Scalable Bayesian Optimization with Metaheuristics for Efficiency and Exploitation

Ibrahim Aydogdu; Akdeniz University; Turkey
Michaela Kempner; Georgia Institute of Technology; United States
Yan Wang; Georgia Institute of Technology; Turkey

Bayesian optimization (BO) is a surrogate-based global optimization approach, where a surrogate of the objective is constructed and used to guide the search. The surrogate is updated iteratively as an active learning strategy. Thus, BO can efficiently solve optimization problems with expensive-to-evaluate black-box objective functions with fewer iterations. Gaussian process (GP) is a common surrogate model for BO which also quantifies uncertainty. Although BO with GP (BOGP) performs well in small-scale optimization problems, its effectiveness decreases in high-dimensional optimization problems because the number of samples required to cover the search space increases exponentially. Calculating the inverse of the covariance matrices in GP models is in the cubic order of the number of samples. High computational cost is the major barrier for BOGP to solve large-scale optimization problems. Some special tactics have been developed to reduce the computational cost and improve the scalability of GP, such as kernel matrix approximations (Sang & Huang, 2012; Smola & Bartlett, 2000; Yang, 2020), the Fully Independent Training Conditional method (Csató & Opper, 2002), and Variational Free Energy method (Titsias, 2009) , which show good performance in the learning process of large data sets. However, these techniques mainly focus on how to represent the whole solution space. These methods do not directly help improve the convergence of optimization. Recently some researchers developed scalable GP methods for global optimization by searching for the optimum samples, namely inducing points, for better representation of the solution space (Fortuin, Dresdner, Strathmann, & Rätsch, 2021; Zhang, Yao, Liu, & Wang, 2019). In these studies, the criteria for selecting the inducing points are to maximize the likelihood of data similarity. However, in the optimization process, the potential to reach the best solution is more important than the similarity of the data. Therefore, the exploitation capabilities of these techniques are limited. The performance of BOGP can be improved if the distance to the best solution is considered when choosing the inducing points instead of the similarity between the samples. In this study, a new method called scalable Bayesian optimization with metaheuristics (SBO-MH) is developed to increase the exploitation capability of the BOGP for solving large-scale global optimization problems. In each iteration of SBO-MH, two optimization procedures are taken to update the surrogate and select the next sample. In the first optimization, SBO-MH selects the optimal inducing points that maximizes the probability of finding a better solution than the current best one. In the second optimization, the next sample is selected to maximize the acquisition function. For both optimization procedures, biogeography based optimization (BBO) (Simon, 2008) is utilized, whose efficiency has been proven in many structural optimization problems (Ibrahim Aydogdu, 2017; Ibrahim Aydogdu & Akin, 2015; I. Aydogdu, Carbas, & Akin, 2017). SBO-MH is tested by three benchmark functions with three different dimensions. The obtained results are compared with the full-scale BOGP method with the criteria of statistical consistency, convergence speed, and computational cost.

References Aydogdu, I. (2017). Cost optimization of reinforced concrete cantilever retaining walls under seismic loading using a biogeography-based optimization algorithm with Levy flights. Engineering Optimization, 49(3), 381-400. doi:10.1080/0305215X.2016.1191837 Aydogdu, I., & Akin, A. (2015). Biogeography Based CO2 and Cost Optimization of RC Cantilever Retaining Walls. Paper presented at the 17th International Conference on Structural Engineering, Paris, France. http://waset.org/abstracts/25608 Aydogdu, I., Carbas, S., & Akin, A. (2017). Effect of Levy Flight on the discrete optimum design of steel skeletal structures using metaheuristics. Steel and Composite Structures, 24(1), 93-112. doi:10.12989/scs.2017.24.1.093 Csató, L., & Opper, M. (2002). Sparse on-line Gaussian processes. Neural computation, 14(3), 641-668. Fortuin, V., Dresdner, G., Strathmann, H., & Rätsch, G. (2021). Sparse Gaussian processes on discrete domains. Ieee Access, 9, 76750-76758. Sang, H., & Huang, J. Z. (2012). A full scale approximation of covariance functions for large spatial data sets. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 74(1), 111-132. Simon, D. (2008). Biogeography-Based Optimization. Ieee Transactions on Evolutionary Computation, 12(6), 702-713. doi:10.1109/tevc.2008.919004 Smola, A., & Bartlett, P. (2000). Sparse greedy Gaussian process

regression. Advances in neural information processing systems, 13. Titsias, M. (2009). Variational learning of inducing variables in sparse Gaussian processes. Paper presented at the Artificial intelligence and statistics. Yang, A. (2020). Scalable Bayesian Optimization with Sparse Gaussian Process Models. arXiv preprint arXiv:2010.13301. Zhang, J., Yao, X., Liu, M., & Wang, Y. (2019). A Bayesian Discrete Optimization Algorithm for Permutation Based Combinatorial Problems. Paper presented at the 2019 IEEE Symposium Series on Computational Intelligence (SSCI).