What's Next?
Predicting Hamiltonian Dynamics from Discrete Observations of a Vector Field

Zi-Yu Khoo[1], Delong Zhang[1], and Stéphane Bressan[1]

National University of Singapore. 21 Lower Kent Ridge Rd, Singapore 119077
{khoozy,zhangdel,steph}@comp.nus.edu.sg

Abstract. We present several methods for predicting the dynamics of Hamiltonian systems from discrete observations of their vector field. Each method is either informed or uninformed of the Hamiltonian property. We empirically and comparatively evaluate the methods and observe that information that the system is Hamiltonian can be effectively informed, and that different methods strike different trade-offs between efficiency and effectiveness for different dynamical systems.

Keywords: trajectory prediction · physics-informed neural network · data analysis · inductive bias · learning bias.

1 Introduction

The prediction of the dynamics of systems is a relevant and crucial task to many applications in domains ranging from the hard to social sciences. The dynamics of a system is captured by the vector field formed by the time derivatives of the variables describing the system's current state. The system evolves along flow lines in the state space. The flow map is the function that, given an initial state and a time interval, outputs the state of the system after the time interval.

A Hamiltonian system [10] is a dynamical system governed by Hamilton's equations. The Hamiltonian property indicates the conservation of some quantity, typically the energy in mechanical and physical systems.

We design, present, and evaluate physics-informed methods for the prediction of the dynamics of a system from the observation of its vector field at discrete locations of the state space. We want to understand and quantify the significance of informing regression and integration of a vector field with physics information. In the spirit of an ablation study, we compare variants of the general method and different devices, a multilayer perceptron and a Gaussian process, with and without the information that the system is Hamiltonian. In the first stage, this concerns whether the vector field is under the constraints of Hamilton's equations. In the second stage, this concerns whether the vector field is integrated with a non-symplectic or symplectic integrator. The empirical comparative performance evaluation is conducted with data from several physical systems of an oscillator, a pendulum, a Henon Heiles system, a Morse potential model

of a dyatomic molecule, and several abstract systems with logarithmic, inverse trigonometric, exponential, radical and polynomial Hamiltonian functions.

2   Related Work

We are interested in the problem of learning the vector field and the flow map from samples of the vector field. There are multiple statistical methods to learn correlated vector-valued functions [9,14,16]. Learning vector fields using machine learning has been addressed as multiple output regression by Hastie et al. [8] while state of the art works use neural networks that model ordinary differential equations [12] and partial differential equations [13]. Learning vector fields using kernel methods with regularization was introduced by Micchelli et al. [11]. Similarly, Boyle and Frean [4] introduced Gaussian processes to learn vector-valued functions. Recent works regularize the vector field [1].

Our work is similar to that of Greydanus et al. [6], who showed that a physics-informed neural network, similar to Bertalan et al.'s [2], can faster learn the Hamiltonian of mechanical systems and better predict their dynamics from selected samples of their vector fields than a neural network. Additionally, following Chen et al. [5], a symplectic integrator [7,15] can integrate the learned Hamiltonian vector field to predict a flow line.

3   Methodology

The general method for predicting the dynamics of a system from the observation of its vector field at discrete locations of its phase space comprises two successive stages: the learning or regression of the vector field from the samples, and the integration of the vector field into the flow map image of a state in the phase space for a prescribed time interval. We consider four variants of the general methods. They result from the obliviousness or awareness of information that a system is Hamiltonian during the first and second stages of the general method.

We consider two non-linear regression devices for the learning of a surrogate of the vector field, a multilayer perceptron neural network and a Gaussian process. The two devices are chosen as the main representatives of parametric and non-parametric non-linear regression statistical machine learning devices. They can learn a surrogate of the vector field, or learn a surrogate of the Hamiltonian and compute a surrogate of the vector field with automatic differentiation.

We consider the supervised learning of a surrogate $\hat{F}$ of the vector field F with two physics oblivious devices: a multilayer perceptron and a Gaussian process. They regress the vector-valued function oblivious to physics information. A training data set Z comprises N samples, $\frac{dx}{dt}$ and $\frac{dy}{dt}$, of the vector field for N states $z = (x, y)$ in the phase space of the system studied. When the surrogate is a multilayer perceptron, the weights minimise the mean squared error between the approximated vector field and the ground truth vector field. When the surrogate is a Gaussian process, the conditional expectation of the Gaussian process for the approximated vector field and the ground truth vector field is maximised.

We consider learning a physics-informed surrogate $\hat{H}$ of the Hamiltonian H learned from the same training data set Z under the constraints that the system is Hamiltonian before deriving the vector field. We adapt the method proposed by [2]. We use the constraints of Equation 1 to define the loss function of the device. f_0 is is an arbitrary pinning term. f_1 and f_2 are Hamilton's equations.

$$f_0 = \left(\hat{H}(x_0, y_0) - H_0\right)^2 \qquad f_1 = \left(\frac{\partial \hat{H}}{\partial y} - \frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 \qquad f_2 = \left(\frac{\partial \hat{H}}{\partial x} + \frac{\mathrm{d}y}{\mathrm{d}t}\right)^2 \quad (1)$$

When the surrogate for the Hamiltonian is a multilayer perceptron neural network, the loss function is a linear combination of f_0, f_1 and f_2. When the surrogate is a Gaussian process, constraining its loss function leads to solving Equation 2. In both cases, the derivative of the Hamiltonian at any new state of the phase space can be obtained from the surrogate by automatic differentiation.

$$\begin{bmatrix} \frac{\partial}{\partial z_1} k(z_1, Z)^\top k(Z, Z')^{-1} \\ \vdots \\ \frac{\partial}{\partial z_3} k(z_N, Z)^\top k(Z, Z')^{-1} \\ k(z_0, Z)^\top k(Z, Z')^{-1} \end{bmatrix} [H(Z)] = \begin{bmatrix} g(Z) \\ H_0 \end{bmatrix} \qquad (2)$$

The flow map for predicting the dynamics of the system is computed by integrating the surrogate vector fields. Ignoring that the system is Hamiltonian, one can use the first order explicit Euler integrator [3]. Knowledge of the Hamiltonian system allows use of the implicit symplectic Euler integrator [7].

4   Performance Evaluation

Two experiments are conducted. In the first, we empirically compare the performance of the two physics-oblivious methods learning the vector field directly and of their two physics-informed counterparts learning the Hamiltonian. We use a testing data set of 20^{2n} vectors at evenly spaced states in the phase space for each system. Effectiveness is measured by mean squared error between the ground truth vectors and approximated vectors, and efficiency by the time taken for early stopping of the multilayer perceptron, or the time taken to fit a Gaussian process. For the second experiment, we evaluate the prediction of the dynamics of each Hamiltonian dynamical system by computing the flow map over the interpolated vector field. The vector field are learned with physics-informed methods in the first experiment, and combined with the Euler or symplectic Euler integrator. We use a testing data set of 5^{2n} evenly spaced states in the phase space for each system. Flow lines are calculated from the differential equations of the system with a symplectic Euler integrator for 50 time steps, with step size $h = 0.1$. The mean squared error between the ground truth flow line and the predicted flow line for each method, and prediction time is computed.

For both experiments, all multilayer perceptrons and physics-informed multilayer perceptrons set aside 20% of the training data set for validation based

early stopping.Other settings follow Bertalan et al. [2]. All surrogates are trained in Google Colaboratory[1]. Training data set of size between 64 and 1024 are sampled uniformly at random from the vector fields of the eight example systems, respectively. The experiments are repeated for 20 unique random seeds, and the mean values are reported and compared. Results are plotted as Pareto plots.
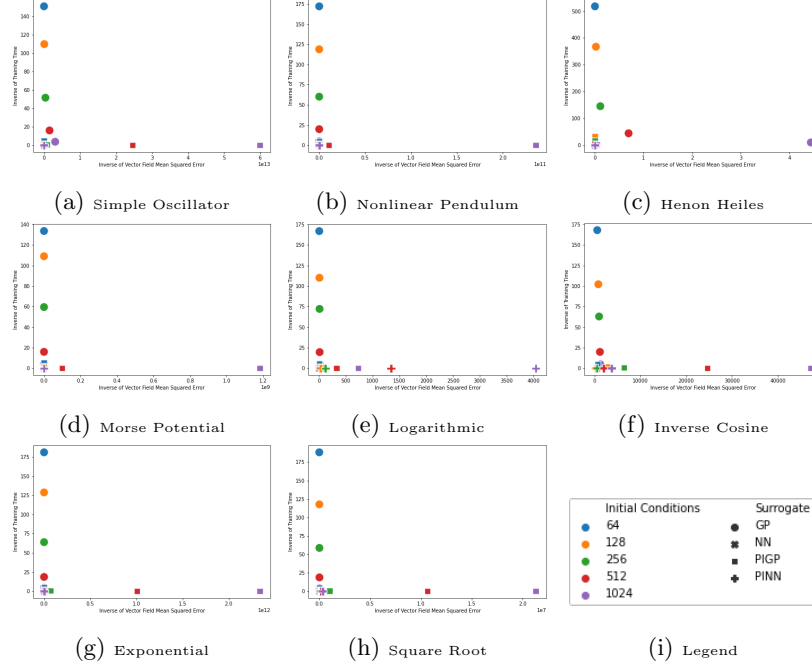

Fig. 1: Learning the vector field

Figure 1 plots the Pareto plot for the inverse of the vector field approximation mean squared error (x axis) and the inverse of surrogate training time (y axis) for the different dynamical system indicated. The colors correspond to the different training data set of varying sizes. The circle, cross, square and plus symbols represent the Gaussian process (GP), multilayer perceptron (NN), physics-informed Gaussian process (PIGP) and physics-informed multilayer perceptrons (PINN) methods respectively. The most efficient method is the physics-informed Gaussian process, while multilayer perceptron can better extrapolate the vector field.

Figure 2 compares inverse prediction error (x axis) and inverse prediction time (y axis). The colors correspond to training data set of varying sizes. The right-pointing triangle, square, down-pointing triangle and empty circle symbols represent the Gaussian process (GPE), physics-informed Gaussian process

---

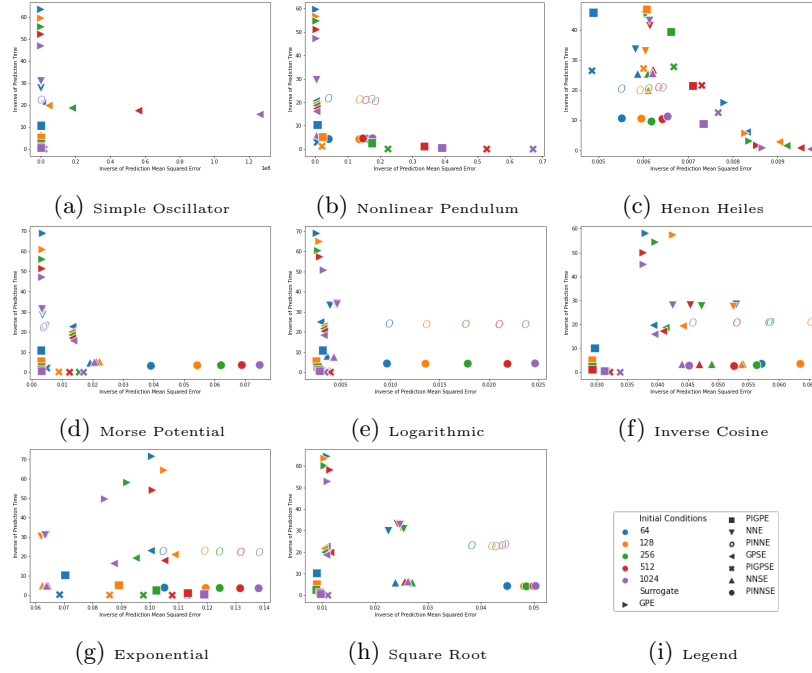[1] Find code and results at github.com/zykhoo/predicting_hamiltonian_dynamics.

Fig. 2: Predicting the flow map

(PIGPE), multilayer perceptron (NNE), and physics-informed multilayer perceptron (PINNE), all with the Euler integrator. The left-pointing triangle, cross, up-pointing triangle and filled circle symbols represent the same surrogates with symplectic Euler integrator. They learn and integrate the vector field. The most efficient method for the prediction of the dynamics is the physics-informed multilayer perceptron with symplectic integrator. This demonstrates the advantage of informing the methods of the Hamiltonian nature of the dynamical systems

5   Conclusion

We design, present, and evaluate physics-informed methods for the prediction of the dynamics of a system from the observation of its vector field at discrete locations of the state space. We show that information that the system is Hamiltonian can be effectively informed in both the regression and integration of the vector field. The methods strike trade-offs between efficiency and effectiveness.

Acknowledgement

References

1. Baldassarre, L., Rosasco, L., Barla, A., Verri, A.: Vector field learning via spectral filtering. In: Balcázar, J.L., Bonchi, F., Gionis, A., Sebag, M. (eds.) Machine Learning and Knowledge Discovery in Databases. pp. 56–71. Springer Berlin Heidelberg, Berlin, Heidelberg (2010)
2. Bertalan, T., Dietrich, F., Mezić, I., Kevrekidis, I.G.: On learning hamiltonian systems from data. Chaos: An Interdisciplinary Journal of Nonlinear Science **29**(12), 121107 (Dec 2019)
3. Boyce, W.E., DiPrima, R.C., Meade, D.B.: Elementary Differential Equations and Boundary Value Problems. John Wiley & Sons, 11 edn. (2017)
4. Boyle, P., Frean, M.: Dependent gaussian processes. In: Saul, L., Weiss, Y., Bottou, L. (eds.) Advances in Neural Information Processing Systems. vol. 17. MIT Press (2004)
5. Chen, Z., Zhang, J., Arjovsky, M., Bottou, L.: Symplectic recurrent neural networks. In: International Conference on Learning Representations (2020)
6. Greydanus, S., Dzamba, M., Yosinski, J.: Hamiltonian neural networks. In: Wallach, H., Larochelle, H., Beygelzimer, A., d'Alché-Buc, F., Fox, E., Garnett, R. (eds.) Advances in Neural Information Processing Systems. vol. 32. Curran Associates, Inc. (2019)
7. Hairer, E., Wanner, G., Lubich, C.: Symplectic Integration of Hamiltonian Systems, pp. 179–263. Springer Berlin Heidelberg, Berlin, Heidelberg (2002)
8. Hastie, T., Tibshirani, R., Friedman, J.: The elements of statistical learning: data mining, inference and prediction. Springer, 2 edn. (2009)
9. Izenman, A.J.: Reduced-rank regression for the multivariate linear model. Journal of Multivariate Analysis **5**(2), 248–264 (1975)
10. Meyer, K.R., Hall, G.R.: Hamiltonian Differential Equations and the N-Body Problem, pp. 1–32. Springer New York, New York, NY (1992)
11. Micchelli, C.A., Pontil, M.: On learning vector-valued functions. Neural Comput **17**(1), 177–204 (Jan 2005)
12. Raissi, M.: Deep hidden physics models: Deep learning of nonlinear partial differential equations. J. Mach. Learn. Res. **19**(1), 932–955 (jan 2018)
13. Ruthotto, L., Haber, E.: Deep neural networks motivated by partial differential equations. Journal of Mathematical Imaging and Vision **62**(3), 352–364 (Apr 2020)
14. Van Der Merwe, A., Zidek, J.: Multivariate regression analysis and canonical variates. Canadian Journal of Statistics **8**(1), 27–39 (1980)
15. de Vogelaere, R.: Methods of integration which preserve the contact transformation property of the hamilton equations (1956)
16. Wold, S., Ruhe, A., Wold, H., Dunn, III, W.J.: The collinearity problem in linear regression. the partial least squares (pls) approach to generalized inverses. SIAM Journal on Scientific and Statistical Computing **5**(3), 735–743 (1984)