparse regression is used to recover the active epidemic and control-related terms from simulated time-series data. A key contribution of this work is a physically consistent model selection criterion that jointly enforces preservation of the dominant physical terms – including the infection term SI, vaccination term Su_1, and treatment term Iu_2 - while minimizing a relative derivative prediction error and penalizing unnecessary active terms. The proposed criterion is evaluated across three noise levels to assess the robustness of the identification procedure. Results demonstrate that SINDYc successfully recovers the physical structure of the one-control model under all noise conditions, while the two-control case reveals limitations of the standard automatically generated library under higher noise levels, motivating the use of physically informed custom libraries in future work.