MVMO with Opposite Gradient Initialization for Single Objective Problems

Thirachit Saenphon(✉)

Chulalongkorn University, Bangkok, Thailand
thirachits@gmail.com

Abstract. The objective of this paper is to describe an opposite gradient initialization concept with mean-variance mapping optimization (OGI-MVMO). OGI-MVMO is an optimization based on the actual manifold of objective function whereas original MVMO based stochastic optimization. Generating the new candidate solution to speed up the solution finding and accuracy of solution are important purposes. The OGI-MVMO algorithm consist of 2 steps: the primary step is generating new solution by OGI and also the second step is mutation between every of selected candidate solution supported the mean and variance of the population. The results showed that OGI-MVMO algorithm has better performance than other algorithm include the original MVMO for 15 real-parameter single objective functions.

Keywords: Continuous function
Mean-variance mapping optimization
Opposite gradient initialization search · Optimization

1 Introduction

The black-box continuous optimization is solving a global optimization solution without explicit knowledge of the form or structure of the objective function. Mean-variance mapping optimization (MVMO) [4] is a mapping function described by the mean and shape variables which both of them are derived from n the best solutions save in the specific archive. MVMO applied to solve the optimization problem such as wind farm [1], electricity pricing [6]. To generate a new offspring before the mutation, the technique based on MVMO also remain use the randomness without consider the interesting population on the real searching surface of the objective function.

An importance task in evolutionary algorithms is a generating first population with neighboring to an expect best solution. Random initialization is appreciated for generating new candidate solutions and selecting low scoring solutions to maximize or minimize the objective function. The geometric structure of the

P. Cong Vinh and V. Alagar (Eds.): ICCASA 2018/ICTCC 2018, LNICST 266, pp. 126–135, 2019.
https://doi.org/10.1007/978-3-030-06152-4_11

objective function is not relevant as a part of solution to generate a new offspring and searching process. The use of searching solutions on a manifold of objective function using opposite gradient search was first introduced in [7]. Fast Opposite Gradient Search (FOGS) does not depend on meta-heuristic search as the others' but it searches the manifold to find the locations with zero gradients and minimum values of objective function.

A new combining approach to improve the accuracy of solution with the better initial population for solving the single objective optimization problem was proposed. The research present an opposite gradient to generate some new offspring and searching a better solution from these first population with mean-variance mapping technique. The technique is called the opposite gradient initialization combined with mean-variance mapping optimization (OGI-MVMO). The method focused on how to enhance a step of generating a new offspring and how to search the better solution. This research also applied mapping function for mutation operation on the basic of the mean and variance of the n-best solutions to adjust the better candidate solution in the black-box problem functions which their surface are difficult and alter dependence on the composite function.

This article has the following sections; Sect. 2, Mean-Variance Mapping Optimization, Opposite Gradient Initialization and the combining OGI and MVMO algorithm are described. Experimental result and analysis are given in Sect. 3. Discussion the experimental results are given in Sect. 4. Finally, the conclusion is given in Sect. 5.

2 Concept of Proposed Algorithm

Our proposed algorithm consists of two main concepts. The first concept concerns the generating new offspring along the manifold of the objective function. The second concept focuses on search of the best solution using an adaptation of mean-variance mapping optimization. The detail of each concept is in the following sections.

2.1 Concept of Opposite Gradient Initialization

For a vector in a D-dimensional space of the objective function, OGI tries to generate some new offspring in the locations of the manifold whose the first derivative $(F'(x))$ are approximate zero since these locations must be the best solution. $\nabla F(x)$ is the unique vector field that satisfies $F'(x) \approx \nabla F(x)$ for vector field x.

Let $\mathbf{P}^{(\alpha)}$ and $\mathbf{P}^{(\beta)}$ be any two vectors on the manifold of the objective function $F(\mathbf{P}^{(i)})$. The point with zero of the gradient must lay on between $\mathbf{P}^{(\alpha)}$ and $\mathbf{P}^{(\beta)}$ if $\nabla F(\mathbf{P}^{(\alpha)})$ and $\nabla F(\mathbf{P}^{(\beta)})$ are difference sign value. A new vector can be computed from a distance δ. The value of δ can be computed by the following equation.

$$\delta = \frac{|\nabla F(\mathbf{P}^{(\beta)})|}{|\nabla F(\mathbf{P}^{(\alpha)})| + |\nabla F(\mathbf{P}^{(\beta)})|} \times ||\mathbf{P}^{(\beta)} - \mathbf{P}^{(\alpha)}|| \times w \qquad (1)$$

A constant $w \in R^+$. After computing δ, two new vectors are generated and computed from $\mathbf{P}^{(\alpha)} + \delta$ and $\mathbf{P}^{(\beta)} - \delta$. These two new vectors are the new offspring which used to be the first population for MVMO in the next step. The principal procedure of the proposed algorithm is as follows.

Algorithm 1. Proposed hybrid opposite gradient initialization with mean-variance mapping algorithm

```
Initialize algorithm parameters and Generation of initial population
while Terminating conditions are not satisfied do
    NP is a size of Q
    for 1 ≤ k ≤ NP do
        Fitness evaluation F(P^(α)).
        Fill or Update individual point using two archives.
        Classification of good gradient vectors and good fitness values.
        Fill the good gradient vectors in the first archive and some vectors with good
        fitness values in the second.
        Using opposite gradient initialization algorithm as described in Algorithm 2
        to compute new set of candidate solutions.
        Fill or Update individual point using solution archives.
        Mutation through mapping of m selected dimensions using local mean and
        variance.
    end for
end while
```

The main procedure begins with an initialization step where the algorithm parameter settings are defined and the first generation of population is generated by OGI within their search boundaries for a set of total number of candidate solutions. It ensures that the generated offspring will always in the search boundaries before fitness evaluation or local search execute. Furthermore, the first generated offspring is relative a local minimum. The heart of the algorithm is contained in the while loop in which for each fitness evaluation of each point, local search, updating the solution archive, fitness and gradient classification of points into selected parents point and offspring generation are performed. The main algorithm is terminated when the termination criterion is satisfied.

According to continually-updated archive in MVMO technique [4], each point has a fixed size compact memory that continuously fill and updated solution archive associated to it, the data stored in the archive are n-best offspring in a descending order of fitness to be knowledge for guiding the search direction. Only a new better solution of each point in the archive can replace later on every fitness evaluation or local search.

In this step of algorithm, the classification of good opposite gradient point and good fitness value for selection to be parents in the opposite gradient algorithm is included. In each generation, two new vectors are generated and lay in between two vectors of opposite gradients in the previous generation. Therefore, the new vectors are computed and lay on the reducing search space. If one of them gives a better cost function, then this new vector along with another vector in the first generation, whose value of cost function is in an acceptable range and its gradient that is opposite to the new vector, is used to generate a new vector in the next generation. Otherwise, any two vectors in the first generation, whose values of cost function are in an acceptable range and their gradients are opposite to each other, are newly selected to generate two new vectors in the second generation.

All $\mathcal{NP}$ vectors are ranked according to their local the best cost function and classified into two groups: Let $\mathcal{G}^+$ be sets of vectors whose gradients of cost function are positive and $\mathcal{G}^-$ be those whose gradients of cost function are negative. All vectors in $\mathcal{G}^+$ and $\mathcal{G}^-$ are already sorted in descending order according to their values of gradient. At the end of algorithm, the locations with zero gradient will be obtained. The detail of this step is given in Algorithm 2.

Algorithm 2. Opposite Gradient Initialization algorithm for generating new offspring

1: Set $count = 1$
2: Separate the vector in archive to $\mathcal{G}^+$ and $\mathcal{G}^-$ groups.
3: **while** $count \leq \mathcal{NP}$ **do**
4: Let $\mathbf{P}^{(\alpha)}$ be the first vector of $\mathcal{G}^+$ and $\mathbf{P}^{(\beta)}$ be the first vector of $\mathcal{G}^-$.
5: Compute vector $\mathbf{P}^{(1)}$ from $\mathbf{P}^{(\alpha)}$ and $\mathbf{P}^{(2)}$ from $\mathbf{P}^{(\beta)}$ by using equation (1).
6: Replace $\mathbf{P}^{(\alpha)}$ with $\mathbf{P}^{(1)}$ in $\mathcal{G}^+$ **if** $|\nabla F(\mathbf{P}^{(1)})| < |\mathbf{P}^{(\alpha)}|$.
7: Replace $\mathbf{P}^{(\beta)}$ with $\mathbf{P}^{(2)}$ in $\mathcal{G}^-$ **if** $|\nabla F(\mathbf{P}^{(2)})| < |\mathbf{P}^{(\beta)}|$.
8: Set $count = count + 1$.
9: **end while**

The new offspring from this step is filled and updated in solution archive, which stores its n-best offspring in descending order of fitness and serve for guiding the search direction. The size of solution archive is not vary along the entire process. After every point evaluate their fitness and local search, each point update its archive takes place only if the new solution is better than those in the archive.

2.2 Combined Mean-Variance Mapping Optimization

The objective of adapting Mean-Variance Optimization as a part of our algorithm is to improve the solution searching process so that the possible best solution can be found in fewer generations. The steps of procedure of MVMO for combine with OGI for the continuous problem are described as follows:

3 Experimental Simulation and Analysis

The proposed algorithm and the three other algorithms (NBIPOPaCMA [2], PVADE [5], MVMO [4]) were implemented for testing the performance of these

Algorithm 3. Combine the MVMO with OGI searching process

1: Set the parameters including $iter_{max}$, d_r, $\triangle d_0$, $f^*_{s\_ini}$, $f^*_{s\_final}$, number of points (NP)
2: Get the initial points which in the range [-100,100] from OGI step algorithm 2.
3: Set $k = 1$, k denotes point counters
4: **while** $iter_{max}$ is not reached **do**
5:   Calculate the cost function f for the problem, store f_{best} and x_{best} in archive.
6:   Increase $i = i + 1$
7:   **if** $i < iter_{max}$ **then**
8:     Check the point for the global best, collect a set of individual solutions.
9:     The $i - th$ point is discarded from the optimization process if the f_i is greater than every f_n in archive
10:     **if** the point is deleted **then**
11:       Increase $k = k + 1$ and go to step 5
12:     **else**
13:       Create offspring generation with algorithm 2
14:     **end if**
15:   **else**
16:     Create offspring generation with algorithm 2
17:   **end if**
18:   **if** $k < NP$ **then**
19:     increase $k = k + 1$ and go to step 5
20:   **end if**
21: **end while**

algorithms, the experimental results were compared with those of the other three algorithms based on 15 test functions on Real-Parameter Single Objective Optimization [3]. In Liang et al. [3], the complete description of the problem definition functions can be found. The set of test functions is uni-modal function. Search range is $[-100, 100]^D$. The optimization is terminated upon completion of the maximum number of function evaluations. In this section, the description of the parameter set-up (Sect. 3.1), the experimental results (Sect. 3.2) are presented.

3.1 Parameter Set-Up

All optimization problems are deciphered by utilizing OGI-MVMO with 150 points. The selected parameters for OGI-MVMO were summarized in Table 1. The first column denotes the names of parameters and the second column shows the value of each parameter.

3.2 Experimental Results

This paper present best, worst, median, mean and standard deviation of the error value between the best fitness values. The results are shown in Tables 2 and 3 respectively.

Table 1. Parameter setting of our proposed algorithm.

Parameters	Values
$f^*_{s\_ini}$	1
$f^*_{s\_final}$	20
γ	15
d_r	1
δd_0	0.05
NP	150
Archive size	5
Maximum number of function evaluations $iter_{max}$	100000
Number of experimental runs in each function	50

OGI-MVMO has capability to search the solution with zero error values for all function in all runs for 10D. The better result for each problem is highlighted in boldface. The result of tests show that OGI-MVMO can successfully solve problems without considering the local search strategy. Thus, OGI-MVMO is an appropriate technique to solve continuous problems.

Solving basic multi-modal functions: For 10D case issues, OGI-MVMO has capability to achieve the results with zero error values for all test function problem except F14, F15. However, the accuracy of the results that are obtained also are terribly satisfactory compared to the compared algorithms.

From the previous tables, OGI-MVMO outperformed others on 15 single objective functions while also get the good results as other 3 techniques on 10 function problems. It means using an opposite gradient initialization, rather than using a random start of the population, close to the best answer, and speed up the search. The examples for performance comparison are shown in Fig. 1. The figures have been also plotted average error (compared to the optimal solution) vs. the number of evaluations (for $D = 10$ dimensions) in Fig. 1. Note that the y-axis (average best value for all 50 run times) is logarithmic scale. Experiments have been repeated 50 times to plot the average error values. (a)–(c) F9, F14, and F15 with $D = 10$.

4 Discussion

The present methodology works on 2 significant ideas. The primary concept is to applied the manifold of objective function perform as a component of initializing and generating a candidate solution. The subsequent generation of solution is created based on the gradient of manifold of objective function. The last significant concept is searching the improve solution by collect the solution in an archive for locating the mean and variance to map a brand new probability of the better solution. These 2 ideas perform along well to attain since the geometrical structure of the cost function manifold and can relieve some problem of the important issues.

Table 2. Result of the comparison from different algorithms for $F1 - F5$

Algorithms	Best	Worst	Median	Mean	Std.
F1					
NBIPOPaCMA	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
PVADE	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
MVMO	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
OGI-MVMO	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
F2					
NBIPOPaCMA	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
PVADE	0.00E+00	7.97E+02	0.00E+00	1.57E+01	1.12E+02
MVMO	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
OGI-MVMO	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
F3					
NBIPOPaCMA	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
PVADE	0.00E+00	5.03E-03	0.00E+00	9.87E-05	7.05E-04
MVMO	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
OGI-MVMO	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
F4					
NBIPOPaCMA	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
PVADE	0.00E+00	2.25E-01	0.00E+00	4.41E-03	3.15E-02
MVMO	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
OGI-MVMO	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
F5					
NBIPOPaCMA	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
PVADE	0.00E+00	8.66E-04	0.00E+00	1.70E-05	1.21E-04
MVMO	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
OGI-MVMO	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00

Table 3. Result of the comparison from different algorithms for *F6–F15*

Algorithms	Best	Worst	Median	Mean	Std.
F6					
NBIPOPaCMA	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
PVADE	0.00E+00	9.89E+00	9.81E+00	7.09E+00	4.25E+00
MVMO	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
OGI-MVMO	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00

(*continued*)

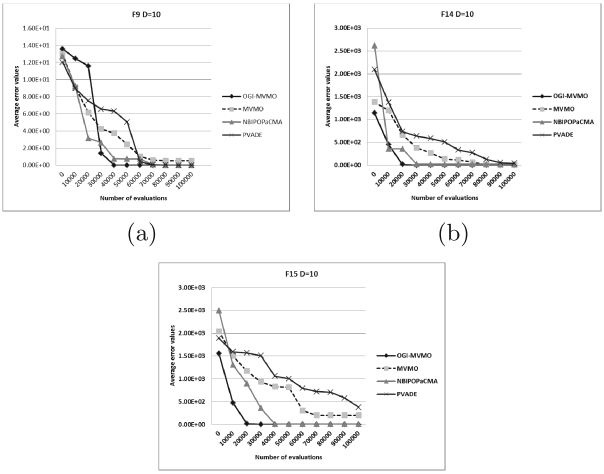
Table 3. (*continued*)

Algorithms	Best	Worst	Median	Mean	Std.
F7					
NBIPOPaCMA	0.00E+00	1.60E+01	3.27E−08	0.14E+01	0.48E+01
PVADE	3.13E−05	1.08E+01	2.64E−03	3.22E−01	1.53E+00
MVMO	6.34E−03	6.34E−03	6.34E−03	6.34E−03	0.00E+00
OGI−MVMO	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
F8					
NBIPOPaCMA	3.63E−08	1.70E−08	1.03E−08	1.03E−08	9.48E−09
PVADE	2.01E+01	2.05E+01	2.04E+01	2.03+01	7.12E−02
MVMO	2.02E+01	2.02E+01	2.02E+01	2.02E+01	8.93E−02
OGI−MVMO	**0.00E+00**	6.21E−08	4.34E−08	1.54E−07	1.0253E−07
F9					
NBIPOPaCMA	1.94E−08	2.53E+01	8.83E−01	3.85E+00	9.48E−09
PVADE	8.15E−03	3.66E+00	1.44E+00	1.56E+00	9.77E−01
MVMO	5.20E−01	2.43E+00	8.50E−01	8.32E−01	6.81E−01
OGI−MVMO	**0.00E+00**	9.66E−08	4.39E−08	7.15E−08	4.69E−08
F10					
NBIPOPaCMA	1.00E−08	6.70E−08	2.88E−08	3.26E−08	1.91E−09
PVADE	0.00E+00	1.77E−01	3.94E−02	5.05E−02	3.82E−02
MVMO	1.01E−02	3.69E−02	1.36E−02	1.67E−02	2.18E−02
OGI−MVMO	**0.00E+00**	3.94E−08	2.98E−08	3.04E−08	7.88E−09
F11					
NBIPOPaCMA	0.00E+00	6.70E−08	2.88E−08	3.26E−08	1.91E−09
PVADE	0.00E+00	1.19E+01	3.39E+00	3.94E+00	2.34E+00
MVMO	0.00E+00	5.97E+00	0.00E+00	2.25E+00	1.38E−02
OGI−MVMO	0.00E+00	1.11E−08	4.04E−09	4.89E−09	3.75E−09
F12					
NBIPOPaCMA	0.00E+00	1.05E+00	1.50E−04	3.50E−01	1.91E−09
PVADE	9.95E−01	1.61E+01	4.97E+00	5.92E+00	3.69E+00
MVMO	1.98E+00	1.57E+01	5.97E+00	5.95E+00	1.38E+00
OGI−MVMO	0.00E+00	3.56E−05	4.41E−08	4.80E−08	7.64E−09
F13					
NBIPOPaCMA	0.00E+00	5.68E+00	2.74E−04	7.8E−01	1.77E+00
PVADE	0.00E+00	2.13E+01	8.07E+00	8.63E+00	4.79E+00
MVMO	1.98e+00	1.97e+01	5.69E+00	9.14e+00	9.36E+00
OGI−MVMO	0.00E+00	3.82E−05	4.76E−02	1.09E−05	1.37E−05

(*continued*)

Table 3. (*continued*)

Algorithms	Best	Worst	Median	Mean	Std.
F14					
NBIPOPaCMA	2.19E+01	7.67E+02	3.77E+02	3.52E+02	2.14E+02
PVADE	4.13E+01	5.28E+02	1.55E+02	1.79E+02	1.09E+02
MVMO	3.41E+00	2.18E+01	3.72E+00	8.05E+00	7.68E+00
OGI−MVMO	**2.02E−05**	2.80E−05	2.68E−05	2.57E−05	4.86E−06
F15					
NBIPOPaCMA	6.89E+00	2.41E+02	1.05E+02	1.19E+02	9.01E+01
PVADE	3.74E+02	1.13E+03	7.96E+02	7.85E+02	1.72E+02
MVMO	1.97E+02	6.67E+02	5.48E+02	5.62E+02	9.99E+01
OGI−MVMO	**2.69E−05**	7.31E−02	6.85E−02	3.23E−02	3.44E−02



(a) (b) (c)

Fig. 1. Examples for performance comparison between OGI-MVMO, NBIPOPaCMA, PVADE, MVMO. F9, F14, and F15 are selected because the OGI-MVMO outperforms others.

5 Conclusion

This paper proposed an algorithm that combines the concept of opposite gradient initiation and the mean variance mapping optimization(OGI-MVMO) for solving the single continuous objective function problems. The algorithm contributed two important issues. The first issue is the opposite gradient initialization technique on the manifold of objective function. The opposite gradient analyses the manifold and generate a new offspring to the global search. The step of mean-variance mapping optimization is the second issue to utilize the result from opposite gradient initialization implementation to enhance the power of global searching. From the experimental results, the best solution quality and average solution quality of OGI-MVMO algorithm showed that OGI-MVMO is attractive for solving single objective functions.

References

1. Erlich, I., Shewarega, F., Feltes, C., Koch, F., Fortmann, J.: Determination of dynamic wind farm equivalents using heuristic optimization. In: Proceedings of 2012 IEEE Power and Energy Society General Meeting, San Diego, USA (2012)
2. Loshchilov, I.: CMA-ES with restarts for solving CEC 2013 benchmark problems. In: Proceedings of Proceedings of IEEE Congress on Evolutionary Computation (CEC), Cancun, Mexico, pp. 369–376 (2013)
3. Liang, J., Qu, B.Y., Suganthan, P.N., Hernandez-Diaz, A.G.: Problem definitions and evaluation criteria for the CEC 2013 special session on real-parameter optimization. Technical reportTR2012, Cancun, Mexico (2013)
4. Rueda, J.L., Erlich, I.: Hybrid mean-variance mapping optimization for solving the IEEE-CEC 2013 competition problems. In: Proceedings of IEEE Congress on Evolutionary Computation (CEC), Cancun, Mexico, pp. 1664–1671 (2013)
5. Coelho, L.D., Ayala, H.V.H., Freire, R.Z.: Population's variance-based Adaptive Differential Evolution for real parameter optimization. In: Proceedings of IEEE Congress on Evolutionary Computation (CEC), Cancun, Mexico, pp. 1672–1677 (2013)
6. Holtschneider, T., Erlich, I.: Optimization of electricity pricing considering neural network based model of consumers' demand response. In: Proceedings of 2013 IEEE Symposium Series on Computational Intelligence, Singapore (2013)
7. Saenphon, T., Phimoltares, S., Lursinsap, C.: Combining new Fast Opposite Gradient Search with Ant Colony Optimization for solving travelling salesman problem. Eng. Appl. Artif. Intell. **35**, 324–334 (2014)