High-Dimensional Bayesian Optimization with a Combination of Kriging models

Tanguy Appriou[†,1,2], Didier Rullière[§,2], David Gaudrie[§,1]

[†] PhD student (presenting author). [§] PhD supervisor

PhD expected duration: Nov. 2021 – Nov. 2024

[1] Mines Saint-Etienne, Univ Clermont Auvergne, CNRS, UMR 6158 LIMOS, Institut Henri Fayol, F - 42023 Saint-Etienne France
tanguy.appriou@emse.fr
drulliere@emse.fr

[2] Stellantis, Centre d'Expertise R&D Carrières-sous-Poissy, F-78955 Carrières-sous-Poissy, France
david.gaudrie@stellantis.com

Abstract

Kriging models, also called Gaussian Process regression, [7] are commonly used as surrogate models of expensive computer codes in various applications. In engineering design optimization, Bayesian approaches based on Kriging models such as efficient global optimization (EGO) [6] are employed to speed-up the optimization process by reducing the number of function evaluations. These methods have been successfully applied to many real-world applications in low dimensions (less than 30 design parameters) [4].

However, engineering designs can often be parametrized by more than 50 parameters in practice. In higher dimensions, Kriging suffers from the curse of dimensionality and building an accurate surrogate model is met with various setbacks. One of the main challenges is related to optimizing the covariance length-scale hyperparameters of the model. These hyperparameters regulate the decay of the correlation between observations when their distance increases, and anisotropic Kriging models consider one length-scale per dimension. Estimating the length-scales correctly is essential to obtain a model with a good accuracy. They are typically obtained by maximizing the log-likelihood of the model. In high-dimension, this inner optimization is problematic due to the exponential growth of the search space with the dimension and to over-fitting issues when there are too few observations. For these reasons, maximum likelihood estimation of the hyperparameters often fails to provide correct values, especially when not enough observations are available [5, 1]. This is problematic for Bayesian optimization where the number of samples at the start of the optimization is low, and where inaccurate surrogates negatively impact the convergence speed. Several papers address these difficulties by either reducing the dimension of the problem [2], or by considering simplifying hypothesis such as additive models [3]. However, both these methods make additional assumption on the underlying function to approximate (low dimension representation or additive structure) which are not necessarily satisfied in practice, thus they are not easily generalizable to any design engineering problems.

In this presentation, we introduce a new method for high-dimensional Bayesian optimization using a convex combination of Kriging models. This method bypasses the length-scales optimization by combining sub-models with random length-scales and its expression is obtained in closed-form expression avoiding any inner optimization. We also describe how to sample suitable length-scales for the sub-models using a criterion based on the entropy of the correlations,

in order to avoid degenerated sub-models with either too large or too small length-scales. Finally, the variance of the combination being not directly available as the correlations between sub-models are unknown, we present a method to compute the prediction variance for any weighting methods. This is done by introducing a global covariance structure also based on a linear combination of covariances. We apply our combined Kriging model to high-dimensional EGO for analytical test functions and for the design of an electric machine. We show that the classical Kriging approach using maximum likelihood estimation fails to properly optimize the length-scale hyperparameters and that our method successfully builds more accurate surrogate models at the beginning of the optimization loop. This results in faster convergence speed for EGO using the combination, thus reducing the number of computer code evaluations to reach the optimum.

Short biography (PhD student)

Tanguy Appriou received an engineering degree from Ecole Centrale de Lyon and a master in aerospace engineering from Tohoku University in 2021. He is currently pursuing a Ph.D. degree in Stellantis and Mines Saint-Etienne. This thesis is partly funded by a CIFRE grant (convention #2021/1284) established between the ANRT and Stellantis.

This thesis is conducted with the support of the consortium in Applied Mathematics CIROQUO (`https://doi.org/10.5281/zenodo.6581217`), gathering partners in technological and academia in the development of advanced methods for Computer Experiments.

References

[1] Tanguy Appriou, Didier Rullière, and David Gaudrie. Combination of optimization-free kriging models for high-dimensional problems. *Computational Statistics*, pages 1–23, 2023.

[2] Paul G Constantine. *Active subspaces: Emerging ideas for dimension reduction in parameter studies.* SIAM, 2015.

[3] Nicolas Durrande, David Ginsbourger, and Olivier Roustant. Additive Covariance kernels for high-dimensional Gaussian Process modeling. *Annales de la Faculté des sciences de Toulouse : Mathématiques*, Ser. 6, 21(3):481–499, 2012.

[4] Alexander Forrester, Andras Sobester, and Andy Keane. *Engineering design via surrogate modelling: a practical guide.* John Wiley & Sons, 2008.

[5] David Ginsbourger, Delphine Dupuy, Anca Badea, Laurent Carraro, and Olivier Roustant. A note on the choice and the estimation of kriging models for the analysis of deterministic computer experiments. *Applied Stochastic Models in Business and Industry*, 25(2):115–131, 2009.

[6] Donald R Jones, Matthias Schonlau, and William J Welch. Efficient global optimization of expensive black-box functions. *Journal of Global optimization*, 13(4):455–492, 1998.

[7] Carl Edward Rasmussen and Christopher KI Williams. *Gaussian processes for machine learning.* MIT press Cambridge, MA, 2006.