Proceedings of the 2022 Winter Simulation Conference
B. Feng, G. Pedrielli, Y. Peng, S. Shashaani, E. Song, C.G. Corlu, L.H. Lee, E.P. Chew, T. Roeder, and P. Lendermann, eds.

MACHINE LEARNING-BASED UNCERTAINTY PREDICTION FOR EFFICIENT GLOBAL OPTIMIZATION

Zizhou Ouyang
Belen Martin-Barragan

Business School
University of Edinburgh
29 Buccleuch Place
Edinburgh, EH8 9JS, UK

Xuefei Lu

SKEMA Business School
Université Côte d'Azur
5 Quai Marcel Dassault
Paris, 92150, France

ABSTRACT

Efficient global optimization (EGO) is widely used to solve black-box expensive-to-evaluate objectives. However, when functions are scaled to high dimensions, classical EGO meets problems as it requires considerable computational overhead (i.e., the huge computation cost of matrix inversion in Kriging). To address this challenge, we propose a machine learning-based uncertainty prediction method as an emerging alternative that replaces Kriging with a machine learning model and uncertainty prediction estimation. The proposed uncertainty prediction method can be applied to most discriminative machine learning models to solve global optimization problems. This study presents an example that leverages the discriminative super vector machine (SVM) to build and update the metamodel. A score function based on metamodel prediction, and an uncertainty estimation method are applied to balance exploration and exploitation. Numerical results show that our method performs well in high-dimensional benchmark test functions.

1 INTRODUCTION

The optimization methods for black-box, expensive-to-evaluate problems have drawn increasing research attention in recent years as these scenarios exist widely in real-world applications. One of the biggest challenges in solving such problems is to find a satisfactory solution with a limited computation budget.

Efficient global optimization (EGO) (Jones et al. 1998) broadly serves as the framework to optimize expensive, black-box functions. The classical EGO sequentially updates the Kriging (also called Gaussian process regression) metamodel and then finds the next sample point based on the expected improvement (EI) infill function under the Gaussian process prior. The EI function includes the prediction of input points and uncertainty prediction based on previous evaluations (Hong et al. 2015).

Uncertainty plays an important role in EGO and helps to find a solution near true global optimum. A point is expected to improve the objective function and be selected for the subsequent evaluation if its prediction is significantly better than the current optimum or has relatively high prediction uncertainty. Without uncertainty, the searching process easily gets stuck in sub-optimal solutions as it ignores exploring unknown regions.

In general, Kriging and EI are the ideal choices for EGO as Kriging provides the prediction variance and helps to formulate the uncertainty prediction in EI. When applied in high dimensions, the Kriging and EI evaluations become complicated and time-consuming. Machine learning models are reliable and efficient in high-dimensional problems. Therefore, we propose an uncertainty prediction method based on machine learning models to replace Kriging and EI, which can solve high-dimensional optimization problems with faster convergence and less computational cost.

2 METHODOLOGY

In EGO, the EI function can be written as:

$$\widehat{\mathrm{EI}}(\mathbf{x}) = \left(f_{min} - \hat{g}(\mathbf{x})\right)\Phi\left(\frac{f_{min} - \hat{g}(\mathbf{x})}{s\{\hat{g}(\mathbf{x})\}}\right) + s\{\hat{g}(\mathbf{x})\}\phi\left(\frac{f_{min} - \hat{g}(\mathbf{x})}{s\{\hat{g}(\mathbf{x})\}}\right) \quad (1)$$

where the f_{min} is the current best solution. The Φ and ϕ represent the cumulative distribution function (CDF) and probability density function (PDF) of normal distributions, respectively. $\hat{g}(\mathbf{x})$ is the Kriging prediction, and $s\{\hat{g}(\mathbf{x})\}$ is the prediction standard deviation. In equation (1), the first item reflects the exploitation process, which evaluates the difference between the current best solution and the prediction of potential solutions. The second item is the exploration process which evaluates the uncertainty of the possible solutions.

Unlike classical EGO heavily rely on the prior distribution of the Gaussian process, our idea is to adopt a simple but powerful machine learning model as the metamodel and measure uncertainty without assuming Gaussian distribution. Take the supper vector machine (SVM) as the metamodel as an example. At first, we distinguish the evaluated points in each iteration into good and bad points based on their properties, and the dataset $\mathcal{D} = \{\mathcal{X}, \mathbf{y}\}$ is created with input points $\mathcal{X}$ and the corresponding labels $\mathbf{y}$ (either +1 or -1). Then, the modified dataset is used to fit the SVM metamodel $g'(\mathbf{x})$. The proposed infill criterion selects a promising point based on the prediction and the uncertainty estimation, and then the dataset is updated. The proposed infill criterion includes a score function for the improvement evaluation and an uncertainty function for uncertainty estimation. The score function can be written as:

$$\widehat{\mathrm{EI\_Score}}(\mathbf{x}) = \sum_{i=1}^{m} a_i \hat{\mathbf{y}}_i K\left(\mathbf{x}_i, \mathbf{x}\right) + b \quad (2)$$

where $\mathbf{x}_i, \hat{\mathbf{y}}_i$ are the *i*th candidate point and the predicted class label (either +1 or -1), respectively. The a_i is the coefficient associated with the *i*th candidate point. The $K(\mathbf{x}_i, \mathbf{x})$ is the kernel function, and b is a scale.

The uncertainty is quantified based on the variance of score prediction in trained SVM metamodels in different iterations. With the help of uncertainty prediction, we get more information about unknown regions and make the solution closer to the true global optimum.

We compare the proposed algorithm with classical EGO on the Griewank test function (Locatelli 2003) on 20 and 30 dimensions. The design space for the Griewank function is $[-10,10]^{\mathrm{D}}$, and the global optimum is 0. The number of initial points is 50, and the max evaluations are 1000. Experimental results are shown in Table 1.

Table 1. Comparison of the proposed method and classical EGO on Griewank 20D and Griewank 30D functions.

	Proposed Method		Classical EGO	
	Mean	Std.	Mean	Std.
Griewank 20D	1.10E-02	6.73E-01	4.71E-02	7.95E-02
Griewank 30D	9.27E-01	8.97E-02	9.96E-01	1.04E-01

REFERENCES

Hong, L. J., B.L. Nelson, and J. Xu. 2015. “Discrete optimization via simulation”. In *Handbook of simulation optimization* 9-44. Springer, New York, NY.

Jones, D. R, and M. Schonlau, and W.J. Welch. 1998. “Efficient Global Optimization of Expensive Black-box Functions”. *Journal of Global optimization* 13(4):455-492.

Locatelli, M. 2003. “A note on the Griewank test function”. *Journal of global optimization* 25(2):169-174.