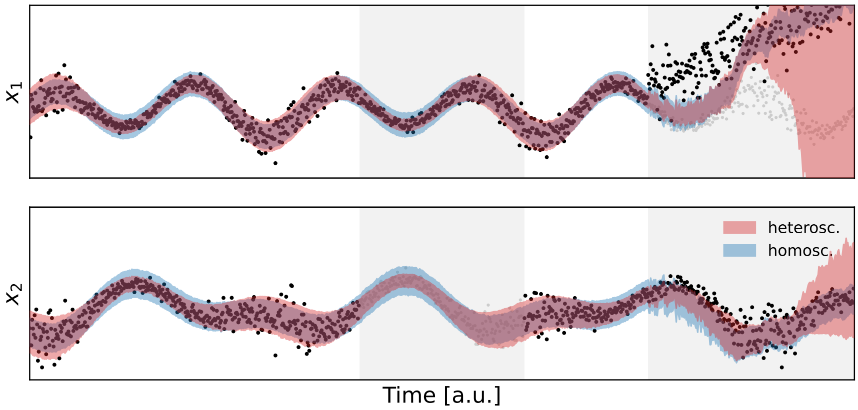
Modelling heteroscedasticity in state-space models

Christian Donner

June 2022



Project description

Motivation State-space models (SSMs) are a well-known tool, to model time-series data. Already in the 1960's the famous *Kalman filter* has been proposed [1], and which has the advantage, that it models time-series data as a single trajectory , but a full (Gaussian) density thereof. Since then many extensions have been proposed to this frame work, such that the model class can be incorporate more complex (non-linear) models, using kernels [2], Gaussian processes [3], or Deep Neural Nets [4].

However, when modelling time-series data with uncertainties, in many cases it is important to model the changes in uncertainty. E.g. solar radiation is more variable in summer than in winter, neuronal activity displays higher fluctuations when eyes are closed in contrast to eyes open, etc. These data are called *heteroscedastic* (in contrast to homoscedastic, where noise is independent of the system's state). Modelling heteroscedasticity in SSMs is quite challenging, but there have been attempts [5, 6, 7]. However, all these approaches are either not designed for multivariate data, or they do not propagate properly uncertainty through time.

Problem statement Recently, we developed a model at the SDSC, which allows modelling multivariate heteroscedastic time-series data (see Figure). We developed an algorithm, that can also learn the model parameters given the data. And most importantly, we can fully propagate uncertainty through time.

However, all this comes at the cost, that the model relatively rigid so far, and we use only very simple functions to model heteroscedastic uncertainty, and hence, the model is not as flexible as e.g. Neural Network competitors. The aim of the project is to introduce one more layer of complexity to increase flexibility of the existing model class.

Student's task The first tasks of the student are to learn the basics of state-space modelling, and the current status of the project. Also she has to familiarize herself with the existing code. Then, together with the supervisors, the extensions of the heteroscedastic models should be developed and implemented. Also it would be nice to establish state-of-the-art baselines within the frame of the project.

Additional information

- **Difficulty of the project:** Learning the basics, getting familiar with existing code (you won't be on your own)
- **What will you learn?** Probabilistic time-series modelling, object-oriented programming in Python, `jax` library
- **Requirements:** Machine Learning fundamentals, basic math skills, probabilistic modelling fundamentals, good Python skills, experience with git.
- **Supervisors:** Dr. Christian Donner (christian.donner@sdsc.ethz.ch).

References

[1] R. E. Kalman. A New Approach to Linear Filtering and Prediction Problems. *Journal of Basic Engineering*, 82(1):35–45, 3 1960.

[2] Sam Roweis and Zoubin Ghahramani. A Unifying Review of Linear Gaussian Models. *Neural Computation*, 11(2):305–345, 2 1999.

[3] Roger Frigola, Yutian Chen, and Carl E Rasmussen. Variational Gaussian process state-space models. In *Advances in Neural Information Processing Systems*, volume 4, pages 3680–3688, 2014.

[4] Emmanuel de Bézenac, Syama Sundar Rangapuram, Konstantinos Benidis, Michael Bohlke-Schneider, Richard Kurle, Lorenzo Stella, Hilaf Hasson, Patrick Gallinari, and Tim Januschowski. Normalizing Kalman filters for multivariate time series analysis. In *Advances in Neural Information Processing Systems*, volume 2020-Decem, 2020.

[5] Kristian Kersting, Christian Plagemann, Patrick Pfaff, and Wolfram Burgard. Most likely heteroscedastic Gaussian process regression. In *Proceedings of the 24th international conference on Machine learning - ICML '07*, pages 393–400, New York, New York, USA, 2007. ACM Press.

[6] Emmanouil A. Platanios and Sotirios P. Chatzis. Gaussian Process-Mixture Conditional Heteroscedasticity. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 36(5):888–900, 5 2014.

[7] Mingliang Wang, Elling W. Jacobsen, Veronique Chotteau, and Hakan Hjalmarsson. Estimation of Heteroscedastic Multilinear Systems. In *2020 59th IEEE Conference on Decision and Control (CDC)*, pages 2875–2880. IEEE, 12 2020.