DOI: 10.11576/dataninja-1162

Finding Commonalities in Dynamical Systems with Gaussian Processes

Andreas Besginow* ANDREAS.BESGINOW@TH-OWL.DE
Technische Hochschule Ostwestfalen-Lippe, Germany

Jan David Hüwel* JAN.HUEWEL@FERNUNI-HAGEN.DE
University in Hagen, Germany

Markus Lange-Hegermann MARKUS.LANGE-HEGERMANN@TH-OWL.DE
Technische Hochschule Ostwestfalen-Lippe, Germany

Christian Beecks CHRISTIAN.BEECKS@FERNUNI-HAGEN.DE
University in Hagen, Germany

Abstract

Gaussian processes can be utilized in the area of equation discovery to identify differential equations describing the physical processes present in time series data. Furthermore, automatically constructed models can be split into components that facilitate comparisons between time series on a structural level. We consider the potential combination of these two methods and describe how they could be used to detect shared physical properties in multiple recordings of dynamical systems as time series. This approach provides insights into the underlying dynamics of the observed systems, facilitating a deeper understanding of complex processes.

Keywords: Gaussian Process, Dynamical Systems, Frequent Itemset Mining, Equation Discovery

1. Gaussian Processes (GPs)

Formally, a GP $g(x) = \mathcal{GP}(\mu(x), k(x, x'))$ defines a probability distribution over the space of functions $\mathbb{R}^d \to \mathbb{R}^\ell$, such that the outputs $g(x_i)$ at any set of inputs $x_i \in \mathbb{R}^d$ are jointly Gaussian [1]. Such a (multi-input multi-output) GP is defined by its mean function (often set to zero for the prior)

$$\mu : \mathbb{R}^d \to \mathbb{R}^\ell : x \mapsto \mathbb{E}(g(x))$$

and its (multi-input multi-output) positive semidefinite covariance function (also called kernel)

$$k : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^\ell_{\succeq 0} : \\ (x, x') \mapsto \mathbb{E}((g(x) - \mu(x))(g(x') - \mu(x'))^T).$$

where (X, y) with $X \in \mathbb{R}^{n \times d}$ and $y \in \mathbb{R}^{n \times \ell}$ are a dataset with n observations. The kernel defines the general form of the GP and usually contains additional hyperparameters, like a lengthscale or the length of a periodic pattern [2]. These hyperparameters directly impact the calculation of the covariance matrix K of observation locations X.

2. CATGP

A GP's kernel encodes prior assumptions about the data, like smoothness or periodicity. The choice of kernel greatly impacts the model's performance, as a strong correlation between these prior assumptions and the actual data allows the model to generate more accurate predictions and fit the data with more precision. Conversely, automatic kernel searches can find a descriptive kernel for a given dataset by evaluating the model performance of different kernels [2, 3]. For complex structures in the data, the best kernel is often a sum or product of simple kernels. In such cases, sums can be interpreted as modelling independent subprocesses which make up the dataset.

Recently, the Component Analysis in Time Series with Gaussian Processes (CATGP) has emerged as a way to use this principle for further analysis of GP models [4]. This algorithm finds commonly appearing kernel components (subkernels) in a collection of GPs by interpreting kernels as sets of such components and applying frequent itemset mining. In the previous publications about this method, the Apriori algorithm [5] was used as a basis, but the same prin-

* These authors contributed equally

ciple can be applied to any frequent itemset mining algorithm.

In this work we outline a potential combination of this method with kernels with inductive bias on systems of differential equations to infer knowledge from data with potential dynamical systems behaviour.

3. LODE-GPs

Consider a system of linear homogenous ordinary differential equations with constant coefficients

$$A \cdot \mathbf{f}(t) = 0 \tag{1}$$

with operator matrix $A \in \mathbb{R}[\partial_t]^{m \times n}$ determining the relationship between the smooth functions $f_i(t) \in C^\infty(\mathbb{R}, \mathbb{R})$ of $\mathbf{f}(t) = \begin{pmatrix} f_1(t) & \dots & f_n(t) \end{pmatrix}^T$. For such systems the main result of [6] holds:

Theorem 1 *(LODE-GPs) For every system as in Equation* (1) *there exists a GP g, such that the set of realizations of g is dense in the set of solutions of $A \cdot \mathbf{f}(t) = 0$.*

As the authors of [6] demonstrate, these LODE-GPs can be constructed algorithmically and are guaranteed to satisfy the original system of linear homogenous system of ordinary differential equations with constant coefficients given by $A \cdot \mathbf{f}(t) = 0$. This, combined with the previously described kernel search approach, enables users to find a fitting system of differential equations for a given dataset.

4. Proposed Method

We propose a combination of these two methods. The resulting process is depicted in Figure 1. While the system in the real world can not be directly examined to identify process components, it's reasonable to assume that these components correspond to time series components in a suitable decomposition. We achieve this decomposition by employing GP models, which adapt to structures present in the data [2], categorizing subkernels by the ordinary differential equation that they are based on, and create analyzable sets of kernels representing present structures in each dataset.

The intent of our method falls under the class of equation discovery methods, which try to discover a fitting dynamic system description for a given dataset [7–9]. Where other works make use of learning this behaviour through direct GP regression for a whole dataset, we propose to find the *most frequently occurring* differential equations as follows.

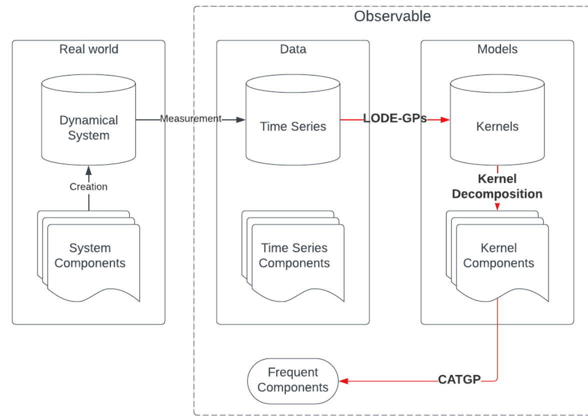

Figure 1: The correlation between different aspects in our proposed method. The red connections symbolize our contribution. Overall, process components correspond to time series components in the recorded data, which in turn correspond to kernel components of descriptive GP kernels. The figure is based on a figure in [4].

The intended objects of analysis for this method are systems, that accumulate multiple subprocesses, that follow physical equations, where the exact correlation between these subprocesses is unknown to the user. To generate insights into such physical systems, we first select descriptive LODE-GPs for time series recordings of those systems. The selected kernels are additive combinations of kernels, that correspond to systems of ordinary differential equations. Thus, each kernel can be equated to a set of kernel components, which can in turn be analysed via CATGP. This analysis identifies the most frequent types of differential equations that appear as subprocesses in the data.

Acknowledgments

This research was supported by the research training group "Dataninja" (Trustworthy AI for Seamless Problem Solving: Next Generation Intelligence Joins Robust Data Analysis) funded by the German federal state of North Rhine-Westphalia.

References

[1] Carl Edward Rasmussen and Christopher K Williams. *Gaussian processes for machine learning*. MIT press Cambridge, MA, 2006.

[2] David Duvenaud, James Lloyd, Roger Grosse, Joshua Tenenbaum, and Ghahramani Zoubin. Structure discovery in nonparametric regression through compositional kernel search. In *International Conference on Machine Learning*, pages 1166–1174. PMLR, 2013.

[3] Fabian Berns, Jan David Hüwel, and Christian Beecks. Automated model inference for gaussian processes: An overview of state-of-the-art methods and algorithms. *SN Comput. Sci.*, 3(4):300, 2022.

[4] Jan David Hüwel and Christian Beecks. Frequent component analysis for large time series databases with gaussian processes. In *EDBT*, pages 617–622. OpenProceedings.org, 2024.

[5] Rakesh Agrawal, Ramakrishnan Srikant, et al. Fast algorithms for mining association rules. In *Proc. 20th int. conf. very large data bases, VLDB*, volume 1215, pages 487–499. Santiago, Chile, 1994.

[6] Andreas Besginow and Markus Lange-Hegermann. Constraining gaussian processes to systems of linear ordinary differential equations. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 35, 2022.

[7] Kevin Egan, Weizhen Li, and Rui Carvalho. Automatically discovering ordinary differential equations from data with sparse regression. *Communications Physics*, 7(1), 2024.

[8] Yi-Chen Zhu, Paul Gardner, Ramon Fuentes, David J Wagg, Elizabeth J Cross, and Robert J Barthorpe. On robust equation discovery: a sparse bayesian and gaussian process approach. *ISMA Conference on Noise and Vibration Engineering and International Conference on Uncertainty in Structural Dynamics*, 2020.

[9] Yi-Chen Zhu, Paul Gardner, David J Wagg, Robert J Barthorpe, Elizabeth J Cross, and Ramon Fuentes. Robust equation discovery considering model discrepancy: A sparse bayesian and gaussian process approach. *Mechanical Systems and Signal Processing*, 168:108717, 2022.