Avrupa Bilim ve Teknoloji Dergisi
Özel Sayı, S. 254-263, Kasım 2020
© Telif hakkı EJOSAT'a aittir
Araştırma Makalesi
www.ejosat.com ISSN:2148-2683
European Journal of Science and Technology
Special Issue, pp. 254-263, November 2020
Copyright © 2020 EJOSAT
Research Article

Model Investigation of Nonlinear Dynamical Systems by Sparse Identification

Nezir Kadah[1*], Necdet Sinan Özbek[2]

[1] Adana Alparslan Türkeş Bilim ve Teknoloji Üniversitesi, Mühendislik Fakültesi, Elektrik-Elektronik Mühendisliği Bölümü, Adana, Türkiye (ORCID: 0000-0001-9320-1140)

[2] Adana Alparslan Türkeş Bilim ve Teknoloji Üniversitesi, Mühendislik Fakültesi, Elektrik-Elektronik Mühendisliği Bölümü, Adana, Türkiye (ORCID: 0000-0002-7184-9015)

(International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT) 2020 – 22-24 October 2020)

(**DOI:** 10.31590/ejosat.822361)

ATIF/REFERENCE: Kadah, N., & Özbek, N.S. (2020). Model Discovery of Nonlinear Dynamical Systems by Sparse Identification. *European Journal of Science and Technology*, (Special Issue), 254-263.

Abstract

The sparse identification of nonlinear dynamics (SINDy), which is based on the sparse regression techniques to identify the nonlinear systems, is one of the recent data-driven model identification methods. The model equations of the system are extracted from the data. Although sufficient data is available from most of the engineering, healthcare, and economic sciences, there are few well-defined models to represent the system behaviour that can also be estimated from data-driven methods. With this motivation in mind, this study presents offline data-driven identification techniques to build the mathematical model of nonlinear systems. The data-based sparse identification of nonlinear systems is elaborated with a number of examples. The performance of the identification procedure is discussed in terms of quantitative metrics in the presence of noisy measurements.

Keywords: Nonlinear System, Sparse Identification, Sparse Regression, Model Discovery, System Identification.

Seyrek Tanılama Yöntemi ile Doğrusal Olmayan Dinamik Sistemlerin Model İncelenmesi

Öz

Doğrusal olmayan sistemleri tanımlamak için seyrek regresyon tekniklerine dayanan doğrusal olmayan dinamiklerin seyrek tanımlanması (SINDy) son yıllarda ortaya konan veriye dayalı model tanımlama yöntemlerinden biridir. Sistem tanılamada sistemin model denklemleri verilerden çıkarılır. Mühendislik, sağlık hizmetleri ve ekonomi bilimlerinin çoğundan yeterli veri mevcut olmasına rağmen, sistem davranışını temsil eden çok az sayıda iyi tanımlanmış model vardır. Sistemin davranışı, veriye dayalı yöntemlerden de tahmin edilebilir. Bu motivasyon göz önünde bulundurularak, bu çalışma doğrusal olmayan sistemlerin matematiksel modelini oluşturmak için çevrimdışı veri odaklı tanımlama tekniklerini ele alır. Doğrusal olmayan sistemlerin veriye dayalı seyrek tanımlanması bir dizi örnekle detaylandırılır. Tanımlama işleminin performansı, gürültülü ölçümlerin varlığında bir takım nicel ölçümler üzerinden tartışılır.

Anahtar Kelimeler: Doğrusal Olmayan Sistemler, Seyrek Tanılama, Seyrek Regresyon, Model Keşfi, Sistem Tanılama.

[*] Corresponding Author: Adana Alparslan Türkeş Science and Technology University, Faculty of Engineering, Departments of Electrical-Electronics Engineering, Adana, Turkey, ORCID: 0000-0001-9320-1140, nkadah@atu.edu.tr

1. Introduction

The mathematical modelling of the real systems, which comprise complex nonlinear dynamics, is a tedious task due to uncertain parameters, nonlinear components, as well as the time-varying nature of the system. Additionally, unlike linear systems, nonlinear equations do not have a specific definition, besides, there is no specific approach or universal methodology in mathematical solutions. Furthermore, closed-form expressions for the solutions of the linear systems are not possible to solve nonlinear systems. Therefore, system identification as well as model discovery are extremely important tasks in the control systems engineering framework.

The advances in hardware and software have made it possible to make data-based predictions and modelling for highly complex real-systems. The investigation of system dynamics and model discovery using measured data has been very attractive research in mathematical physics and engineering. System identification enables modelling and design of high-performance systems. Thus, system identification and data-driven modelling have been developing rapidly in the last decades (Ljung, 2010), (Brunton & Kutz, 2019), (Kaheman et al., 2020). Great research efforts have been devoted to developing the most efficient method.

The identification strategies are defined as parametric or nonparametric with respect to control criteria and architecture limitations (Ayyad et al., 2020). Physical rules are paired with so-called white-box formulas in a conventional approach to achieving a system model. Alternatively, grey box models that have a fixed model structure may be used, but the selected model parameters are tuned by using measurement data (Fey et al., 2020).

System identification is challenging for complex nonlinear systems (Brunton et al., 2016), (Ranković et al., 2012). There is no specific mathematical approach for the solution of nonlinear systems, which may exhibit chaotic behaviour (Lusch et al., 2018), (Kadah, 2019). However, a breakthrough was made in data analysis with the advancements in artificial intelligence and machine learning (Ayyad et al., 2020), (Brunton et al., 2016), (Zucatti et al., 2020). It is to be noted that data-based model discovery was first tried in the studies of Kepler and Newton. Both scientists explored the dynamic relationship between variables using data (Brunton et al., 2016), (Niall M. Mangan, Steven L. Brunton, Member, Joshua L. Proctor, 2016), (Cortiella et al., 2020). Recently, much faster and more effective methods are used for data-driven modelling by using methods such as machine learning, artificial neural networks, and deep learning. However, these methods require large datasets and time to characterize dynamics (Quade et al., 2018), (Callaham et al., 2019).

It should be pointed out that complex dynamics affect the systems, which may behave uniquely at different times (Champion et al., 2019). Therefore, it is crucial to model the system mathematically for controlling the nonlinear system dynamics. Towards this goal, data-based learning online and offline methods have been developed to model dynamical systems (Quade et al., 2018), (Maheshwari et al., 2018). Moreover, generating models and estimating system parameters from observed data can effectively be used for systems comprising complex and chaotic behaviour. It is the sparse identification method, which is widely used in the literature and one of the most effective methods to generate models from data (Cortiella et al., 2020), (Maheshwari et al., 2018), (Wen et al., 2020). A few approaches exploit sparse approximations for system identification. However, all these methods are mainly based on sparse regression.

In this work, an offline identification method is proposed to obtain the mathematical model of the system by making use of the data of the system dynamics. To do this, SINDy (Sparse Identification of Nonlinear Dynamics) algorithm, which is a recent invention for system identification, is used through the sparse regression method (Chu & Hayashibe, 2020), (Horrocks & Bauch, 2020). Also, the equations of dynamical systems can be identified via SINDy from measurement data. SINDy technique essentially is based on the notion of the Koopman operator and of sparsity (Fey et al., 2020). In this algorithm, a library containing the variation of linear or nonlinear candidate basis functions is created first. Secondly, the active terms of the coefficients vector are computed via sparse regression. Finally, the model is updated with active terms and the remaining terms are ignored depending on the regularization parameter using the sparse regression. This method has been used successfully by many researchers to diagnose different systems (Brunton & Kutz, 2019), (Bhadriraju et al., 2019), (Goharoodi et al., 2018). The scheme of data-driven modelling is depicted in Fig. 1.

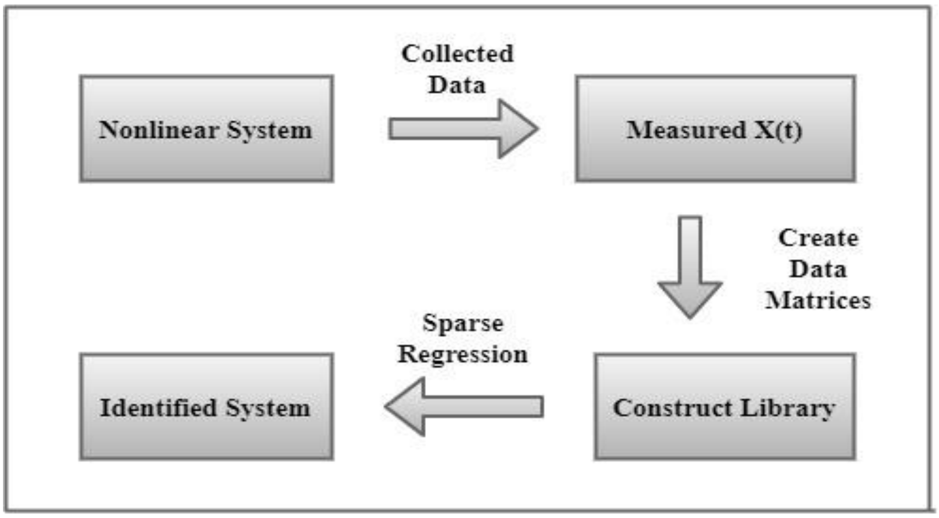

Fig. 1. The scheme of SINDy

In this study, the dynamics of different systems are measured, and the models are estimated with the SINDy method. One of the major challenges in data-driven modelling via the SINDy method is undoubtedly noisy measurements. A number of problems can occur in the presence of noise. To mention a few, noise may cause erroneous parameter estimates (Cortiella et al., 2020). Further, the computing of the derivatives of the states will be big trouble, especially in data measurements. Thus, pre-processing methods such as noise filtering methods are suggested to reduce the effect of noise (Brunton & Kutz, 2019), (Chartrand, 2011), (H. Li et al., 2020).

The contributions of this study are given as follows: Models of nonlinear systems are discovered by using SINDy method in the presence of noisy measurement data. The robustness of sparse regression is investigated through some applications on several benchmark systems.

This paper is organized as follows: The theoretical background of the elaborated strategies is given in the second section. The sparse identification of nonlinear systems is also addressed in the second section. Different applications are highlighted in the third section. Finally, concluding remarks are given in the last section.

2. Material and Method

2.1. Theoritical Background

In this section, the central notion of the algorithm is investigated theoretically. A commonly used model for the dynamical system is

$$x'(t) = \frac{dx(t)}{dt} = f\left(x(t)\right) \tag{1}$$

f is dynamic, $x = \left[x_1\ x_2\ \cdots\ x_n\right] \in \Re^n$ is state vector and the components of f are sparse in a given function space. The discrete-time dynamical system of eq. (1) is

$$x_{k+1} = F\left(x_k\right) = x_k + \Delta x_k \tag{2}$$

where x_k can be got by sampling the trajectory.

Discretization is the process of separating continuous functions and equations into discrete components in mathematic. This is referring to the discretization of features and variables in machine learning as well as statistics.

In this paper, Euler Step Method is used for discretization. Since the Euler step gives time step of $t + \Delta T$. Hence, this structure gives a time map of the data. Numerically discretization would be used for a time map in the future.

- Input is the state of the system at time t
- Output is the state of the system at time $t + \Delta T$

The discrete-time F is given by the flow map (Champion et al., 2019):

$$F\left(x_k\right) = x_k + \int_{k\Delta t}^{(k+1)\Delta t} f\left(x(\tau)\right) d\tau \tag{3}$$

m step of time-series data is collected from eq. (1) and given as a data matrix X :

$$X = \begin{bmatrix} x^T{}_1 \\ \vdots \\ x^T{}_{m-1} \end{bmatrix} = \begin{bmatrix} | & | & | & \dots & | \\ x_1 & x_2 & x_3 & \dots & x_{m-1} \\ | & | & | & \dots & | \end{bmatrix} \updownarrow \text{time} \tag{4}$$

where $[.]^T$ donates transpose of the matrix and m is the count of measurements. Similarly, the derivative of (4) yields by shifting the data

$$X' = \begin{bmatrix} x_1'^T \\ \vdots \\ x_{m-1}'^T \end{bmatrix} = \begin{bmatrix} | & | & | & \dots & | \\ x_1' & x_2' & x_3' & \dots & x_{m-1}' \\ | & | & | & \dots & | \end{bmatrix} = \begin{bmatrix} | & | & | & \dots & | \\ x_2 & x_3 & x_4 & \dots & x_m \\ | & | & | & \dots & | \end{bmatrix} \updownarrow \text{time} \tag{5}$$

The central derivative method (6) can be used to compute the numerical derivative of a function f . In this paper, X' will be generated by using the first-order central derivative method from the measured data. The first-order central derivative can be given as follows:

$$f'(x) = \frac{f(x+\Delta x) - f(x-\Delta x)}{2\Delta x} \tag{6}$$

2.1.1. Time-Delay Embedding

It is significant to make measurements accurately, in data-driven modelling, control, and prediction. However, the measurement data may not contain all the dynamics of a system. Time-delay embedding method can be implemented to reconstruct the system from a sequence of observations of the state and an idea about the dynamics of the system can be obtained with this method (Brunton & Kutz, 2019), (Champion et al., 2019), (Brunton et al., 2017), (De Silva, Callaham, et al., 2020).

The time-delay embedding method allows reconstructing data of dynamic systems. With this method, a one-dimensional data set is configured to be multi-dimensional. Thus, it is possible to enrich $x(t)$ measurement by using $x(t-\tau)$, which are delay coordinates (Brunton et al., 2017), (Brunton & Nathan Kutz, 2019). This method does not reconstruct the properties or behaviour of the dynamic system, but only changes its structure in phase space.

2.1.2. Hankel Matrix and Singular Value Decomposition

Hankel matrix is the time-shift "ΔT" of the previous rows. It is a time-delay embedding. So, the Hankel matrix H can be generated from a time-series of measurement (Brunton & Kutz, 2019), (Brunton et al., 2017), (Jain & Pachori, 2014), (Jain & Pachori, 2015).

$$H = \begin{bmatrix} x(t_1) & x(t_2) & \cdots & x(t_q) \\ x(t_2) & x(t_3) & \cdots & x(t_{q+1}) \\ \vdots & \vdots & \ddots & \vdots \\ x(t_p) & y(t_{p+1}) & \cdots & x(t_m) \end{bmatrix} \tag{7}$$

Then dominant time-delay coordinates are found by taking SVD of Hankel Matrix. Moreover, the order of the system can be defined by the Hankel matrix rank and this order can be computed with the left-singular vector of SVD (Champion et al., 2019), (Brunton et al., 2017), (Lim et al., 1998).

$$H = U \sum V^* \tag{8}$$

where "*" denotes the conjugate transpose of the matrix. SVD, which is one of the most useful matrix decompositions, can be applied to any matrix in linear algebra (Ford, 2014). SVD allows us to decompose the matrix as the product of three matrices. U and V are the unitary or orthogonal matrices and $\sum$ is a diagonal matrix. SVD is also a coordinate transformation or mapping strategy, thus it is not unique (Brunton & Kutz, 2019).

2.2. Sparse Identification of Nonlinear Dynamics (SINDy)

The performance of data-driven identification can reduce for strongly nonlinear and high dimensional systems. However, SINDy is a perfect method in the low-data limit (Quade et al., 2018), (Kaiser et al., 2018). SINDy method can discover mathematical equations of a system with the comparatively little data-set provided that calculates the derivative (Champion et al., 2019).

This algorithm is mainly concerned with the dominant terms of the function and ignores the terms that low effect (Brunton et al., 2017). Because the function f , which contains the dynamics of the system, is dominantly affected by only a few terms (De Silva, Higdon, et al., 2020).

SINDy that uses sparse regression to find coefficients from measured data. Each of the constants is the coefficients of terms that affect system dynamics. Then, the mathematical model of the system is created by using dominant terms from the library that contains all candidate terms of a dynamic system. This library contains all possible combinations of terms (Niall M. Mangan, Steven L. Brunton, Member, Joshua L. Proctor, 2016), (Quade et al., 2018), (Champion et al., 2019), (Kaiser et al., 2018).

Consider a library of possible nonlinear functions $\Theta(X)$ that may consist of polynomial terms, trigonometric functions or etc., (Corbetta, 2020), see Eq. (9).

$$\Theta(X)=\begin{bmatrix} | & | & | & \ldots & | & | & \cdots & | & \cdots \\ 1 & X & X^2 & \ldots & X^a & \cos(X) & \cdots & e^X & \cdots \\ | & | & | & \ldots & | & | & \cdots & | & \cdots \end{bmatrix} \tag{9}$$

Equation (9), the matrix $\Theta(X)$ contains many columns and each of these columns represents one of the candidate nonlinear functions. The matrix of coefficient Ξ has only a few active or dominant terms. The remaining terms are almost zero (Quade et al., 2018), (Brunton & Kutz, 2019), (Brunton et al., 2017), (Rudy et al., 2017).

$$\Xi=\begin{bmatrix} | & | & \cdots & | \\ \zeta_1 & \zeta_1 & \cdots & \zeta_m \\ | & | & \cdots & | \end{bmatrix} \tag{10}$$

2.3. Sparse Regression Method

Most of the nonlinear systems may contain polynomial terms. However, only a few of these terms affect or represent system dynamics, while many are insignificant, therefore they can be removed. To calculate the subset of candidate terms of the system, sparse regression method like “least absolute shrinkage and selection operator (LASSO)”, ElasticNet, “least-square method (LSM)” and “sequential thresholded least-squares (STLSQ)” can be used. Unlike other methods, LASSO and STLSQ have noise elimination and improved robustness performance (Brunton & Kutz, 2019), (Brunton et al., 2017), (Rudy et al., 2017), (Kukreja et al., 2006), (Calafiore et al., 2015).

"Minimum norm least-square", “Moore-Penrose Pseudoinverse methods” and “backslash” are also regression methods, which can be used for linear systems represented by $Ax=b$. However, none of these methods has penalty parameters. Therefore they do not provide performance like LASSO or ElasticNet (Brunton & Kutz, 2019).

LASSO, which is a widely used regression method for data-driven modelling, learns the linear relationship between a dependent variable and explanatory variables (Misra et al., 2020), (J. Li & Li, 2020). Besides, LASSO is used extensively for the model or feature selection and system identification in statistics, machine learning, and control theory (Brunton & Kutz, 2019), (Kukreja et al., 2006), (Calafiore et al., 2015), (J. Li & Li, 2020). LASSO implements an ℓ_1 regularization term that can produce sparse coefficients.

LASSO algorithm is expressed as

$$\min_w \frac{1}{2n}\|Xw-Y\|_2^2+\lambda\|w\|_1 \tag{11}$$

where X and Y represent the input and target vector; w is sparse or coefficients vector and n is the number of sample data. It is to be noted that LASSO requires the determination of a regularization term that penalizes the ℓ_1 norm, which is λ for the penalty parameter. ℓ_1 norm forces particular elements to be exactly zero. As the value of λ parameter decreases, the degree of sparsity decreases and consequently the error in the model decreases. Because, when the value of λ rises more terms are ignored (Corbetta, 2020), (Calafiore et al., 2015), (Misra et al., 2020), (Tibshirani, 1996).

Finally, the mathematical model of the dynamical system can be generated from (12) for continuous and discrete-time systems.

$$X' \approx \Theta\left(X\left(t\right)\right)\Xi$$
$$X_{k+1} \approx \Theta\left(X_k\right)\Xi \quad (12)$$

3. Results and Discussion

The application of SINDy for nonlinear systems is discussed in this section. Two nonlinear systems namely, Rössler and Lotka-Volterra equations are chosen for system identification applications. Suppose that in the first example, initially only one state is measured to calculate the number of states. However, suppose in the second step, all states are measured for both applications. All results and simulations are generated using MATLAB 2020a.

3.1. Rössler Attractor System

In the first case study, predictions will be made about the behaviour of a system with limited measurements of the trajectory data. Also, in the case of full-state measurements, the equations of the model are identified correctly by SINDy.

Rössler attractor equation is a very popular example of a nonlinear system. Consider Rössler system represented by three differential equations as follow:

$$x' = -y - z$$
$$y' = x + \alpha y \quad (13)$$
$$z' = \beta - \gamma z + xz$$

with the numerical parameters $\alpha = 0.4$, $\beta = 2$ and $\gamma = 4$.

Firstly, Data will be collected from (13) only for one dimension (x) or variable for random initial conditions. Then using the time-delay embedding method, Hankel Matrix will be generated.

Secondly, the SVD of the Hankel matrix can be calculated. The rank of this system can be obtained with SVD of the Hankel Matrix. The first three modes of the Hankel matrix are dominants as can be seen in Fig. 2. Thus, it can be truncated at three (x, y, z).

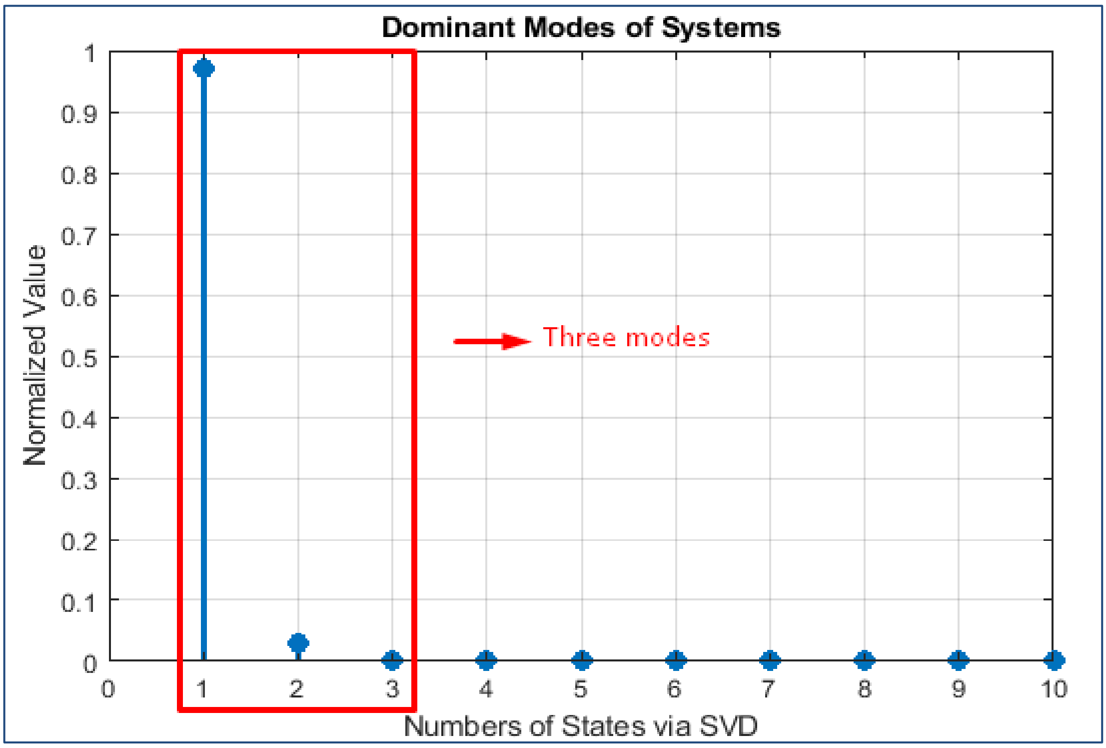

Fig. 2. Selection of dominant states of system

The type of functions, which can be used to represent the system, is obtained when the left-singular value matrix (U) created by SVD is examined. The system has nearly polynomial degrees for Rossler attractor. So, the time-delay embedding of Hankel Matrix produces the basis set of candidate functions.

Time-delay embedding theorem reconstructed manifold produced from lags of a single variable. Each point in the three-dimensional reconstruction can be thought of as a time segment with different points capturing a different segment of the history of variable x and the reconstructed manifold is then the library or collection of the historical behaviour of lags. This method represents a one-to-one mapping between the original manifold and reconstruction, see Fig. 3.

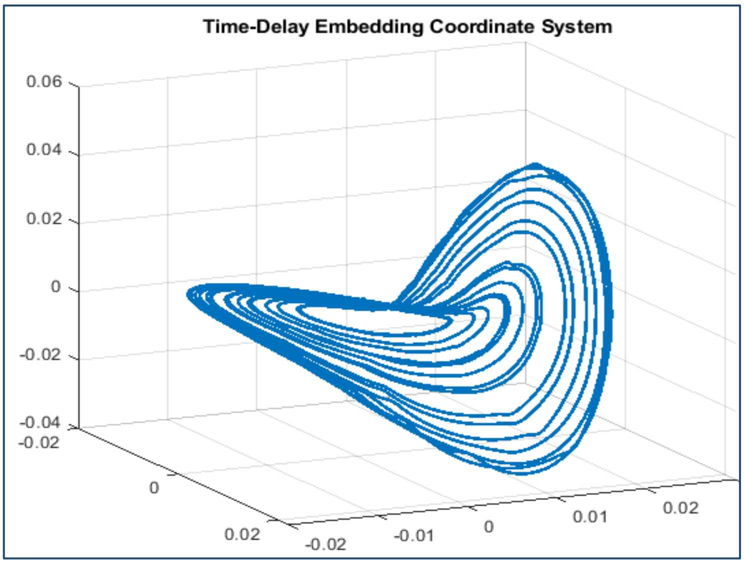

Fig. 3. Phase space mapping

Vectors of all possible polynomial nonlinearities can be structured from data as in (14). Then regression is applied to find which linear combination of these nonlinearity terms represent x', y', and z'. These three variables will actually be the state equations of the system. Because the arithmetic sequence of the term and coefficient of each state variable will be generated.

$$\underbrace{\begin{bmatrix} x' & y' & z' \end{bmatrix}}_{X'} = \underbrace{\begin{bmatrix} | & | & | & | & | & | & | & .... \\ 1 & x & y & z & x^2 & xy & xz & .... \\ | & | & | & | & | & | & | & .... \end{bmatrix}}_{\Theta(X)} \underbrace{\begin{bmatrix} \xi_1 & \xi_2 & \xi_3 \end{bmatrix}}_{\Xi} \tag{14}$$

Finally, the coefficients of the dominant term are obtained with the SINDy algorithm. Basically, the equation of the nonlinear system is represented by (12).

The data of the library containing the system dynamics are calculated, but the dominant ones of these variables are used. To do this, three type regression model is used. In Fig. 4, the values of the polynomial terms, which are obtained by the regression method, are given. In addition, clearly active terms and coefficients can be seen from this figure.

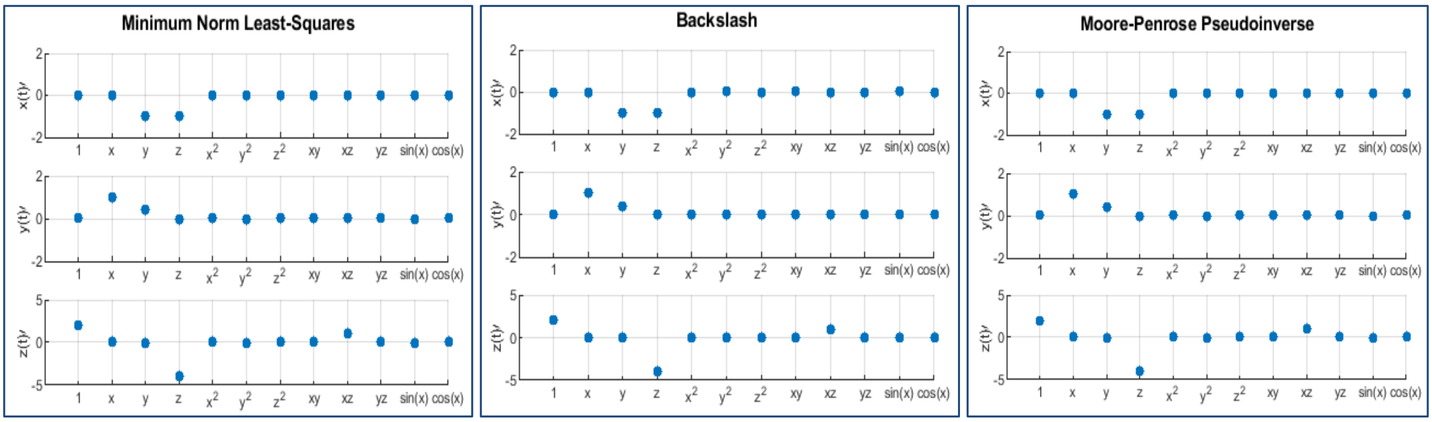

Fig. 4. Determined active terms of the library for Rössler system

Using SINDy with Minimum norm least-square, backslash and Moore-Penrose Pseudoinverse methods, it is determined that Eq. (15) of the system consists of three state variables.

$$\begin{aligned} x' &= -y - 0.99z \\ y' &= x + 0.4y \\ z' &= 2 - 3.98z + xz \end{aligned} \tag{15}$$

3.2. Lotka-Volterra System

In this application, the system is identified by using SINDy when the system dynamics are measured full-state and noisy.

Consider the Lotka-Volterre system in (16) with numerical parameters $\alpha = 2.5$, $\beta = -\delta = -0.1$ and $\gamma = -1$. $x_1(t)$ and $x_2(t)$ are predator and prey population respectively.

$$\begin{aligned} x_1' &= \alpha x_1 + \beta x_1 x_2 \\ x_2' &= \delta x_1 x_2 + \gamma x_2 \end{aligned} \tag{16}$$

Data matrices will be generated from (16) for both states with initial conditions $x_1(0) = 8$ and $x_2(0) = 3$, and noisy measurement. Then, derivative matrices are created for both states by first-order central derivative method with the collected data. Then, candidate library of possible nonlinear terms can be structured from data as in (17).

$$\begin{bmatrix} x_1' & x_2' \end{bmatrix} = \begin{bmatrix} | & | & | & .... & | & .... \\ x_1 & x_2 & x_1x_2 & .... & \sin(x_1) & .... \\ | & | & | & .... & | & .... \end{bmatrix} \begin{bmatrix} \xi_1 & \xi_2 \end{bmatrix} \tag{17}$$

Finally, the LASSO regression is applied to find the linear combination of these nonlinearity terms represented by x_1' and x_2'. These two variables will be the state equations of the system. Since, the arithmetic sequence of the term and coefficient of each state variable will be generated. Towards this goal, LASSO regression model is used and as a result, non-dominant terms are eliminated.

In Fig. 5, the coefficients of possible terms that represent the system are given. In addition, active terms can be easily recognized from this figure.

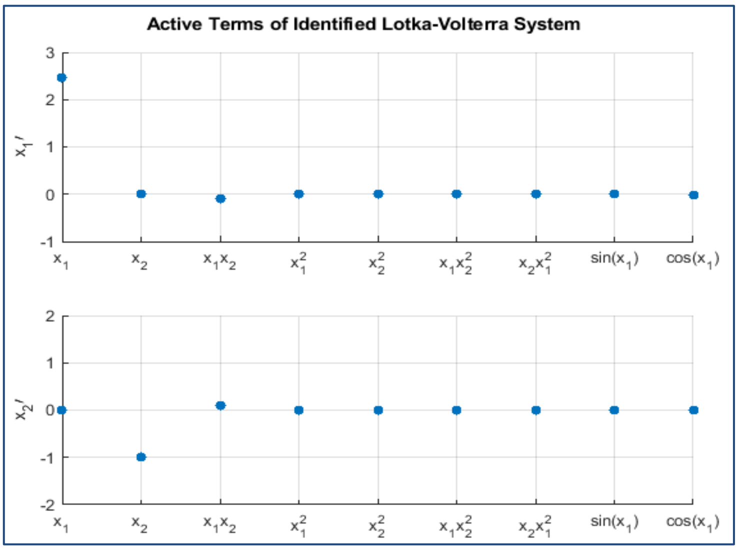

Fig. 5. Determined terms of the library for Lotka-Volterra system

As a result, using SINDy, it was determined that the equation (18) of the system consists of two state variables, and this is achieved despite being noisy data.

$$x_1' = 2.48x_1 - 0.01x_1x_2$$
$$x_2' = -0.1x_2 + 0.01x_1x_2 \quad (18)$$

4. Conclusions and Recommendations

This paper investigates the modelling of dynamical systems using SINDy method, with an emphasis on different regression methods. The modelling of nonlinear dynamical systems from data that is collected or measured from the system is a key insight for "System Identification". SINDy method, which is based on sparse regression, performs quite satisfactory results for full-state measurements. However, the time-delay embedding method allows the designer to construct dynamic models in the presence of partial measurements. Also, this study has been verified in which Lasso regression can find the best coefficient vector. This makes the equation of the system simpler and easier to understand.

This manuscript contributes to the behaviour of system variables, the system model, and methods to predict the system's future plots. It was also emphasized that the model of an unknown system can be easily obtained with measurements. Moreover, SINDy method can be easily integrated with areas such as controller design and model prediction.

References

Ayyad, A., Chehadeh, M., Awad, M., & Zweiri, Y. (2020). Real-Time System Identification Using Deep Learning for Linear Processes With Application to Unmanned Aerial Vehicles. *IEEE Access*, *8*, 122539–122553. https://doi.org/10.1109/ACCESS.2020.3006277

Bhadriraju, B., Narasingam, A., & Kwon, J. S. Il. (2019). Machine learning-based adaptive model identification of systems: Application to a chemical process. *Chemical Engineering Research and Design*, *152*, 372–383. https://doi.org/10.1016/j.cherd.2019.09.009

Brunton, S. L., Brunton, B. W., Proctor, J. L., Kaiser, E., & Kutz, J. N. (2017). Chaos as an Intermittently Forced Linear System. *Nature Communications*, *8*(19), 34. http://faculty.washington.edu/sbrunton/HAVOK.zip

Brunton, S. L., & Kutz, J. N. (2019). Data-Driven Science and Engineering: Machine Learning, Dynamical Systems and Control. In *Cambridge University Press*. Cambridge University Press. https://doi.org/10.1017/9781108380690

Brunton, S. L., & Nathan Kutz, J. (2019). Methods for data-driven multiscale model discovery for materials. *J. Phys.: Mater*, *2*, 44002. https://doi.org/10.1088/2515-7639/ab291e

Brunton, S. L., Proctor, J. L., & Nathan Kutz, J. (2016). Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *PNAS*, *113*(15). https://doi.org/10.1073/pnas.1517384113

Calafiore, G. C., El Ghaoui, L. M., & Novara, C. (2015). Sparse identification of posynomial models. *Automatica*, *59*, 27–34. https://doi.org/10.1016/j.automatica.2015.06.003

Callaham, J. L., Maeda, K., & Brunton, S. L. (2019). Robust flow reconstruction from limited measurements via sparse representation. *PHYSICAL REVIEW FLUIDS*, *4*, 103907. https://doi.org/10.1103/PhysRevFluids.4.103907

Champion, K. P., Brunton \ddagger, S. L., & Nathan Kutz, J. (2019). Discovery of Nonlinear Multiscale Systems: Sampling Strategies and Embeddings. *SIAM J. APPLIED DYNAMICAL SYSTEMS*, *18*(1), 312–333. https://doi.org/10.1137/18M1188227

Chartrand, R. (2011). Numerical Differentiation of Noisy, Nonsmooth Data. *ISRN Applied Mathematics*, *2011*, 1–11. https://doi.org/10.5402/2011/164564

Chu, H. K., & Hayashibe, M. (2020). Discovering Interpretable Dynamics by Sparsity Promotion on Energy and the Lagrangian. *IEEE Robotics and Automation Letters*, *5*(2), 2154–2160. https://doi.org/10.1109/LRA.2020.2970626

Corbetta, M. (2020). Application of sparse identification of nonlinear dynamics for physics-informed learning. *2020 IEEE Aerospace Conference*, 1–8. https://doi.org/10.1109/aero47225.2020.9172386

Cortiella, A., Park, K.-C., & Doostan, A. (2020). *Sparse Identification of Nonlinear Dynamical Systems via Reweighted ℓ_1-regularized Least Squares*. http://arxiv.org/abs/2005.13232

De Silva, B. M., Callaham, J., Jonker, J., Goebel, N., Klemisch, J., Mcdonald, D., Hicks, N., Nathan Kutz, J., Brunton, S. L., & Aravkin, A. Y. (2020). *Physics-informed machine learning for sensor fault detection with flight test data*. 21.

De Silva, B. M., Higdon, D. M., Brunton, S. L., & Kutz, J. N. (2020). Discovery of Physics From Data: Universal Laws and Discrepancies. *Frontiers in Artificial Intelligence*, *3*(25), 17. https://doi.org/10.3389/frai.2020.00025

Fey, A., Thiele, G., & Krüger, J. (2020). *System identification of a hysteresis-controlled chiller plant using SINDy*. 8. http://arxiv.org/abs/2003.07465

Ford, W. (2014). Numerical Linear Algebra with Applications: Using MATLAB. In *Academic Press*. Elsevier Inc. https://doi.org/10.1016/C2011-0-07533-6

Goharoodi, S. K., Dekemele, K., Dupre, L., Loccufier, M., & Crevecoeur, G. (2018). Sparse Identification of Nonlinear Duffing Oscillator From Measurement Data. *IFAC-PapersOnLine*, *51*(33), 162–167. https://doi.org/10.1016/j.ifacol.2018.12.111

Horrocks, J., & Bauch, C. T. (2020). Algorithmic discovery of dynamic models from infectious disease data. *Scientific Reports*, *10*(1), 1–18. https://doi.org/10.1038/s41598-020-63877-w

Jain, P., & Pachori, R. B. (2014). Event-Based Method for Instantaneous Fundamental Frequency Estimation from Voiced Speech Based on Eigenvalue Decomposition of the Hankel matrix. *IEEE Transactions on Audio, Speech and Language Processing*, *22*(10), 1467–1482. https://doi.org/10.1109/TASLP.2014.2335056

Jain, P., & Pachori, R. B. (2015). An iterative approach for decomposition of multi-component non-stationary signals based on eigenvalue decomposition of the Hankel matrix. *Journal of the Franklin Institute*, *352*(10), 4017–4044. https://doi.org/10.1016/j.jfranklin.2015.05.038

Kadah, N. (2019). *Doğrusal Olmayan RLC Devrelerinin Kararlılık ve Pasiflik Analizi*. Van Yuzuncu Yil University.

Kaheman, K., Kutz, J. N., & Brunton, S. L. (2020). *SINDy-PI: A Robust Algorithm for Parallel Implicit Sparse Identification of Nonlinear Dynamics*. http://arxiv.org/abs/2004.02322

Kaiser, E., Nathan Kutz, J., & Brunton, S. L. (2018). Sparse identification of nonlinear dynamics for model predictive control in the low-data limit. *Proceedings of The Royal Society A Mathematical Physical and Engineering Sciences*, *474*(2219), 14. https://doi.org/https://doi.org/10.1098/rspa.2018.0335

Kukreja, S. L., Löfberg, J., & Brenner, M. J. (2006). a Least Absolute Shrinkage and Selection Operator (Lasso) for Nonlinear System Identification. *IFAC Proceedings Volumes*, *39*(1), 814–819. https://doi.org/10.3182/20060329-3-au-2901.00128

Li, H., Wang, Z., & Wang, W. (2020). A Local Sparse Screening Identification Algorithm with Applications. *Computer Modeling in Engineering & Sciences*, *124*(2), 765–782. https://doi.org/10.32604/cmes.2020.010061

Li, J., & Li, X. (2020). Online sparse identification for regression models. *Systems and Control Letters*, *141*, 104710. https://doi.org/10.1016/j.sysconle.2020.104710

Lim, R. K., Phan, M. Q., & Longman, R. W. (1998). State-Space System Identification with Identified Hankel Matrix. *Department of Mechanical and Aerospace Engineering Technical Report*, *3045*, 1–36.

Ljung, L. (2010). Perspectives on system identification. *Annual Reviews in Control*, *34*(1), 1–12. https://doi.org/10.1016/j.arcontrol.2009.12.001

Lusch, B., Kutz, J. N., & Brunton, S. L. (2018). Deep learning for universal linear embeddings of nonlinear dynamics. *Nature Communications*, *9*(1), 4950. https://doi.org/10.1038/s41467-018-07210-0

Maheshwari, J., Jariwala, R., Pradhan, S., & George, N. V. (2018). Online Least Angle Regression Algorithm for Sparse System Identification. *2017 IEEE International Symposium on Signal Processing and Information Technology, ISSPIT 2017*, 191–195. https://doi.org/10.1109/ISSPIT.2017.8388640

Misra, S., Li, H., & He, J. (2020). Robust geomechanical characterization by analyzing the performance of shallow-learning regression methods using unsupervised clustering methods. In *Machine Learning for Subsurface Characterization* (pp. 129–155). Elsevier Inc. https://doi.org/10.1016/b978-0-12-817736-5.00005-3

Niall M. Mangan, Steven L. Brunton, Member, Joshua L. Proctor, and J. N. K. (2016). Inferring Biological Networks by Sparse Identification of Nonlinear Dynamics. *IEEE Transactions on Molecular, Biological and Multi-Scale Communications*, *22*(1), 12. https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=7809160

Quade, M., Abel, M., Nathan Kutz, J., & Brunton, S. L. (2018). Sparse Identification of Nonlinear Dynamics for Rapid Model Recovery. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 10. https://github.com/Ohjeah/sparsereg

Ranković, V., Radulović, J., Grujović, N., & Divac, D. (2012). Neural Network Model Predictive Control of Nonlinear Systems Using Genetic Algorithms. *International Journal of Computers, Communications and Control*, *7*(3), 540–549. https://doi.org/10.15837/ijccc.2012.3.1394

Rudy, S. H., Brunton, S. L., Proctor, J. L., & Kutz, J. N. (2017). Data-driven discovery of partial differential equations. *Science Advances*, *3*(4). http://advances.sciencemag.org/

Tibshirani, R. (1996). Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society*, *58*(1), 267–288.

Wen, H. X., Yang, S. Q., Hong, Y. Q., & Luo, H. (2020). A Partial Update Adaptive Algorithm for Sparse System Identification. *IEEE/ACM Transactions on Audio Speech and Language Processing*, *28*, 240–255. https://doi.org/10.1109/TASLP.2019.2949928

Zucatti, V., Lui, H. F. S., Pitz, D. B., & Wolf, W. R. (2020). Assessment of reduced-order modeling strategies for convective heat transfer. *Numerical Heat Transfer; Part A: Applications*, *77*(7), 702–729. https://doi.org/10.1080/10407782.2020.1714330