Sparse Estimation of Dynamical System Based on Hamiltonian Mechanics

Yuya Note[1], Takaharu Yaguchi[2*], Toshiaki Omori[1†]

[1]Department of Electrical and Electronic Engineering, Graduate School of Engineering, Kobe University, Japan
[2]Department of Mathematics, Graduate School of Science, Kobe University, Japan

Abstract

In this study, we propose a new data-driven approach to estimating dynamical systems with sparse modeling by incorporating physical constraints derived from Hamiltonian mechanics. In the proposed method, we estimate governing dynamics from candidate nonlinear terms by using a sparse representation of Hamiltonian, rather than that of individual dynamical equations for coordinate and momentum. Experiments with noisy observed data show that the proposed method provides accurate parameter estimation and extraction of the necessary nonlinear terms from the candidates compared to conventional methods. Furthermore, we show that the estimation based on the energy conservation law provides superior accuracy in long-term forecasts.

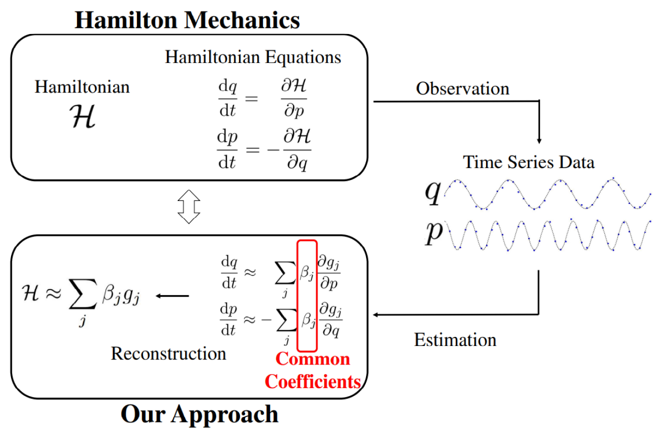

Figure 1: Overview of the proposed method.

Introduction

Recent advances in measurement and information technologies have improved both the quantity and quality of available data. Due to these developments, data-driven approaches have attracted attention as a method to extract the structure and characteristics of the underlying systems from the data. Notably, the extraction of dynamical systems from time-series data stands out as a significant challenge in diverse fields, encompassing natural sciences such as physics and neuro science, along with engineering disciplines like thermal and fluid engineering (Liu et al. 2023; Chen and Poor 2022; Yin et al. 2021; Ansari et al. 2021).

In this study, we propose a sparse modeling based on Hamiltonian mechanics. In this method, constraints derived from physical laws are imposed on the estimation. The results demonstrate that our proposed approach achieves highly interpretable and accurate estimates while maintaining physical consistency.

Related Work

Estimation of Dynamical Systems

For a data-driven approach to estimating dynamical systems, various machine learning techniques are employed including Bayesian estimation and neural network (Roda 2020; Omori et al. 2016). Bayesian estimation involves parameter estimation when the model of the dynamical system is known. Estimation using neural networks allows for estimation even when the model is unknown, but the predicted system tends to be treated as a black box.

Against this background, there is a growing emphasis on methods capable of white-box type estimation when the model is unknown. One approach involves the utilization of sparse estimation (Brunton, Proctor, and Kutz 2016). The method extracts essential bases through sparse estimation in scenarios involving a set of prepared basis functions. Consequently, it remains applicable even in cases where the model is unknown. However, in situations with numerous candidates of basis functions or noisy observation, accurate estimation of the system behavior may be compromised. To address these problems, it is important to incorporates constraints such as physical property.

Methodology

Hamiltonian Mechanics

Hamiltonian mechanics offers a framework for describing the time evolution of states of physical system that admit the energy conservation law. In Hamiltonian mechanics, governing dynamics is mathematically formulated by two indepen-

*yaguchi@pearl.kobe-u.ac.jp
†omori@eedept.kobe-u.ac.jp

dent variables: generalized coordinates $\boldsymbol{q} = [q_1, ..., q_n]$ and generalized momentum $\boldsymbol{p} = [p_1, ..., p_n]$, where n is the degree of freedom of the system. Using these variables, the system's dynamical behavior is described by the following equations:

$$\frac{\mathrm{d}q_i}{\mathrm{d}t} = \frac{\partial \mathcal{H}}{\partial p_i}, \quad \frac{\mathrm{d}p_i}{\mathrm{d}t} = -\frac{\partial \mathcal{H}}{\partial q_i} \quad \text{for} \quad i = 1, ..., n \tag{1}$$

Here, $\mathcal{H} = \mathcal{H}(\boldsymbol{p}, \boldsymbol{q}, t)$ is called the Hamiltonian, which is a function of $\boldsymbol{q}$ and $\boldsymbol{p}$ that represents the energy of the entire system. By using the Hamiltonian, the system can be expressed in a symmetric expression. Furthermore, if the Hamiltonian does not depend explicitly on time ($\mathcal{H} = \mathcal{H}(\boldsymbol{p}, \boldsymbol{q})$), the relation $\mathrm{d}\mathcal{H}/\mathrm{d}t = 0$ holds, i.e., conservation of energy is guaranteed. In addition, it is known as the Noether theorem that a symmetry of the system leads to a conservation law. The Hamilton equations includes the Newton equation and can describe broader physical phenomena in a manner consistent with the original laws of nature.

Proposed Method

Here, we develop a data driven method for estimating dynamical equations from time-series data (Figure 1). Our method imposes the physical information possessed by Hamiltonian dynamics as a constraint. By imposing the physical constraint, dynamical equations (Eq. (1)) can be expressed for all dimension i as follows:

$$\frac{\mathrm{d}q_i}{\mathrm{d}t} \approx \sum_{j=1}^{m} \beta_j \frac{\partial g_j(\boldsymbol{q}, \boldsymbol{p}, t)}{\partial p_i}, \; \frac{\mathrm{d}p_i}{\mathrm{d}t} \approx -\sum_{j=1}^{m} \beta_j \frac{\partial g_j(\boldsymbol{q}, \boldsymbol{p}, t)}{\partial q_i} \tag{2}$$

where $g_1(\boldsymbol{q}, \boldsymbol{p}, t), ..., g_m(\boldsymbol{q}, \boldsymbol{p}, t)$ are nonlinear functions that can be candidates for the m basis functions prepared in advance, and $\boldsymbol{\beta} = [\beta_1, ..., \beta_m]$ is the weight coefficient.

Note that a shared weight coefficient β_j are applied to two equations for $\mathrm{d}q_i/\mathrm{d}t$ and $\mathrm{d}p_i/\mathrm{d}t$ with subscript j, and this common property is hold for all subscript j. This commonality can be achieved by decomposing the Hamiltonian into m terms as follows:

$$\mathcal{H} \approx \sum_{j=1}^{m} \beta_j g_j(\boldsymbol{q}, \boldsymbol{p}, t) \tag{3}$$

Representing the Hamiltonian as Eq. (3), we needs to consider partially differentiated. The partial derivatives of the basis functions can be easily calculated either analytically or numerically, assuming the functions are provided.

To extract only essential nonlinear terms we propose a sparse modeling approach for Hamiltonian dynamics. We assume that the state values $\boldsymbol{q}$, $\boldsymbol{p}$ and their time derivatives $\mathrm{d}\boldsymbol{q}/\mathrm{d}t$, $\mathrm{d}\boldsymbol{p}/\mathrm{d}t$ are given as data. State values are utilized to create nonlinear basis functions constituting the approximation of the Hamiltonian, while time derivatives serve as objective variables. The objective function is as follows:

$$\mathcal{L}(\boldsymbol{\beta}) = \left\| \begin{bmatrix} \dot{\boldsymbol{q}}_{\mathrm{obs}} \\ \dot{\boldsymbol{p}}_{\mathrm{obs}} \end{bmatrix} - \sum_{j=1}^{m} \beta_j \begin{bmatrix} \frac{\partial g_j(\boldsymbol{q}_{\mathrm{obs}}, \boldsymbol{p}_{\mathrm{obs}}, t)}{\partial \boldsymbol{p}} \\ -\frac{\partial g_j(\boldsymbol{q}_{\mathrm{obs}}, \boldsymbol{p}_{\mathrm{obs}}, t)}{\partial \boldsymbol{q}} \end{bmatrix} \right\|_2^2 + \lambda \|\boldsymbol{\beta}\|_1 \tag{4}$$

where $\boldsymbol{q}_{\mathrm{obs}}$ and $\boldsymbol{p}_{\mathrm{obs}}$ represent observed values of $\boldsymbol{q}$ and $\boldsymbol{p}$, while $\dot{\boldsymbol{q}}_{\mathrm{obs}}$ $\dot{\boldsymbol{p}}_{\mathrm{obs}}$ represent observed values of $\mathrm{d}\boldsymbol{q}/\mathrm{d}t$, $\mathrm{d}\boldsymbol{p}/\mathrm{d}t$, respectively. L_1-norm regularization is introduced for the weight coefficients $\boldsymbol{\beta}$ for sparse estimation. By assuming sparse representation in Hamiltonian $\mathcal{H}(\boldsymbol{q}, \boldsymbol{p})$, we realize simultaneous estimation of dynamical equations for both $\boldsymbol{q}$ and $\boldsymbol{p}$. Furthermore, this sparsity reduces the dimension, leading to a lightweight simulation. The regularization parameter λ determines the strength of sparsity. This allows only terms related to the Hamiltonian to be extracted from a large number of basis functions, contributing to the identification of dynamics and the interpretability of the model.

Finally, we highlight the distinctions among conventional and proposed methods. The conventional approach necessitates the consideration of the following concerning Eq. (1):

$$\frac{\mathrm{d}q_i}{\mathrm{d}t} \approx \sum_{j=1}^{m} \mu_j g_j(\boldsymbol{q}, \boldsymbol{p}, t), \quad \frac{\mathrm{d}p_i}{\mathrm{d}t} \approx \sum_{j=1}^{m} \nu_j g_j(\boldsymbol{q}, \boldsymbol{p}, t) \tag{5}$$

where μ_j and ν_j are independent weight coefficients corresponding to $\boldsymbol{q}$ and $\boldsymbol{p}$, respectively. These mean that the estimation is conducted independently based only on the data, without considering the relevance of each state value. On the other hand, concerning Eq. (2), it is possible to relate the equations according to natural laws. From this, it can be argued that the proposed approach considering Hamiltonian mechanics (Eqs. (2)–(4)) can introduce regularization based on physical laws.

Results

To evaluate the performance of our proposed method, we present experimental results from two Hamiltonian dynamical systems: Duffing oscillator and Rayleigh-Bénard convection. In these experiments, the observed data was created by adding Gaussian noise to the true value of each state values and their time derivatives.

To demonstrate the superiority of the proposed method, we employed the least-squares method without regularization (LSM) and the least-squares method with L_1-norm regularization (LASSO) as comparative methods. These methods do not consider Hamiltonian mechanics and utilize the time derivatives of the state variables as the objective variables. The evaluation criteria consist of assessing the accuracy of estimated weight coefficients and comparing the accuracy of the long-term forecasts. In order to achieve optimal sparsity, the regularization parameters were automatically selected to have the best generalization performance in the proposed method and LASSO.

Duffing Oscillator

In the first experiment, the Duffing oscillator was examined. We consider the Hamiltonian expressed as follows:

$$\mathcal{H} = \frac{1}{2}p^2 + \frac{1}{2}aq^2 + \frac{1}{4}bq^4 \tag{6}$$

where we consider the system does not have a dissipative term and hence conserves energy in this experiment. For the observed values, the data were prepared at 0.001 time intervals from $t = 0$ to 20. The first 25% of the data points were

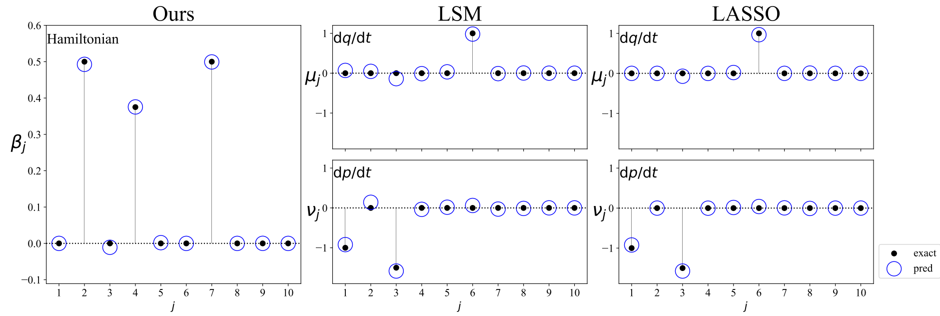

Figure 2: Estimated coefficients in Duffing oscillator. Common coefficients $\boldsymbol{\beta}$ estimated by proposed method (left) and coefficients $\boldsymbol{\mu}$ and $\boldsymbol{\nu}$ estimated separately for $\mathrm{d}q/\mathrm{d}t$ by conventional methods (upper figure) and $\mathrm{d}p/\mathrm{d}t$ (lower figure), respectively. Horizontal axis denotes basis function type, and vertical axis shows coefficient values. True values (black dots) and predicted values (blue circles) are shown.

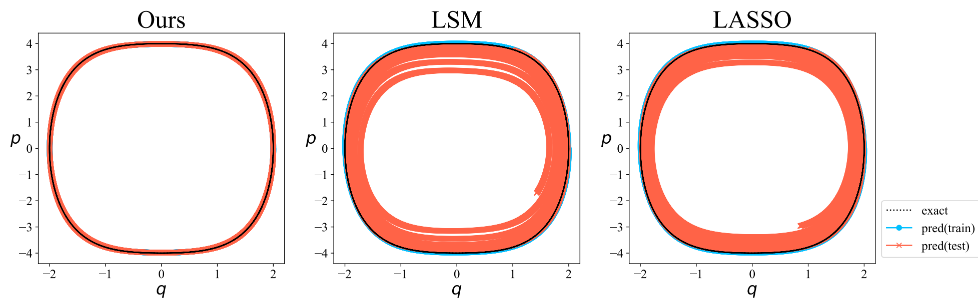

Figure 3: Comparison of the long-term forecasts in Duffing oscillator. True trajectory (black dashed line) and predicted ones for the training period (blue dots) and the test period (red crosses) are shown.

used for training, whereas the remaining 75% were used as the test set. The 10 candidate functions are as follows:

$$\{q^u, p^u \mid u \in \mathbb{N}, 1 \le u \le 5\}$$

The results of estimated coefficients of the basis functions for each method are shown in Figure 2. Next, the results of the long-term forecasts are shown in Figure 3. Mean squared error for the test portions are also shown in Table 1. Details of the results will be mentioned in Discussion.

Rayleigh-Bénard Convection

In the second experiment, Rayleigh-Bénard convection was examined. Here, the following Hamiltonian is considered:

$$\mathcal{H} = \frac{A}{k}\sin(kp)\sin(\pi q) \tag{7}$$

where we consider energy-stable system in this experiment. For the observed values, the data were prepared at 0.001 time intervals from $t = 0$ to 19.2. The first half of the data points were used for training and the other as a test set. The following 54 candidates for basis functions were prepared:

$$\begin{aligned}&\{\sin(\omega t)\sin(ukp)\sin(\pi q), \cos(\omega t)\sin(ukp)\sin(\pi q),\\&\sin(\omega t)\cos(ukp)\sin(\pi q), \cos(\omega t)\cos(ukp)\sin(\pi q),\\&\sin(\omega t)\sin(ukp)\cos(\pi q), \cos(\omega t)\sin(ukp)\cos(\pi q),\\&\sin(\omega t)\cos(ukp)\cos(\pi q), \cos(\omega t)\cos(ukp)\cos(\pi q),\\&\sin(\omega t), \sin(\omega t)p, \sin(\omega t)q, \cos(\omega t), \cos(\omega t)p, \cos(\omega t)q,\\&\sin(ukp)\sin(\pi q), \cos(ukp)\sin(\pi q), \sin(ukp)\cos(\pi q),\\&\cos(ukp)\cos(\pi q) \mid u \in \mathbb{N}, 1 \le u \le 4, \omega = 2\pi/T\}\end{aligned}$$

The results of estimating the coefficients of the basis functions for each method are shown in Figure 4. Next, the results of the long-term forecasts comparison are shown in Figure 5. Mean squared error for the test portions are also shown in Table 1. Details of the results will be mentioned in Discussion.

Discussion

As shown in Figures 2 and 4, the proposed method estimates both true nonzero and true zero coefficients with high accuracy. As shown in Figures 3 and 5, we find that the proposed method accurately estimate the true trajectory, whereas the exisiting methods show large deviation from true ones. Quantitive evaluation using mean squared error for test data (Table 1) also shows high accuracy of the proposed method.

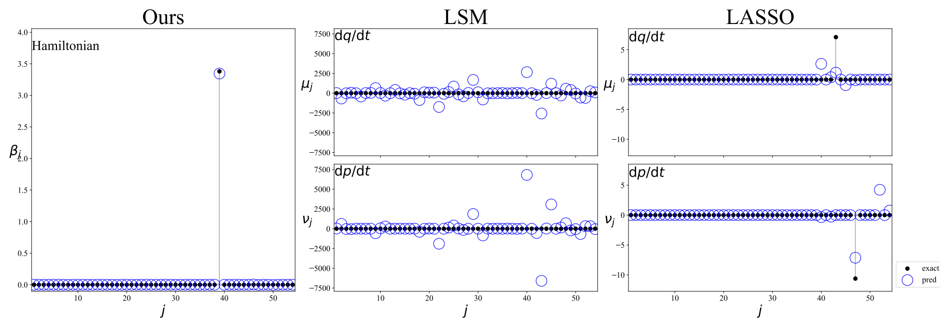

Figure 4: Estimated coefficients in Rayleigh-Bénard convection: See the caption for Figure 2.

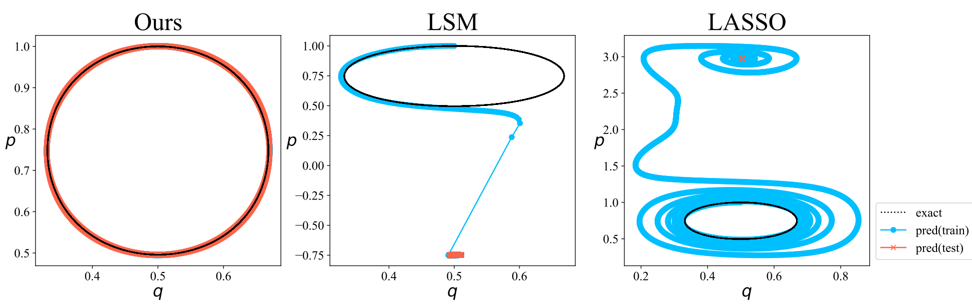

Figure 5: Comparison of the long-term forecasts in Rayleigh-Bénard convection: See the caption for Figure 3.

Table 1: Comparison of MSE for test data

	Duffing oscillator	Rayleigh-Bénard convection
Ours	$\mathbf{4.99 \times 10^{-2}}$	$\mathbf{1.17 \times 10^{-1}}$
LSM	4.42	2.29
LASSO	2.58	5.00

The results for the Duffing oscillator (Figures 2 and 3) show that the proposed method excels in both the coefficients' estimation and the long-term forecasts. The conventional methods have shown a small discrepancy in estimated coefficients, but a significant impact on long-term forecasts. This suggests that the proposed method achieves stable learning, where the behavior does not change even for long-term forecasts, i.e., that energy is conserved.

The results for the Rayleigh-Bénard convection (Figures 4 and 5) show that, even when there are many candidates, the proposed method outperforms the conventional method in the both aspects. This suggests that the symmetry and the conservation law of the Hamilton equations prevent the selection of basis functions that differ from the true system, even when there are many candidates for the terms.

Conclusion

Our proposed method has demonstrated superior accuracy in long-term forecasts and parameters estimation compared to conventional methods in experiments on noisy data. Our approach achieves more physically consistent predictions for dynamical systems by introducing the energy conservation law and symmetries based on Hamiltonian mechanics into sparse modeling.

Acknowledgments

This work was partially supported by Grant-in-Aid for Scientific Research (B) (No. JP21H03509), and a Fund for the Promotion of Joint International Research (International Collaborative Research) (No. JP23KK0184), MEXT, Japan, CREST (No. JPMJCR1914), JST, Japan.

References

Ansari, A. F.; Benidis, K.; Kurle, R.; Turkmen, A. C.; Soh, H.; Smola, A. J.; Wang, B.; and Januschowski, T. 2021. Deep explicit duration switching models for time series. *NeurIPS*, 34: 29949–29961.

Brunton, S. L.; Proctor, J. L.; and Kutz, J. N. 2016. Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proc. Nat. Acad. Sci.*, 113(15): 3932–3937.

Chen, Y.; and Poor, H. V. 2022. Learning mixtures of linear dynamical systems. *Proc. 39th ICML*, 162: 3507–3557.

Liu, Y.; Magliacane, S.; Kofinas, M.; and Gavves, E. 2023. Graph switching dynamical systems. 202: 21867–21883.

Omori, T.; Kuwatani, T.; Okamoto, A.; and Hukushima, K. 2016. Bayesian inversion analysis of nonlinear dynamics in surface heterogeneous reactions. *Phys. Rev. E*, 94: 033305.

Roda, W. C. 2020. Bayesian inference for dynamical systems. *Infectious Disease Modelling*, 5: 221–232.

Yin, Y.; Ayed, I.; de Bézenac, E.; Baskiotis, N.; and Gallinari, P. 2021. LEADS: Learning Dynamical Systems that Generalize Across Environments. *NeurIPS*, 34: 7561–7573.