Multi-objective Bayesian global optimization for continuous problems and applications

Yang, K.

Chapter 8

Conclusions and Outlooks

Multi-objective Bayesian global optimization is an effective method when the objective functions cannot be directly measured or evaluation of the objective functions is too expensive. The basic concept behind MOBGO is using Kriging/Gaussian Process to build the independent models based on the relationship between inputs and outputs of a real problem. After this step, an optimization algorithm is applied to find an optimal solution based on the Kriging models and an infill criterion. Then, the optimal solution will be evaluated by the real objective functions to update the Kriging models. This loop will be repeated until it meets the stop criterion.

Compared to evolutionary algorithms, Bayesian global optimization is more efficient when dealing with expensive function evaluation problems. However, there still exists much space to improve the efficiency of MOBGO. The efficiency of MOBGO is mainly determined by three aspects: the computational complexity of updating a Kriging model, the computational complexity of infill criteria, and efficiency of a single objective optimization to find the optimal solution based on the Kriging models. Targeting at reducing the computational complexity of EHVI calculation and improving the performance of MOBGO in this dissertation, an efficient EHVI calculation algorithm, two new infill criterion (TEHVI and EHVIG) were proposed, and a thorough research on them was conducted.

Section 8.1 provides the conclusions drawn from works in this dissertation and Section 8.2 discusses an outlook and directions of the future research.

8.1 Conclusions

During the process of searching the optimal solution, the infill criterion in MOBGO acts as a pre-selection criterion and the maximizing of the infill criterion is equivalent to achieving an optimal Pareto front set. The infill criterion in MOBGO is very crucial, considering that it merges each objective into a single value. Then, according to the value of this infill criterion, the optimization algorithm can select the optimal solution by choosing the solution which has the maximum/minimum infill criterion value. Among the infill criteria, EHVI performs much better than the other criteria, because it can balance the exploration and exploitation well in the objective space. However, EHVI is rarely applied in real applications for its high computational complexity. An efficient EHVI calculation algorithm is proposed. The computational complexity of EHVI is decreased to $\Theta(n \log n)$, for both 2-D and 3-D cases, by an efficient non-dominate space partitioning method and a new EHVI calculation formula. It is also proven that the number of partitioned slices in non-dominated space is always $n + 1$ and $2n + 1$ for 2-D and 3-D, respectively. For the high dimensional cases ($d \geq 4$), the proposed algorithm is much faster than the previous methods. Moreover, the proposed algorithm can also be extended to other integral-based criteria calculation, like PoI, TEHVI and EHVIG.

The inherent assumption of EHVI is that each objective function follows a normal distribution and ranges from $-\infty$ to ∞. However, the rough range of an objective function is obtained from experts in some real applications, and this a-prior knowledge was not utilized in EHVI. To solve this problem, a new criterion, TEHVI derived from EHVI, was proposed in this dissertation. Compared to EHVI, TEHVI follows a truncated normal distribution within a lower bound and an upper bound for each objective function. The experimental results show that TEHVI outperforms EHVI on benchmarks and PID robust parameter tuning problems.

Moreover, TEHVI can also be used to solve the preference-based multi-objective optimization problems, since the boundary information is related to the objective space and can be set by a decision maker according to his/her preferences. Inspired by TEHVI, THV is proposed and applied to solve preference-based multi-objective optimization problems. Compared to TEHVI, THV can be used to any evolutionary algorithms, while TEHVI can only be utilized in BGO. The experimental results that show both of two approaches are effective.

Based on parts of the works in this dissertation, the computation of EHVI and TEHVI is considered to be efficient, as illustrated in Chapter 3 and Chapter

4. However, the maximization of EHVI still requires significantly amount of time in MOBGO, because EHVI needs to be calculated for many times in each iteration. Since the landscape of a Kriging model is continuous, and considering the definition of EHVI that the integral of hypervolume improvement times its corresponding probability density function, EHVI should be differentiable at a target point, and EHVIG is proposed to speed up the convergence of the optimizer in MOBGO by using GAA. However, compared to other optimizers (CMA-ES, GA), GAA is much faster, but it is very easy to be stuck at a local optimal point. To improve the convergence of the optimizer, EHVIG is applied as a stopping criterion in EAs, as the gradient of EHVI at an optimal point should be a zero vector and EHVIG can force an EA to stop earlier when an EA find the optimal solution. The experimental results show that using EHVIG as a stopping criterion needs less execution time and improves the quality of the final Pareto front approximation on the benchmarks.

8.2 Outlooks

This dissertation aims to improve the efficiency and effectiveness of MOBGO by reducing the computational complexity of integral-based infill criterion and proposing two new infill criteria. Some open questions related to this dissertation for future research are:

Improving the efficiency of EHVI calculation for $d \geq 4$ case: The partitioning algorithm of EHVI for high dimensional case ($d \geq 4$) is based on two state-of-the-art algorithms, DKLV17 and LKF17. These two algorithms are very efficient to partition the dominated space, but not non-dominated space. Therefore, the research of efficient partitioning non-dominated space is highly recommended.

Multivariate surrogate models with output dependence assumption: Currently, Kriging/Gaussian processes build a surrogate model for each objective function and surrogate models are independent. However, in some real applications, the objective functions are dependent. This can lead to inappropriate representations of joint uncertainty. The potential solutions could be achieved by principal component analysis (PCA) and borrowing the idea of decoupling method in control theory.

Reference point free in infill criteria: Setting a reference point for hypervolume-based criteria is very tricky. For a minimization problem, a big reference point can more easily lead an algorithm to focus on the extreme points. On the other hand, the optimization algorithm would omit the extreme points if the reference point is small. Recently, a dynamic reference point strategy, which uses a big reference point in the early iteration and decreases the reference point in the later iterations, is applied in some papers. This strategy is useful but will also lead to another question: how to set the dynamic reference point strategy effectively? Actually, the most effective solution to this problem is to remove the concept of a reference point of an infill criterion theoretically, and this will be an interesting research topic in the future.