PhD offer: High-dimensional Bayesian optimization

Period: November 2021 - November 2024
Location: Saint-Étienne and Paris area

Funding:

- CIFRE Stellantis (Industrial Agreement for Training through Research). The CIFRE fellow will sign a 3 years full time work contract with Stellantis.
- Year gross salary: of the order of 30k€.

Advisors: Didier Rullière (École des Mines de Saint-Étienne) and David Gaudrie (Stellantis)

Application: `https://jobs.groupe-psa.com/offre-de-emploi/emploi-optimisation-bayesienne-en-grande-dimension-h-f_9458.aspx`

Application profile: Master in applied mathematics, data science, machine learning, statistics or engineering with mathematical background. Good programming skills in R and/or Python. Knowledge or relevant experience in the field of Gaussian Processes and/or optimization is a plus.

Context: Numerical simulations have been employed for many years at Stellantis in various engineering domains (aerodynamics, combustion, electromagnetic, crash, ...) to predict the performance of a parameterized system, usually built via CAD tools. Optimization methods are used to find the optimal settings of this system. Employed algorithms cope with the fact that the objective function(s) stem from a numerical simulator which takes several minutes/hours for performing one simulation, and can therefore only be called with parsimony. While these methods have demonstrated good performance in low dimensional settings ($\approx$ 5-10 design parameters), difficulties are encountered when the amount of design variables is large, typically over 50.

Objectives of the PhD: The objective of the PhD is to devise and implement an optimization method of high-dimensional time-consuming simulators. Due to the non-linearity and the computation time (few minutes to several hours) of the objective function, surrogate-based approaches in the vein of EGO [1] will be used. Those methods are particularly efficient in the context of expensive black-box simulations.

There are nonetheless some limitations to these methods. First, they behave poorly in larger dimensions. Recent research has attempted to tackle the problem, e.g. by emphasizing the most relevant variables [2], or by projecting the designs in a lower-dimensional subspace where to carry out the optimization [3, 4]. The first objective of the PhD is to devise a method to build an accurate surrogate-model and perform the surrogate-based optimization in

spite of the high dimensionality.

Another critical issue of EGO-like methods is their scalability with respect to the amount of observations: for technical reasons (matrix products and inversions), Gaussian Processes (GP) are limited to approximately 1,000-2,000 observations. However, in applications where the objective function is relatively cheap (say a few minutes) and/or when large parallel computing capabilities are available, this upper bound may be attained and overshot. Building GPs that can efficiently cope with moderate sizes of datasets (up to ≈10,000 observations) is a recent research effort [5, 6] which will be followed during this thesis Remark that building GPs allowing more data than the classical limit (≈1,000) is also a way to improve the predictivity in high dimensional spaces.

Last but not least, the devised optimization method should fully exploit parallel computing capabilities: since it is possible to compute the objective function simultaneously on a large number of computers/nodes of a cluster, the method should not provide a single new design per iteration, but rather a large batch of designs. Techniques for parallelizing the EGO procedure have already been discussed in the literature [7, 8]. But such approaches hinge on Monte Carlo simulations or blindly trust the surrogate model. They are typically limited to 2-4 designs per iteration, especially in high dimensional problems. Additionally, the time spent for optimizing such an acquisition criterion might no longer be negligible when simulations take a few minutes. Devising an efficient and large-scale (≈20-50 new designs) acquisition function is an objective of the thesis.

References

[1] Jones, D. R., Schonlau, M., and Welch, W. J. Efficient global optimization of expensive black-box functions. *Journal of Global optimization*, 1998, vol. 13, no 4, p. 455-492.

[2] Marrel, A., Iooss, B., and Chabridon, V. Statistical Identification of Penalizing Configurations in High-dimensional Thermal-hydraulic Numerical Experiments: The ICSCREAM Methodology. arXiv preprint arXiv:2004.04663, 2020.

[3] Gaudrie, D., Le Riche, R., Picheny, V., Enaux, B., and Herbert, V. Modeling and optimization with Gaussian processes in reduced eigenbases. *Structural and Multidisciplinary Optimization*, 2020, vol. 61, no 6, p. 2343-2361.

[4] Constantine, P. G., Dow, E., and Wang, Q. Active subspace methods in theory and practice: applications to kriging surfaces. *SIAM Journal on Scientific Computing*, 2014, vol. 36, no 4, p. A1500-A1524.

[5] Rullière, D., Durrande, N., Bachoc, F., and Chevalier, C.. Nested Kriging predictions for datasets with a large number of observations. *Statistics and Computing*, 2018, vol. 28, no 4, p. 849-867.

[6] Hensman, J., Fusi, N., and Lawrence, N. D. Gaussian processes for big data. arXiv preprint arXiv:1309.6835, 2013.

[7] Ginsbourger, D., Le Riche, R., and Carraro, L. Kriging is well-suited to parallelize optimization. *Computational intelligence in expensive optimization problems*. Springer, Berlin, Heidelberg, 2010. p. 131-162.

[8] Chevalier, C. and Ginsbourger, D. Fast computation of the multi-points expected improvement with applications in batch selection. *International Conference on Learning and Intelligent Optimization*. Springer, Berlin, Heidelberg, 2013. p. 59-69.