Active learning-assisted neutron spectroscopy with log-Gaussian processes

Mario Teixeira Parente
(Department of Statistics, LMU Munich)

08.11.2023, 16.00 (c.t.)

Department of Statistics, Ludwigstr. 33, Room 144
and online via Zoom (Link)
(Meeting-ID: 913-2473-4411; Password: StatsCol22)

From a methodological perspective, experiments at three-axes spectrometers (TAS) collect noisy observations of a two-dimensional intensity distribution to investigate a material of interest. If the intensity distribution is only partly or even fully unknown, experimenters usually decide manually where to place measurement points for a rapid overview. Active learning methods can assist this process by choosing informative measurement locations while taking instrument costs into account and hence optimize the use of beam time. For example, avoiding measurements in the background (no or parasitic signal) but preferring regions of signal leads to more efficient experiments (Fig. 1). Our method [1] for discovering regions of signal is based on Gaussian Process Regression as a technique for probabilistic function approximation and makes use of log-normal distributions. More concretely, we fit a Gaussian process on logarithmic intensity data and exponentiate it to get a log-Gaussian process (Fig. 2). The uncertainty of this log-Gaussian process (in terms of its posterior variance) can now be taken to identify regions of signal as we shall demonstrate. Fortunately, it can be shown that, for large intensities, the resulting log-normal noise distribution converges to a normal distribution which is a natural approximation to Poisson distributed noise in intensity observations. In order to quantify the benefit of our approach, we present results of a benchmarking procedure that we developed as a cost-benefit analysis in a synthetic but still representative setting [2]. Furthermore, we describe the outcome of experiments in real environments at the TAS EIGER (PSI) and IN12 (ILL).

References:

[1] Teixeira Parente et al. *Nat. Commun.* **14**, 2246 (2023)
[2] Teixeira Parente et al. *Front. Mater.* **8**, 772014 (2022)

Figures:

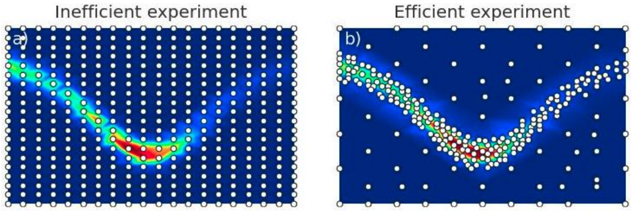

Figure 1: **Example of an inefficient and efficient experiment.** (a) Inefficient experiment with a large part of measurement locations (dots) in the background (dark blue area). (b) More efficient experiment with most measurement locations in the region of signal (remaining area).

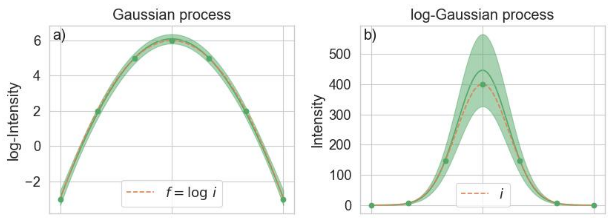

Figure 2: **Transformation of a Gaussian process to a corresponding log-Gaussian process.** The one-dimensional example intensity function i and its logarithm are displayed with dashed orange lines. (a) Gaussian process with observations (green dots) of $\log i$ and its uncertainties (light green area) around its mean function (solid green line). (b) Corresponding log-Gaussian process. Maximizing the uncertainty of the log-Gaussian process enables to find regions of signal.

Biography:

Mario Teixeira Parente is currently an interim professor for Computational Statistics and Data Science at the LMU Department of Statistics. Before he was a postdoc at the Jülich Centre for Neutron Science located at the research reactor FRM2 in Garching. There, he worked on a Helmholtz AI project to autonomize experimental strategies for neutron spectroscopy methods by applying log-Gaussian processes.