technische universität dortmund

fakultät statistik

December 2023

Gaussian Processes and semiparametric modelling for Bayesian Optimization: a benchmark study

Problem definition:

Bayesian Optimization (BO) is a tool for black-box optimization. It is a sequential global optimization technique and addresses problems of the form

$$\boldsymbol{x}^* \in \underset{\boldsymbol{x}\in\boldsymbol{X}}{\mathbf{argmax}}\, \boldsymbol{f}(\boldsymbol{x}),$$

i.e. the aim is to find optimal parameters $\boldsymbol{x}^*$ that maximize an objective function $\boldsymbol{f}$ over a domain $\boldsymbol{X}$ (Garnett, 2023). The gist of BO is to approximate the function $\boldsymbol{f}$ we are trying to optimize with a regression model that is uncertainty aware, and to use these uncertainty estimates to determine what our next test point should be. It is usual practice to do BO using Gaussian Processes (GPs) (Rasmussen and Williams, 2006). An alternative to GPs are semiparametric models using for example splines (Eilers and Marx, 1996).

The aims of this thesis are as follows: First, state-of-the-art software for BO using either GPs or semiparametric models and the corresponding literature is reviewed. Then, having introduced both frameworks, pros and cons of the two modelling approaches should be investigated and benchmarked against each other in a detailed simulation study. Common functions for testing optimization algorithms should be used for this purpose (Finck et al., 2009).

References:

Eilers, P. H.C. and Marx, B. D. (1996). Flexible Smoothing with B-splines and Penalties. Statistical Science, 11(2), 89-121.

Finck, S., Hansen, N., Ros, R., and Auger A. (2009). Real-parameter Black-Box Optimization Benchmarking 2009: Presentation of the noiseless functions. Technical Report 2009/20, Research Center PPE. Updated February 2010.

Garnett, R. (2023). Bayesian Optimization. Cambridge University Press.

Rasmussen, C. E. and Williams, C. K. I. (2006). Gaussian Processes for Machine Learning. The MIT Press. www.GaussianProcess.org/gpml.

Contacts:

Prof. Dr. Katja Ickstadt (ickstast@statistik.uni-dortmund.de)
Prof. Dr. Nadja Klein (nadja.klein@tu-dortmund.de)