

HMMOCS 2022**International Workshop "Hybrid methods of modeling and optimization in complex systems"****AN INVESTIGATION OF THE HYBRIDIZATION OF
DE AND BFGS ALGORITHMS**

Anton Sopov (a)*, Pavel Sherstnev (b)

*Corresponding author

(a) Reshetnev Siberian State University of Science and Technology, Krasnoyarsk, Russia,
antonsopov2004@gmail.com

(b) Siberian Federal University, Krasnoyarsk, Russia, sherstpasha99@gmail.com

Abstract

Many real-world global optimization problems are presented by a black-box model, therefore there is no information on properties of the objective function and its derivatives. Differential evolution (DE) is one of search algorithms, which demonstrates the high performance in solving global black-box optimization problems. Usually DE shows good global convergence, but it is slow in the local convergence. A local search can efficiently find a local optimum with high accuracy but cannot identify a basin of the global optimum. In the study we have proposed a hybrid of DE and the Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm. In BFGS derivatives are substituted by their approximations using the finite difference method. We have tested 3 different schemes of the cooperation of DE and BFGS. Using a set of benchmark optimization problems, the experimental results have shown that the hybridization can improve the performance of the standard DE algorithm. At the same time the choice of the hybridization scheme affects the results.

2672-8834 © 2023 Published by European Publisher.

Keywords: Global optimization, black-box optimization, differential evolution, quasi-Newton methods, BFGS

1. Introduction

Many real-world global optimization problems are presented by a black-box model, therefore there is no information on properties of the objective function and its derivatives. In this case different blind search methods, heuristics, and metaheuristics are applied (Zhan et al., 2022). Evolutionary algorithms are one of the most popular and efficient metaheuristics for solving hard optimization problems. Evolutionary algorithms are stochastic and population-based approaches (Del Ser et al., 2019).

Differential evolution (DE) is one of search algorithms, which demonstrates the high performance in solving global black-box optimization problems (Storn & Price, 1997). DE uses the difference-based mutation of random solution vectors and demonstrates the ability of global exploration of the search space. When a basin of the global optima was found, DE converges to the optima using the selection operator (Das & Suganthan, 2011). Because of the stochastic behaviour, DE demonstrates poor local convergence in the basin of attraction. A local search can efficiently find a local optimum with high accuracy but cannot identify a basin of the global optimum. Broyden–Fletcher–Goldfarb–Shanno algorithm (BFGS) is known as one of the fastest local search algorithms (Fletcher, 1987). BFGS is a quasi-Newton method that uses approximation instead of computing the exact Hessian matrix. At the same time BFGS uses first order derivatives, which can be replaced with their numerical estimates.

The performance of evolutionary algorithms can be improved by their hybridization with other methods. One of the most popular approaches is a combination of local and global search algorithms [also known as memetic evolutionary algorithms (Hart et al., 2005; Vakhnin & Sopov, 2020)]. In the study we have proposed a hybrid of DE and the BFGS algorithm.

A cooperation of DE and BFGS in the hybrid can be implemented in different ways:

- using the BFGS method for the final improvement of the best-found solutions, obtained by the DE algorithm,
- using the BFGS method for the intermediate improvement of trial solutions in the main DE's loop,
- and alternating DE and BFGS methods during the run.

2. Problem Statement

The problem statement of global black-box optimization can be stated in the following way (1):

$$f(\mathbf{x}) \rightarrow \min_{\mathbf{x} \in R^n}, f: R^n \rightarrow R^1, \quad (1)$$

here $f(\mathbf{x})$ is an objective function, $f: R^n \rightarrow R^1$, $\mathbf{x} = (x_1, \dots, x_n)$ is an n -dimensional solution, the search space (S) is defined by box constrains $x_i^l \leq x_i \leq x_i^u$, $i = \overline{1, n}$, x_i^l and x_i^u are the lower and upper values.

The solution to the problem (1) is

$$\mathbf{x}^*: \forall \mathbf{x} \in S \subset R^N, f(\mathbf{x}^*) \leq f(\mathbf{x}). \quad (2)$$

We do not set any requirements for the objective function f and don't use any properties of the objective function and its derivatives in search algorithms. Thus, the objective function is viewed as a black-box model.

3. Research Questions

In the study we use implementations of DE and BFGS algorithm with pseudocodes presented in Table 1 and 2.

Table 1. Pseudocode of the DE algorithm

Line	Pseudocode
1	Generate an initial population randomly
2	while <i>stop criteria</i> is not satisfied do
3	Perform <i>mutation</i>
4	Perform <i>crossover</i> and get a <i>trial</i> vector
5	Perform <i>selection</i>
6	end while

Table 2. Pseudocode of the BFGS algorithm

Line	Pseudocode
1	Set a starting point, the accuracy, and initial approximation of Hessian
2	while <i>stop criteria</i> is not satisfied do
3	Calculate a <i>search direction</i> using the finite difference approximation of gradients
4	Calculate the <i>step length</i> along the direction using the line search for satisfying Wolfe conditions
5	Update the Hessian approximation
6	end while

Three hybrids of DE and BFGS mentioned above have been implemented as described below:

- The first algorithm (titled as DE_BFGS 1) performs DE and BFGS consequentially. The best-found solution from the DE run is used as the starting point in the BFGS algorithm for making the final local improvement.
- The second algorithm (DE_BFGS 2) runs the BFGS algorithm after performing crossover (step 4 in Table 1) for each trial solution.
- The third algorithm (DE_BFGS 3) alternates DE and BFGS methods. We apply the BFGS algorithm after every 10 generations of DE. DE continues its evolution with the updated population.

Because of many variants of hybrid implementations, it's necessary to estimate the performance of them and to choose the best one.

4. Purpose of the Study

The purpose of the study is the experimental comparison of the proposed ways of hybridization. For making that, it's necessary to complete software implementation of algorithms, to obtain statistical data on numerical experiments, and to proceed the data for establishing the best hybrid.

5. Research Methods

In this study, we have compared four different approaches, including the original DE algorithm and three hybrid schemes. All of them were tested on 4 benchmark problems: Ackley, Griewank, Rastrigin, and Rosenbrock functions (Awad et al., 2016).

The number of variables of each problem is equal to 10 and 20. We have used the maximum number of fitness evaluations as the stopping criteria. The maximum number of fitness evaluations is 200000 for problems with 10 variables and 1000000 for problems with 20 variables. All algorithms are evaluated over 40 independent runs. The comparison of algorithms is based on medians of the best-found values and errors. The error measure is used to show the distance from the known global optimum if an algorithm hasn't reached the minimum of the objective function.

All algorithms in the comparison are implemented using the Python programming language with no external packages. We have used a multi-CPU parallelization for improving the computational performance in our numerical experiments.

6. Findings

The results of experiments are presented in Tables 3 and 4.

Table 3. The experimental results for problems with 10 variables

Algorithm	DE		DE_BFGS 1		DE_BFGS 2		DE_BFGS 3	
Function	f	Error	f	Error	f	Error	f	Error
Ackley	3.1E-15	4.2E-15	9.0E-18	7.1E-18	1.8E+1	3.7E+1	1.3E-6	3.2E-6
Griewank	7.4E-3	5.4E+0	7.3E-3	5.4E+0	5.2E-8	1.2E-3	9.1E-15	4.5E-8
Rastrigin	1.9E+0	1.4E+0	9.4E-1	1.0E+0	4.9E+1	6.9E+0	9.9E+0	3.1E+0
Rosenbrock	6.1E-28	1.4E-15	6.1E-4	4.0E-2	3.4E+5	1.2E+1	2.3E+4	7.9E+0

Table 4. The experimental results for problems with 20 variables

Algorithm	DE		DE_BFGS 1		DE_BFGS 2		DE_BFGS 3	
Function	f	Error	f	Error	f	Error	f	Error
Ackley	4.1E-9	6.3E-9	3.1E-12	3.7E-12	1.8E+1	5.6E+1	4.9E-3	6.3E-3
Griewank	6.6E-2	1.3E+1	3.7E-2	1.3E+1	1.1E-2	6.6E+0	1.0E-13	2.8E-7
Rastrigin	2.3E+1	3.2E+1	8.4E+0	2.8E+0	1.10E+2	1.0E+1	4.2E+1	6.5E+0
Rosenbrock	1.9E-4	1.0E-3	5.5E-3	3.9E-2	1.2E+6	1.9E+1	3.4E+5	1.4E+1

Figures 1-4 shows the variance of the results in the runs using box-plot diagrams. As we can see from figures 1, 2, 3, and 4, DE_BFGS 1 and DE_BFGS 3 give more stable results. At the same time, DE_BFGS 1 wins 2 of 4 times on hard multimodal problems. The standard DE outperforms hybrids with the Rosenbrock function, this can indicate that BFGS performs not well with ravines.

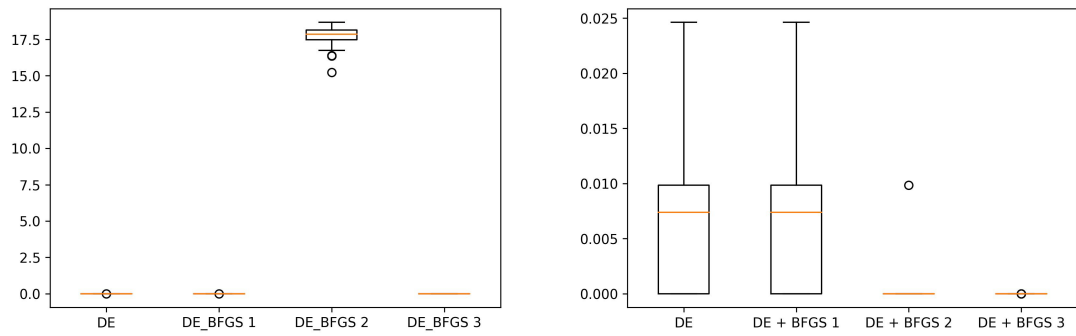


Figure 1. The variance of the results on Ackley (left) and Griewank (right) with 10 variables

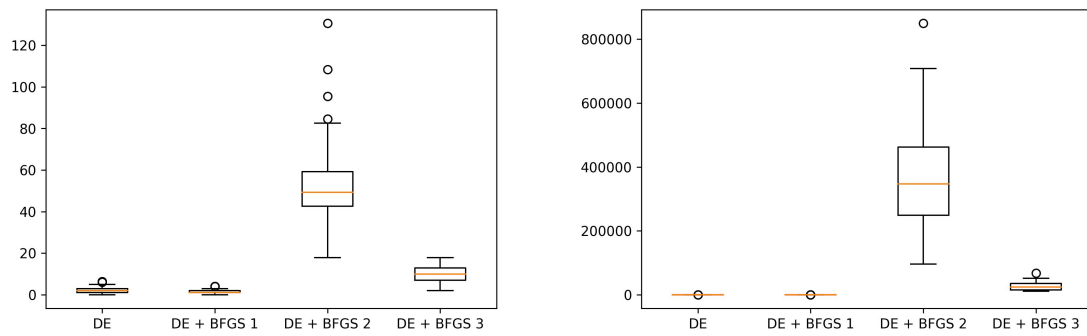


Figure 2. The variance of the results on Rastrigin (left) and Rosenbrock (right) with 10 variables

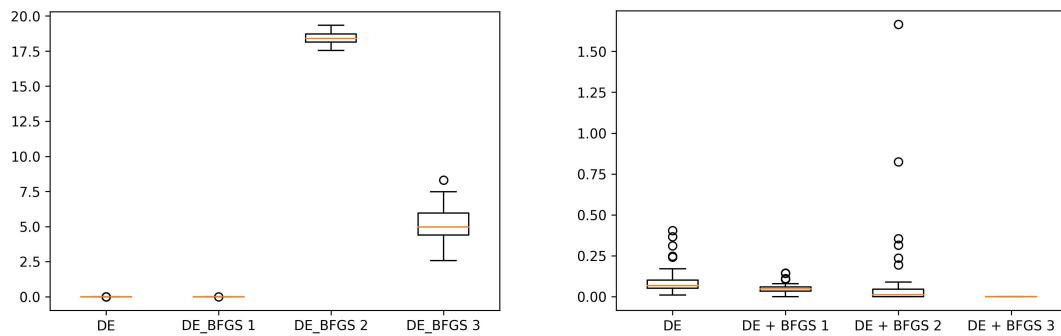


Figure 3. The variance of the results on Ackley (left) and Griewank (right) with 20 variables

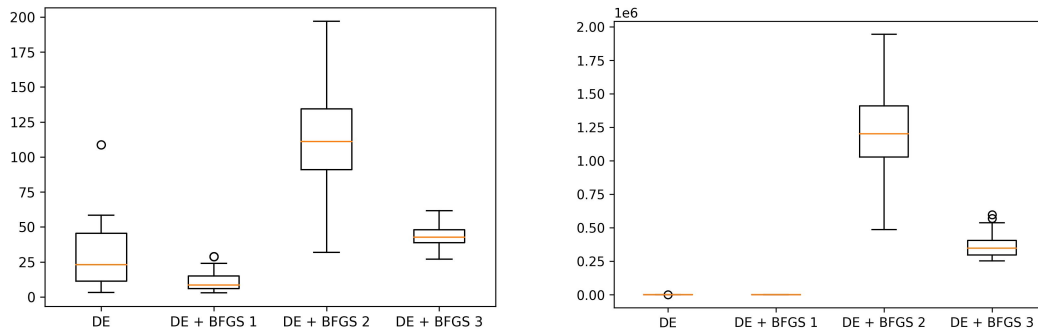


Figure 4. The variance of the results on Rastrigin (left) and Rosenbrock (right) with 20 variables

We have also visualized the convergence in one run to demonstrate how algorithms improve the best-found solution. An example on the Rastrigin problem with 10 variables is shown in Figure 5.

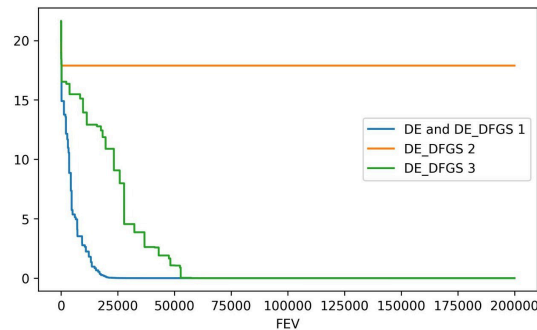


Figure 5. The convergence plot on the Rastrigin problem

Plots for DE and DE_BFGS 1 are combined, because DE_BFGS 1 doesn't use BFGS in the main run. As we can see in Figure 5, DE_BFGS 2 uses too greedy search and cannot escape a local optimum. DE_BFGS 3 can be slower than DE, but as we can see from the average results, it outperforms DE.

7. Conclusion

In the study, we have compared the performance of different schemes of combining global and local search in order to estimate how the hybridization affects the performance of DE. BFGS have been used as a local search in the hybrid. The experimental results have shown that the hybridization improves the performance on hard multimodal problems but yields to the standard DE on the unimodal problem with a ravine.

DE_BFGS 2 has shown the lowest performance, because the intensive local search leads to premature convergence. DE_BFGS 1 has shown that the original DE converges not well in the basin of attraction of the identified optimum, thus the usage of a local search after the run of DE is necessary. DE_BFGS 3 has a potential for improvement, because there are many different options for setting its parameters.

In further work, we will try different ways of implementing DE_BFGS 3, including a self-adaptive scheme. Also, the standard DE can be substituted by more advanced global search algorithms (for example, SaDE or SHADE). Finally, the set of benchmark problems can be extended.

Acknowledgments

This research was funded by the Ministry of Science and Higher Education of the Russian Federation, Grant No. 075-15-2022-1121.

References

- Awad, N., Ali, M., Liang, J., Qu, B., & Suganthan, P. (2016). *Problem definitions and evaluation criteria for the cec 2017 special session and competition on single objective bound constrained real-parameter numerical optimization*. Technical Report, Nanyang Technological University, Singapore.
- Das, S., & Suganthan, P. N. (2011). Differential Evolution: A Survey of the State-of-the-Art. *In IEEE Transactions on Evolutionary Computation*, 15(1), 4-31. <https://doi.org/10.1109/tevc.2010.205903>
- Del Ser, J., Osaba, E., Molina, D., Yang, X.-S., Salcedo-Sanz, S., Camacho, D., Das, S., Suganthan, P. N., Coello Coello, C. A., & Herrera, F. (2019). Bio-inspired computation: Where we stand and what's next. *Swarm and Evolutionary Computation*, 48, 220-250. <https://doi.org/10.1016/j.swevo.2019.04.008>
- Fletcher, R. (1987). *Practical methods of optimization* (2nd ed.). John Wiley & Sons.
- Hart, W. E., Krasnogor, N., & Smith, J. E. (2005). Memetic Evolutionary Algorithms. In W. E. Hart, J. E. Smith, & N. Krasnogor (Eds.), *Recent Advances in Memetic Algorithms. Studies in Fuzziness and Soft Computing* (vol. 166, pp. 3-27). Springer. https://doi.org/10.1007/3-540-32363-5_1
- Storn, R., & Price, K. (1997). Differential Evolution – A Simple and Efficient Heuristic for global Optimization over Continuous Spaces. *Journal of Global Optimization*, 11, 341-359. <https://doi.org/10.1023/A:1008202821328>
- Vakhnin, A., & Sopov, E. (2020). Investigation of the iCC Framework Performance for Solving Constrained LSGO Problems. *Algorithms*, 13(5). <https://doi.org/10.3390/A13050108>
- Zhan, Z.-H., Shi, L., Tan, K. C., & Zhang, J. (2022). A survey on evolutionary computation for complex continuous optimization. *Artificial Intelligence Review*, 55(1), 59-110. <https://doi.org/10.1007/s10462-021-10042-y>