

# System Identification with Copula Entropy

Jian MA\*

Hitachi China Research Laboratory

## Abstract

Identifying differential equation governing dynamical system is an important problem with wide applications. Copula Entropy (CE) is a mathematical concept for measuring statistical independence in information theory. In this paper we propose a method for identifying differential equation of dynamical systems with CE. The problem is considered as a variable selection problem and solved with the previously proposed CE-based method for variable selection. The proposed method is composed of two components: the difference operator and the CE estimator. Since both components can be done non-parametrically, the proposed method is therefore model-free and hyperparameter-free. The simulation experiment with the 3D Lorenz system verified the effectiveness of the proposed method.

**Keywords:** Copula Entropy; System Identification; Dynamical Systems

## 1 Introduction

Differential equation is a main tool for modelling dynamical systems and has very wide applications in diverse fields. Discovering differential equation from data is a important problem in dynamical systems and is becoming popular recently. Many different approaches have been proposed to tackle this problem [1].

Such equation discovery problem is usually considered as a regression problem in which the regression function between the states of systems and the derivatives of the system states is identified from data. Given a dynamical system that has mathematical model as

$$\frac{dx_i}{dt} = f(\mathbf{x}, t), \quad (1)$$

where  $x_i, i = 1, \dots, N$  are system states. The equation discovery is to identify the function  $f$ . Many regressions models have been applied to this problem, such as Gaussian processs [2], SINDy with sparsity assumption [3].

In this paper, we propose to use copula entropy (CE) to identify the regression function in differential equations. CE is a mathematical concept for measuring statistical independence. It is defined with the copula theory by

---

\*Email: majian@hitachi.cn

Ma and Sun [4] and has many properties, such as non-positive, multivariate, invariant to monotonic transformation, equivalent to correlation coefficient in Gaussian cases. A non-parametric estimator of CE was also proposed [4].

Recently, CE was proposed to solve the variable selection problem [5] and shown to be advantageous than the tradition similar methods, such as AIC, Lasso, etc. The CE-based variable selection method has been applied to many other fields, such as hydrology [6], medicine [7], manufacturing [8], reliability [9], energy [10], etc.

We propose a method for equation discovery based on the CE-based variable selection method. It is to estimate the CE values between the derivatives of system states and many candidate covariates derived from system states and then select those covariates associated with high CE values for the identified function  $f$ . The proposed method is composed of two components: the difference operator for calculating the derivatives of system states and the non-parametric CE estimator. Since both components can be done without any assumption on the underlying system, the proposed method is therefore model-free. In this paper we will verify the proposed method on the 3D Lorenz system, in which both first-order and second-order covariates are included in its system equations.

This paper is organized as follows: the theory and estimation of CE is introduced in Section 2, the proposed method is presented in Section 3, the simulation experiment is given in Section 4, followed by some discussion in Section 5, finally Section 6 concludes the paper.

## 2 Copula Entropy

### 2.1 Theory of CE

Copula theory unifies representation of multivariate dependence [11, 12]. According to Sklar theorem [13], a multivariate joint density function can be represented as a product of its marginals and copula density which represents dependence structure among random variables. Please refer to [4] for notations.

With copula density, Ma and Sun [4] defined a new mathematical concept, called Copula Entropy, as follows:

**Definition 1** (Copula Entropy). *Let  $\mathbf{X}$  be random variables with marginals  $\mathbf{u}$  and copula density  $c(\mathbf{u})$ . The CE of  $\mathbf{X}$  is defined as*

$$H_c(\mathbf{x}) = - \int_{\mathbf{u}} c(\mathbf{u}) \log c(\mathbf{u}) d\mathbf{u}. \quad (2)$$

In information theory, MI is a fundamental concept different from entropy [14]. Ma and Sun [4] proved that MI is essentially negative CE, as follows:

**Theorem 1.** *MI of random variables is equivalent to negative CE:*

$$I(\mathbf{x}) = -H_c(\mathbf{x}), \quad (3)$$

where  $I$  denotes MI.

Theorem 1 has a simple proof [4] and an instant corollary on the relationship between the information contained in joint density function, marginals, and copula density.

**Corollary 1.** *The joint entropy of random variables equals to the sum of the entropies of each variable and the CE of the random variables:*

$$H(\mathbf{x}) = \sum_i H_i(x_i) + H_c(\mathbf{x}), \quad (4)$$

where  $H$  denotes entropy.

The above worthy-a-thousand-word results cast insight into the relationship between MI and copula and therefore build a bridge between information theory and copula theory.

## 2.2 Estimating CE

MI, as a fundamental concept in information theory, has wide applications in different fields. However, estimating it has been notoriously difficult. Under the blessing of Theorem 1, Ma and Sun [4] proposed a simple and elegant non-parametric method for estimating CE(MI) from data, which is composed of two steps:

1. estimating Empirical Copula Density (ECD);
2. estimating CE from the estimated ECD.

Given a sample  $\{X_1, \dots, X_T\}$  i.i.d. generated from random variables  $\mathbf{X} = \{X_1, \dots, X_N\}$ , one can easily derive ECD using empirical functions:

$$F_i(x_i) = \frac{1}{T} \sum_{t=1}^T I(X_t^i \leq x_t^i), \quad (5)$$

where  $i = 1, \dots, N$  and  $I$  denotes indicator function. Once ECD is estimated, estimating CE becomes a problem of entropy estimation which can be tackled with many existing methods. The k-Nearest Neighbor method [15] was suggested in [4] for such entropy estimation. In this way, a nonparametric method for estimating CE is derived.

## 3 Proposed Method

In this section we propose a method for identifying dynamical system based on CE. Identifying dynamical system is treated as a problem of variable selection in which the covariates related to the derivatives of the system are selected. In this way, the CE-based method for variable selection can be applied to solve this problem. Particularly, the estimated CEs between system variables and derivatives are considered as criteria for variable selection.

The proposed method is composed of three steps:

1. calculating the derivative of system variables with differential operator;
2. estimating the CEs between the calculated derivatives and the covariates of the system;
3. selecting the covariates with high CE value for each derivatives.

With a time series of system variables, the derivative of them can be derived with difference operator. Given a pair of variable  $X, Y$ , the derivative of  $y$  with respect to  $x$  can be simply calculated as follows:

$$\frac{dx}{dt}|_{t=t_0} = \frac{x_{t_1} - x_{t_0}}{t_1 - t_0}. \quad (6)$$

For the second step of the proposed method, the non-parametric estimator of CE can be used.

Since the difference operator and the CE estimator are both non-parametric, the proposed method is therefore model-free and can be applied to any system without any assumptions.

## 4 Simulations

### 4.1 Experiments

We conducted simulation experiments to verify the effectiveness of the proposed method. In the simulation, the 3D Lorenz system [16] is used to simulate time series data.

**Lorenz system** The mathematical equations of the 3D Lorenz system are as follows:

$$\begin{aligned} \frac{dx}{dt} &= \sigma(y - x), \\ \frac{dy}{dt} &= \rho x - y - xz, \\ \frac{dz}{dt} &= -\beta z + xy, \end{aligned} \quad (7)$$

where  $\sigma, \rho, \beta$  are the parameters of the system for Prandtl number, Rayleigh number, and geometric factor.

In the experiment, the time series data is first simulated from the 3D Lorenz system. Then the derivatives of  $x, y, z$  with respect to  $t$  are calculated from the simulate data based on (6). Then the CE between the calculated derivatives and the covariates of the system equation are estimated from the simulated data using the non-parametric CE estimator [4]. Here, six covariates are considered:  $x, y, z, xy, xz, yz$  as candidates.

Two packages in R are used for the implementation of the experiments: the `nonlinearTseries` package for simulating the 3D Lorenz system and the `copent` package for estimating CE from data. In the simulation experiment, we set  $\sigma = 10, \beta = 8/3, \rho = 28$  respectively. The starting points of the simulated time series is randomly generated. The time horizon is 30 and the sample rate is 100, which leads to 3000 samples for our experiment. The default values of the hyperparameters of the implemented CE estimator were used in the experiments.

### 4.2 Results

The simulated data from the 3D Lorenz system are shown in Figure 1 and Figure 2.

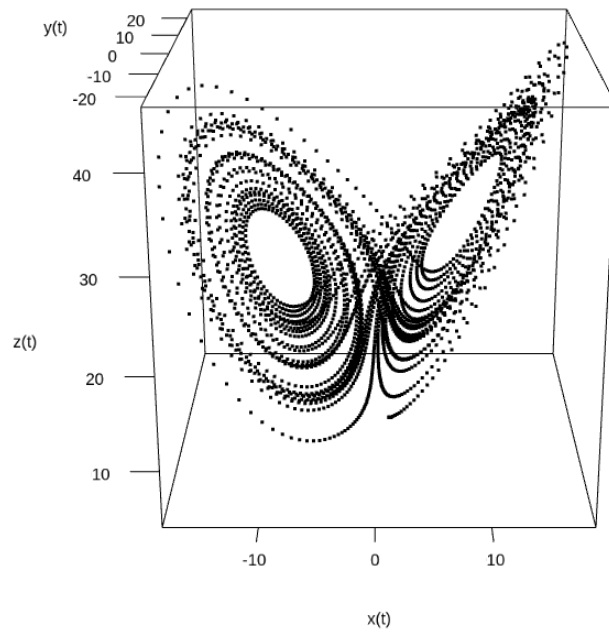


Figure 1: 3D plot of the data simulated from the 3D Lorenz system.

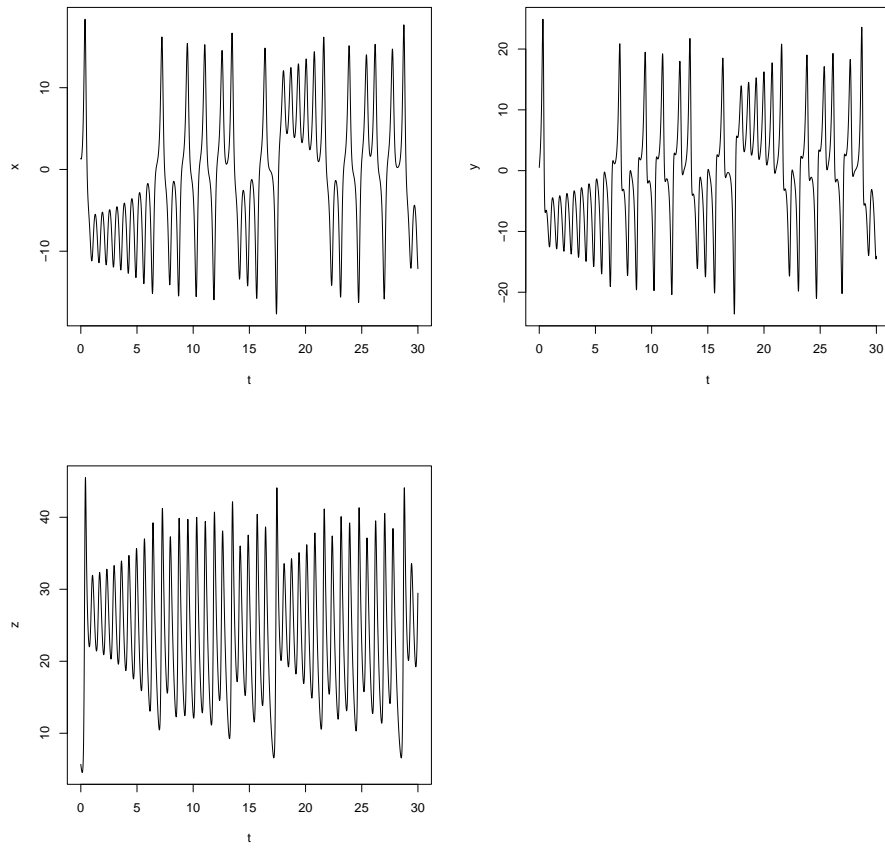


Figure 2: 2D plots of the data simulated from the 3D Lorenz system.

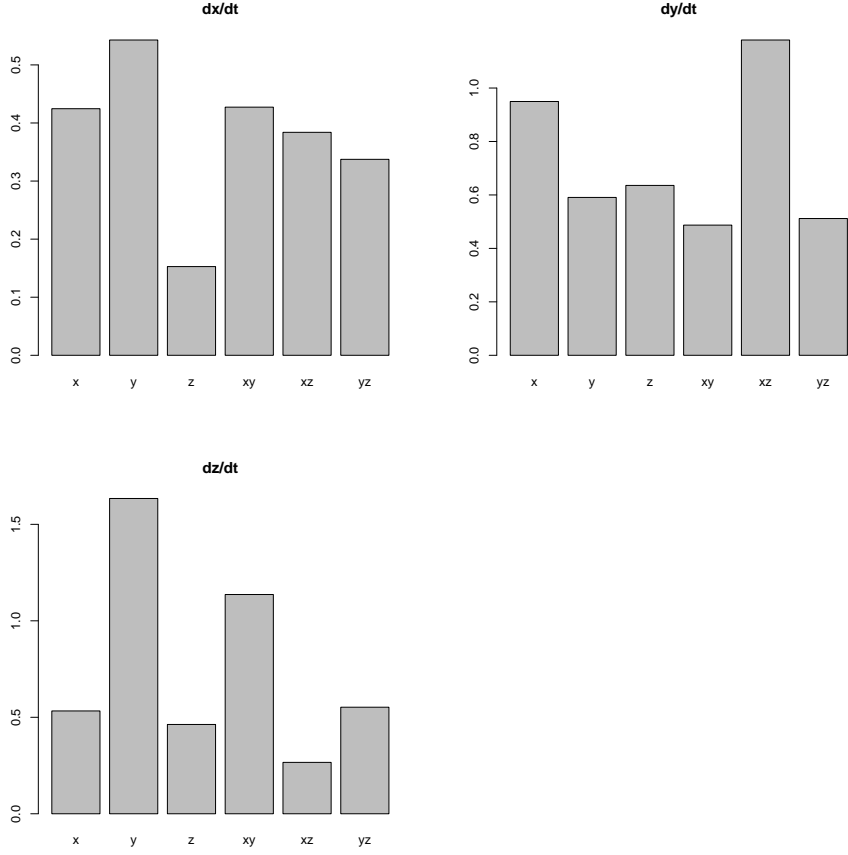


Figure 3: Estimated CEs between the candidate covariates and the derivatives of the states of the Lorenz system.

After calculating the derivatives from the simulated data, the CEs between the derivatives and the covariates are estimated as shown in Figure 3. It can be learned from it that: 1) for the derivative of  $x$ , the covariates  $x, y$  has large CE value while that of  $z$  has small CE value; 2) for the derivative of  $y$ , the covariates  $x, xz$  have large CE value; 3) for the derivative of  $z$ , the covariates  $y, xy$  have large CE values. The large CE values means the relationships between the derivatives and the covariates.

Compared the estimation results with the system equation (7), one can learn that the proposed method successfully identified the relationships between: the derivative of  $x$  and  $x, y$ , the derivative of  $y$  and  $x, xz$ , and the derivative of  $z$  and  $xy$ . The two second-order covariates in (7) are both identified from the simulated data.

## 5 Discussion

In this paper, we proposed a method for system identification based on CE. the identification problem is treated as variable selection which has been solved with CE. Since CE can be estimated non-parametrically, the proposed method is also non-parametric and model-free. In practice, almost no hyper-parameter need to be tuned in estimating CE, which makes our method advantageous than the traditional methods that can only derive the results after hyperparameters-tuning.

CE is a rigorously defined mathematical concept in information theory, which makes it can be applied to not only statistical problems but also deterministic systems. It can measure the static relationships governed by system equations. It can also measure the dynamic relationships in the systems as information flow estimated with the CE-based estimator of transfer entropy [17], even in the systems with time delay [18].

In our experiment, only the first-order covariates  $x, y, z$  and the second-order covariates  $xy, xz, yz$  are considered due to the already known system equation. It can be easily extended to the cases where the system equation is unknown by considering more high-order covariates.

When selecting covariates based on the estimated CE values, the interrelationship between covariates should be considered. For example, the covariates related to the derivative of  $x$  in our experiments include not only  $x, y$  but  $xy, xz, yz$  which are associated with high CE values. Such selection should be done carefully by considering these interrelationships between covariates.

## 6 Conclusions

In this paper we propose a method for identifying differential equation of dynamical systems with CE. The problem is considered as a variable selection problem and solved with the previously proposed CE-based method. The proposed method is composed of two components: the difference operator and the CE estimator. Since both components can be done non-parametrically, the proposed method is therefore model-free and hyperparameter-free. The simulation experiment with the 3D Lorenz system verified the effectiveness of the proposed method.

## A Code

The codes are available at <https://github.com/majianthu/sysid>.

## References

- [1] Alan A. Kaptanoglu, Lanyue Zhang, Zachary G. Nicolaou, Urban Fasel, and Steven L. Brunton. Benchmarking sparse system identification with low-dimensional chaos. *arXiv e-prints*, page arXiv:2302.10787, February 2023.



- [2] Maziar Raissi, Paris Perdikaris, and George Em Karniadakis. Machine learning of linear differential equations using Gaussian processes. *Journal of Computational Physics*, 348:683–693, 2017.
- [3] Steven L. Brunton, Joshua L. Proctor, and J. Nathan Kutz. Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proceedings of the National Academy of Sciences*, 113(15):3932–3937, 2016.
- [4] Jian Ma and Zengqi Sun. Mutual information is copula entropy. *Tsinghua Science & Technology*, 16(1):51–54, 2011. See also arXiv preprint arXiv:0808.0845, 2008.
- [5] Jian Ma. Variable selection with copula entropy. *Chinese Journal of Applied Probability and Statistics*, 37(4):405–420, 2021.
- [6] Lu Chen, Vijay P. Singh, and Shenglian Guo. Measure of correlation between river flows using the copula-entropy method. *Journal of Hydrologic Engineering*, 18(12):1591–1606, 2013.
- [7] Radko Mesiar and Ayyub Sheikhi. Nonlinear random forest classification, a copula-based approach. *Applied Sciences*, 11(15), 2021.
- [8] Yan-Ning Sun, Yu Chen, Wu-Yin Wang, Hong-Wei Xu, and Wei Qin. Modelling and prediction of injection molding process using copula entropy and multi-output SVR. In *2021 IEEE 17th International Conference on Automation Science and Engineering (CASE)*, pages 1677–1682, 2021.
- [9] Fuqiang Sun, Wendi Zhang, Ning Wang, and Wei Zhang. A copula entropy approach to dependence measurement for multiple degradation processes. *Entropy*, 21(8):724, 2019.
- [10] Peili Liu, Song Han, Na Rong, and Junqiu Fan. Frequency stability prediction of power systems using vision transformer and copula entropy. *Entropy*, 24(8):1165, 2022.
- [11] Roger B Nelsen. *An introduction to copulas*. Springer Science & Business Media, 2007.
- [12] Harry Joe. *Dependence modeling with copulas*. CRC press, 2014.
- [13] M Sklar. Fonctions de repartition an dimensions et leurs marges. *Publ. inst. statist. univ. Paris*, 8:229–231, 1959.
- [14] Thomas M Cover and Joy A Thomas. *Elements of information theory*. John Wiley & Sons, 2012.
- [15] Alexander Kraskov, Harald Stögbauer, and Peter Grassberger. Estimating mutual information. *Physical Review E*, 69(6):066138, 2004.
- [16] Edward N. Lorenz. Deterministic nonperiodic flow. *Journal of Atmospheric Sciences*, 20(2):130 – 141, 1963.
- [17] Jian Ma. Estimating transfer entropy via copula entropy. *arXiv preprint arXiv:1910.04375*, 2019.

- [18] Jian Ma. Identifying time lag in dynamical systems with copula entropy based transfer entropy. *arXiv preprint arXiv:2301.06037*, 2023.