A Hybrid of Bayesian-Based Global Search With Hooke–Jeeves Local Refinement for Multi-Objective Optimization Problems

Linas Litvinas

Vilnius University, Faculty of Mathematics and Informatics

e-mail: linas.litvinas@mif.vu.lt

Introduction

The proposed multi-objective optimization algorithm hybridizes random global search with a local refinement algorithm [1]. The global search algorithm mimics the Bayesian multi-objective optimization algorithm. The site of current computation of the objective functions by the proposed algorithm is selected by randomized simulation of the bi-objective selection by the Bayesian-based algorithm. The advantage of the new algorithm is that it avoids the inner complexity of Bayesian algorithms. A version of the Hooke–Jeeves algorithm is adapted for the local refinement of the approximation of the Pareto front. The developed hybrid algorithm is tested under conditions previously applied to test other Bayesian algorithms so that performance could be compared. Other experiments were performed to assess the efficiency of the proposed algorithm under conditions where the previous versions of Bayesian algorithms were not appropriate because of the number of objectives and/or dimensionality of the decision space.

The proposed hybrid algorithm

The multi-objective minimization problems:

$$\min F(x), x \in A \subset \mathbb{R}^d, F(x)=\left(f_1(x),\dots,f_m(x)\right)^T$$

are considered assuming that the feasible region (decision space) is a unit hypercube $A=[0;1]^d$ to which any hyperrectangular region can be rescaled.

The algorithm consists of two alternating counterparts carried out multiple times: random global search and local refinement of the Pareto front approximation found by the global search algorithm. The Bayesian global optimization strategy would be preferable because of the rational balancing of exploration and exploitation. However, the number of iterations (computations of the objective functions) of the Bayesian algorithms is considerably limited due to their inner computational complexity. The number of iterations is limited to several hundred for the standardly implemented Bayesian algorithms. The idea is to mimic the search strategy of the Bayesian algorithm without using a stochastic model of the objective functions and thus avoiding the basic computational burden. A randomized algorithm is proposed where the criteria for selecting a point for computing the objective functions are similar but much simpler than the Bayesian approach-based criteria [2]. The reduced computational burden allows a considerably larger number of iterations. The global search phase interchanges with the local refinement phase. The approximation of the Pareto front is refined by a version of the Hooke–Jeeves algorithm adapted to multi-objective optimization.

The considered algorithm is initialized by the predefined number of computations of the objective functions at the random points uniformly distributed in the feasible region A. Let us consider the current start of global search; the set of points where the objective functions are already computed is denoted by $U_A=\{x_i \in A | i=1,\dots,N\}$ and the corresponding set of function values is $U=\{F(x) | x \in U_A\}$. Non-dominated points of U constitute current Pareto optimal solutions set in objective space $P \subseteq U$ and corresponding Pareto optimal solutions set in decision space $P_A \subseteq U_A$.

Strategies to mimic Bayesian search should combine simplicity and numerical substantiation. Combining both mentioned properties the exploration-exploitation strategy can be used, i.e. let say new random points $U_A^{new}=\{x_i \in A | i=1,\dots,q \cdot N\}$ are uniformly generated in feasible region A (q is parameter value) and for every point $x_i \in U_A^{new}$ bi-objective selection functions calculated: $\theta_1(x_i), \theta_2(x_i)$. $\theta_1(x_i)$ is a generated point's x_i Euclidean distance to the closest known point $x_{min} \in U_A$ and $\theta_2(x_i)$ is the closest known point's $x_{min} \in U_A$ function evaluation vector's $F(x_{min}) \in U$ Euclidean distance to the closest current Pareto optimal solution $F_{min} \in P$. The exploration-exploitation strategy is achieved by making new function evaluations at points $x_i \in U_A^{new}$ having minimal trade-off Pareto optimal values of selection functions $(-\theta_1(x_i), \theta_2(x_i))$. Function $\theta_1(x_i)$ has a negative sign because it is maximized. After new function evaluations are made, the sets U_A, U, P_A, P are updated and reused in next phases and next iterations. To induce function evaluations clustering near Pareto optimal solutions, the global search has two random points generation modes: global uniform generation in all feasible region A and local generation near current Pareto optimal solutions.

The proposed exploration-exploitation global search strategy lacks effective local refinement since randomly generated points lack improvement direction. On the contrary, the Hooke–Jeeves single-objective optimization algorithm searches function improvement direction by taking a step in all decision variables [3]. A multi-objective optimization problem is reduced to a single-objective optimization problem by novel approach without the use of the weight vectors.

The multi-objective function is converted to a single-objective surrogate function $f_s(x), x \in A$, which has a current solution: a decision vector x_{cur} and multi-objective function evaluation vector $F(x_{cur})$. Initially, the surrogate function has value $f_s(x_{cur})=0$. When the new solution x_{new} dominates the current solution $F(x_{cur}) \prec F(x_{new})$, the surrogate function value at new location x_{new} decreases $f_s(x_{new})=f_s(x_{cur})-1$ and the current solution is updated $x_{cur}=x_{new}$. Otherwise when the new solution x_{new} does not dominate the current solution the surrogate function value at new location x_{new} remains the same $f_s(x_{new})=f_s(x_{cur})$ and the current solution x_{cur} is not updated, i.e.,

$$f_s(x)=\begin{cases} f_s(x_{cur})-1, when\ F(x_{cur}) \prec F(x) \\ (current\ solution\ is\ updated: x_{cur}=x), \\ f_s(x_{cur}), when\ F(x_{cur}) \not\prec F(x). \end{cases}$$

For every current Pareto optimal solution $x_P \in P_A$ a separate surrogate function having Pareto optimal solution as current solution $x_{cur}=x_P$ is defined. Every defined surrogate function is optimized using the Hooke–Jeeves optimization algorithm taking x_P as the start point:

$$\min f_s(x), x \in A \subset \mathbb{R}^d.$$

The global search phase interchanges with the local refinement phase multiple times. The algorithm works until the maximum allowed number of function evaluations or the maximum iteration number of global searches with local refinement is reached. The algorithm returns a set of all function evaluations and a set of non-dominated Pareto optimal solutions.

Numerical experiments

In case when experiments have low functions evaluation budget, the results of the proposed algorithm were compared with the results of Bayesian rooted optimization algorithms: standard and partition-based implementations of the P-algorithm [4]. First bi-objective test problem is solved to illustrate the performance:

$$f_k(x)=1-e^{-\sum_{i=1}^{d}\left(x_i+(-1)^k/\sqrt{d}\right)^2}, d=2, -4 \leqslant x_1, x_2 \leqslant 4, k=1,2.$$

The second bi-objective test problem is two Shekel functions frequently used to evaluate global optimization algorithms. For the comparison performance NN, GD, EI metrics applied. Where NN is number of non-dominated solutions found by an optimization Algorithm. GD is computed as the maximum of distances between the found non-dominated solutions and their closest neighbors from the Pareto front. EI is computed as the maximum of distances between the true Pareto front and their closest neighbors from found non-dominated solutions.

Since the proposed algorithm and P-algorithm are stochastic, the test problems were solved 100 times. The mean values and standard deviations of the considered metrics are present in two columns of Table 1. Otherwise, the hyperrectangle partition based P-algorithm is deterministic, so its results occupy a single column for each test problem. In case of first problem the proposed algorithm gives better NN and EI values than the P-algorithm. In case of second problem the proposed algorithm gives the best NN value and gives better GD and EI values than the hyperrectangle partition-based P-algorithm.

Table 1. Mean values and standard deviations of performance criteria (NN, GD, EI).

Criteria	P-algorithm				Partition-based		Proposed algorithm			
	Problem I		Problem II		Problem I	Problem II	Problem I		Problem II	
NN	9.87	1.4	15.7	2.0	27	18	12.61	3.11	25.35	5.378
GD	0.015	0.0061	0.07	0.051	0.015	0.21	0.052	0.025	0.161	0.084
EI	0.2	0.034	0.13	0.053	0.092	0.25	0.139	0.018	0.204	0.069

Numerical experiments using test problems having many decision variables and using test problems having many-objective function are performed. Due to limited format of poster presentation results are not presented here, but can be found in [1].

Conclusions

The hybrid multi-objective optimization algorithm is proposed combining random global search and local refinement of the found approximation of the Pareto front. The global search algorithm mimics the Bayesian algorithm. The Hooke–Jeeves algorithm is used for local refinement. At the local optimization phase, the multi-objective optimization problem is converted to a single-objective optimization problem by introducing a surrogate function without the use of the weight vectors. The developed algorithm was tested in the case of extremely low functions evaluation budget, and the proposed algorithm gives decent performance comparing results with the results of Bayesian rooted optimization algorithms. Also, the proposed optimization algorithm was tested with many decision variables test suite and with many-objective test suite, and the results of numerical experiments showed good performance compared with the results of popular evolutionary optimization algorithms NSGA2 and NSGA3. In the case of many-objective optimization, the proposed algorithm need no predefined reference points, so raw solutions selection can be done by a human expert to be tuned to final solutions of needed properties. Future plans include the development parallel version of the proposed algorithm to optimize functions with an extreme computational burden. Another research theme of interest is integrating the proposed algorithm execution with human expert decisions to get solutions of needed properties.

References

[1] L. Litvinas, A hybrid of Bayesian-based global search with Hooke–Jeeves local refinement for multi-objective optimization problems, *Nonlinear Analysis: Modelling and Control*, 27(3), pp. 534–555, 2022..

[2] A. Zhigljavsky, A. Žilinskas, *Bayesian and High-Dimensional Stochastic Optimization*, Springer, Cham, 2021

[3] C.T. Kelley, *Iterative Methods for Optimization,* SIAM, Philadelphia, 1999.

[4] A. Žilinskas, L. Litvinas, A partition based Bayesian multi-objective optimization algorithm, *Numerical Computations: Theory and Algorithms*, pp. 511–518, 2020.