Karlsruhe Institute of Technology

Autonome Lernende Roboter (ALR)

Prof. Gerhard Neumann

Project Type

- ☑ Master Thesis
- ☑ Bachelor Thesis
- ☐ Research Project

Supervisors

Maximilian Hüttenrauch

m.huettenrauch@kit.edu

Difficulty

Algorithmic

Math

Application

Recursive Surrogate-Modeling for Stochastic Search

Description

Model-Based Relative Entropy Stochastic Search (MORE) [1] is a recent policy search method for episodic reinforcement learning tasks such as the optimization of movement primitive parameters [2]. It is an information-theoretic black-box optimizer which requires no gradient information or anything else besides the function evaluations of the objective function. It relies on a local quadratic surrogate model of the objective which is estimated from samples collected during the optimization process. While MORE achieves state-of-the-art results in terms of the optimization accuracy, estimating accurate local quadratic models is data-intensive and therefore a large bottle-neck.

So far, the model has been estimated from scratch in each iteration using a form of (weighted) least squares regression. This procedure, however, neglects that subsequent models are highly correlated due to the locality of the data. Hence, a more sample efficient way to estimate the model parameters of the quadratic model is to use recursive estimation approaches such as Kalman filtering. These filtering approaches can, at a later stage of the project, also be replaced with deep neural network filters to allow for transferring knowledge between tasks.

Tasks

In this thesis, you should

- Get familiar with the MORE algorithm and Kalman filtering
- Derive a new algorithm connecting them
- Learn a dynamical system that can be used for filtering of the model parameters
- Test it on existing benchmark functions for optimization and robotics.

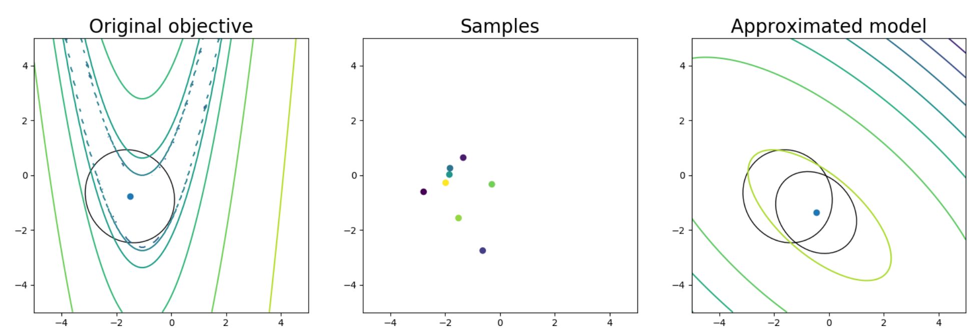

Figure 1: Left: Rosenbrock function in 2D with a search distribution centered at blue dot. Middle: Samples from the distribution with colors according to their function value. Right: The resulting model of the objective function with the old and a new search distribution (centered at blue dot).

References

[1] Abbas Abdolmaleki, Rudolf Lioutikov, Jan R Peters, Nuno Lau, Luis Pualo Reis, and Gerhard Neumann. Model-based relative entropy stochastic search. In *Advances in Neural Information Processing Systems*, pages 3537–3545, 2015.

[2] Auke J. Ijspeert, Jun Nakanishi, and Stefan Schaal. Learning attractor landscapes for learning motor primitives. In S. Becker, S. Thrun, and K. Obermayer, editors, *Advances in Neural Information Processing Systems 15*, pages 1547–1554. MIT Press, 2003.