PHYSBO – optimization tools for PHYSics based on Bayesian Optimization–

Ryo TAMURA
International Center for Materials Nanoarchitectonics,
National Institute for Materials Science,
1-1 Namiki, Tsukuba, Ibaraki, 305-0044

Bayesian optimization (BO) is machine-learning based black-box optimization technique and has recently garnered significant attention in physics, chemistry, and materials science[1]. For example, for materials developments, the trial-and-error process to find better material is regarded as optimizing a black-box function where input is the composition, structure, and process and output will be materials property. In this algorithm, first, a Gaussian process regression that predicts the expected property and variance is constructed from the already observed input-output pairs. Next, through the trained Gaussian process, the probable input data that will yield the desired output value are also selected based on the acquisition function using expected property and variance. Then, the true output value for the selected candidate is obtained by experiments or simulations as black-box functions. BO repeats these processes to find better inputs. Although BO is powerful tool for black-box optimization, BO is generally computationally expensive in two parts: training Gaussian process regression and optimizing acquisition function.

COMBO (COMmon Bayesian Optimization) has been developed mainly for researchers in the materials science field[2]. In the Gaussian process, two hyperparameters, i.e., parameters whose values were given prior to learning, existed: the Gaussian kernel width and noise variance. Using the COMBO package, these hyperparameters were automatically determined by maximizing the Type-II likelihood. In addition, to avoid computationally expensive for training Gaussian process regression, COMBO achieves high scalability using Thompson sampling, random feature map, and one-rank Cholesky update,

To accelerate COMBO package further, PHYSBO (optimization tools for PHYSics based on Bayesian Optimization) package is developed as Python 3 code[3]. In PHYSBO, the massive parallelization using ISSP supercomputer can be performed for optimizing acquisition function, and then both computationally expensive parts in BO can be resolved. In addition, new function to perform multiobjective optimization is implemented.

In physics field, so far, BO has been applied to some problems such as autonomous X-ray scattering experiments [4], inverse scattering [5], crystal structure prediction [6], and effective model estimation[7]. Thus, PHYSBO package can accelerate these problems, and will solve more complex physical problems using supercomputer.

This package is developed with Tsuyoshi Ueno, Kei Terayama, Koji Tsuda, Yuichi Motoyama, Kazuyoshi Yoshimi, and Naoki Kawashima. We would like to thank the support from “Project for advancement of software usability in materials science” by The Institute for Solid State Physics, The University of Tokyo, for development of PHYSBO.

References

[1] K. Terayama, M. Sumita, R. Tamura, and K. Tsuda, Black-box optimization for automated discovery, Accounts of Chemical Research **54**, 1334 (2021).

[2] T. Ueno, T. D. Rhone, Z. Hou, T. Mizoguchi, and K. Tsuda, COMBO: An efficient Bayesian optimization library for materials science, Materials Discovery, **4**, 18 (2016).

[3] https://www.pasums.issp.u-tokyo.ac.jp/physbo/

[4] M. M. Noack, K. G. Yager, M. Fukuto, G. S. Doerk, R. Li, and J. A. Sethian, A kriging-based approach to autonomous experimentation with applications to x-ray scattering, Scientific Reports **9**, 1 (2019).

[5] R. Vargas-Hernández, Y. Guan, D. Zhang, and R. Krems, Bayesian optimization for the inverse scattering problem in quantum reaction dynamics, New Journal of Physics **21**, 022001 (2019).

[6] T. Yamashita, N. Sato, H. Kino, T. Miyake, K. Tsuda, and T. Oguchi, Crystal structure prediction accelerated by bayesian optimization, Physical Review Materials **2**, 013803 (2018).

[7] R. Tamura and K. Hukushima, Bayesian optimization for computationally extensive probability distributions, PLoS ONE **13**, e0193785 (2018).