Journal of Nonlinear Modeling and Analysis
Volume 6, Number 2, June 2024, 333–359
http://jnma-online.com
DOI:10.12150/jnma.2024.333

Model Selection of Dynamical Systems via Entropic Regression and Bayesian Information Criteria*

Jinhui Li[1] and Aiyong Chen[2,†]

Abstract Recovering system model from noisy data is a key challenge in the analysis of dynamical systems. Based on a data-driven identification approach, we develop a model selection algorithm called Entropy Regression Bayesian Information Criterion (ER-BIC). First, the entropy regression identification algorithm (ER) is used to obtain candidate models that are close to the Pareto optimum and combine as a library of candidate models. Second, BIC score in the candidate models library is calculated using the Bayesian information criterion (BIC) and ranked from smallest to largest. Third, the model with the smallest BIC score is selected as the one we need to optimize. Finally, the ER-BIC algorithm is applied to several classical dynamical systems, including one-dimensional polynomial and RC circuit systems, two-dimensional Duffing and classical ODE systems, three-dimensional Lorenz 63 and Lorenz 84 systems. The results show that the new algorithm accurately identifies the system model under noise and time variable t, laying the foundation for nonlinear analysis.

Keywords Data-driven, system identification, model selection, ER algorithm, BIC

MSC(2010) 37M05, 34A34, 93E24.

1. Introduction

Nonlinear dynamical systems have been one of the hot topics of fundamental theoretical research in recent years, both at home and abroad. the usual dynamical systems include Lorenz 84 system [5, 12, 34, 35], Lorenz 63 system [11, 21], Duffing system [23, 28, 29], RC circuit system [15] with trigonometric term, etc. The task that follows is to recover the periodic dynamical equations in order to study their properties. Thus, the data-driven approach [16, 24] that has emerged in recent years–obtaining dynamical behavior from data–can be an excellent solution to this problem.

For data-driven methods, common approaches include dynamic mode decomposition (DMD) [6, 9, 32, 33], Koopman theory [8, 17, 19], sparse representation, etc. Considering that most dynamical systems only contain a few terms, Brunton et

†the corresponding author.
Email address: lijinhuixlq@163.com(J. Li), aiyongchen@163.com(A. Chen)
[1]School of Mathematics and Computing Science, Guilin University of Electronic Technology, Guilin, Guangxi 541004, China
[2]Department of Mathematics, Hunan First Normal University, Changsha, Hunan 410205, China
*The authors were supported by Natural Science Foundation of Hunan Province (No. 2021JJ30166).

al. [7] proposed a nonlinear sparse regression algorithm (SINDy) in 2016. This algorithm obtains the sparse matrix based on the sequential least squares method and obtains the system model from. AlRahman et al. [2] proposed a new information-theoretic regression method (called entropy regression (ER)) to infer the sparse structure of the general model and its parameters based on the cost function of the metric to determine the error quantification that makes the existing methods fail under large noise and outliers. The ER method is robust in the presence of noise and outliers in the data. The ER algorithm [2] emphasizes information relevant to selecting models according to the entropy standard without models. The underlying terms will be incorporated into the model simply because they are correlated and thus together constitute the sparsest model.

For parameter estimation problems, most scholars choose the likelihood function as the objective function in their study, that is, the best model fit is achieved by maximizing the likelihood function. In this case, making a judgment requires the use of a criterion that can balance the accuracy and complexity of the model. The AIC criterion and the BIC criterion are commonly used [18]. Japanese statistician Akaike [1] proposed the AIC criterion for model selection in 1974. Schwarz [31] proposed the Bayesian information criterion (BIC) in 1978. It is also used for model selection and normally selects the model with the lowest BIC value. Thus AIC can be considered when the objective of the study is predictive and BIC can be considered when the objective of the study is descriptive [4].

Mangan et al. [26] proposed a new mathematical framework to evaluate a large set of candidate models for combinatorics using Akaike information criteria (AIC) with sparse regression for model selection. Sparse regression for nonlinear system identification was first performed by SINDy, and model selection was then performed via an information criterion. SINDy-AIC [26] sparsely selects dynamic models from information theoretic criteria, ranks the candidate models, and further shows that the correct model is strongly supported by the AIC score. Moreover, it utilizes data to uncover governing equations that are data-driven among the candidate models, where the data has undergone cross-validation and ranking.

In this work, we propose a model selection algorithm (ER-BIC algorithm) based on entropy regression (ER) and BIC, which are similar to SINDy-AIC [26] algorithm. First, we use the ER algorithm to obtain near-Pareto optimal models for nonlinear systems with a relatively small number of models. Second, the BIC scores of a limited number of the obtained models are calculated and sorted from small to large. Finally, the model with the lowest BIC score is chosen as the optimal model for the system. Our algorithm is robust to identify nonlinear systems, and can select a smaller number of models near Pareto optimality than the SINDy-AIC algorithm, but with higher accuracy. In particular, when the trigonometric function term is included, the model accuracy of the ER-BIC algorithm is higher than that of the SINDy-AIC algorithm.

The paper is divided into the following subsections. In section 2, we introduce the framework of the ER-BIC algorithm, where the theory of the ER algorithm, the notion of the BIC score, and the specific algorithmic procedure are presented. In section 3, some key applications and results of the ER-BIC algorithm are presented. The algorithm is applied to ordinary dynamical system models and shows robustness even under noisy data. In Section 4, we consider the accuracy of the Lorenz 63 system at different noise amplitudes and different amounts of data for the ER-BIC algorithm selection. The conclusions and discussion are provided in section 5.