Theoretical Background for Many-Objective Optimization

Monalisa Pal, Sriparna Saha and Sanghamitra Bandyopadhyay

1 Introduction

A brief description of the chief terminologies is assembled in this material for easy comprehension of many-objective optimization algorithms. For a more detailed description of the concepts, the reference accompanying the terminologies should be looked up.

The material describes the following concepts: definition of many-objective optimization problems in Section 2, pareto-optimality and related notions in Section 3, non-conflicting objective set in context of objective reduction in Section 4, steps of Differential Evolution for Multi-objective Optimization in Section 5, various benchmark problems to test and compare the performance proposed algorithms with existing algorithms in Section 6 and several performance metrics to asess the results from the optimization algorithms in Section 7.

2 Many-objective Optimization Problem

The optimization problems dealing with multiple objectives form the multi-objective optimization domain. It involes the mapping of an N-dimensional vector $X = [x_1, x_2, \cdots, x_N]^T$ in the decision space i.e. $X \in S$, to an M-dimensional vector $F(X) = [f_1(X), f_2(X), \cdots, f_M(X)]^T$ in the objective space such that searching the objective space yields a set of decision vectors which presents the optimal trade-off in terms of all the M-objectives. An M-objective minimization problem is mathematically formulated as in Eq. (1). It is also essential to list all the K equality constraints (Eq. (2)) and J inequality constraints (Eq. (3)) while defining the optimization problem [9]. Such a problem is called many-objective optimization (MaOO) problem when $M > 3$ [9].

Minimize

$$F(X) = \{f_1(X), f_2(X), \cdots, f_M(X)\} \tag{1}$$

subject to constraints (if any)

$$\begin{aligned} H(X) = \{h_1(X), h_2(X), \cdots, h_K(X)\} \\ \text{such that } h_k(X) = 0,\ k = 1, \cdots, K \end{aligned} \tag{2}$$

and

$$\begin{aligned} G(X) = \{g_1(X), g_2(X), \cdots, g_J(X)\} \\ \text{such that } g_j(X) \geq 0,\, j = 1, \cdots, J \end{aligned} \tag{3}$$

where

$$x_i^L \leq x_i \leq x_i^U,\, i = 1, \cdots, N \tag{4}$$

The common region which satisfies all the constraints (Eq. (2) and (3)) and is bounded by the lower and upper bounds (Eq. (4)) of the decision variables, creates the entire decision space S.

3 Concept of Pareto-Optimality

The optimal state of trade-off is called Pareto-optimality which is the state where improvement in performance for no objective is possible without deterioration with respect to any other objective(s) [9, 2]. Since, the objective is not scalar valued, two solution vectors cannot be compared based on greater than or less than relation. Pareto-dominance relation is commonly used for comparison of two solution vectors. Given two feasible solution vectors X and Y, if Eq. (5) follows, X Pareto-dominates Y i.e. $X \succ Y$.

$$\begin{aligned} \forall i \in \{1, \cdots, M\}, f_i(X) \leq f_i(Y) \text{ and} \\ \exists j \in \{1, \cdots, M\}, f_j(X) < f_j(Y) \end{aligned} \tag{5}$$

Pareto-Optimal Set (PS) is a set of solution vectors in S such that there exists no other solution vector dominating any constituent of PS. The set of objective vectors corresponding to the Pareto-Optimal Set (PS) forms the Pareto-Front (PF). MaOO algorithm provides an approximation in terms of PS and PF which is given by Eq. (6) and (7), respectively.

$$PS = \{X \in S \mid X \text{ is a Pareto-Optimal solution}\} \tag{6}$$

$$PF = \{y \in \mathbb{R}^M \mid y = F(X), X \in PS\} \tag{7}$$

4 Non-Conflicting Objective Set

Let X and Y be two solution vectors belonging to the decision space, then two out of the M objectives viz. $f_i(.)$ and $f_j(.)$ are related according to Eq. (8) when these are conflicting [2]. In case there is atleast one pair solutions (X and Y) which are non-dominated in terms of $f_i(.)$ and $f_j(.)$, then these two objectives are conflicting with each other. A more recent approach towards conflicting objectives is defined in terms of the Pareto-dominance relation ($\preceq_{F_i}$) induced by a objective set ($F_i()$). According to this, two sets of objectives viz. $F_i(.)$ and $F_j(.)$ are said to be conflicting when $\preceq_{F_i} \neq \preceq_{F_j}$ and thus leading to different

Pareto-Optimal Set. A Non-Conflicting Objective Set ($F'(.)$) [3] is a subset of the entire objective set i.e. $F' \subseteq F$, has a cardinality of $2 \leq |F'| \leq |F|$ and follows $\preceq_{F'} = \preceq_F$. This indicates that excluding $F - F'$ objectives will not alter the Pareto-dominance relation induced by F. Objective reduction techniques are employed to look for the smallest Non-Conflicting Objective Set.

$$\exists(X, Y) \text{ such that } (f_i(X) > f_j(Y)) \text{ and } (f_i(X) < f_j(Y)) \tag{8}$$

5 Differential Evolution for Multi-objective Optimization (DEMO)

Differential Evolution for Multi-objective Optimization (DEMO) [11] has four stages, similar to the single-objective Differential Evolution. These stages are Initialization, Mutation, Recombination and Selection. The first three stages are similar for both single-objective and multi-objective versions of the Differential Evolution. The fourth stage i.e., selection, is where the two approaches differ. A brief description of the DEMO model is presented next which is called DE/rand/1/bin.

5.1 Initialization

In this context, a population is represented by a matrix of order $NP \times N$ implying NP candidates where each candidate is a vector with N decision variables. For the initial population, the j-th decision variable of every i-th candidate vector is randomly initialized as per Eq. (9) where x_j^U and x_j^L are the upper and lower bounds of the j-th decision variable, respectively.

$$\begin{aligned} &x_{ij,0} = x_j^L + rand(0,1) \times \left(x_j^U - x_j^L\right) \\ &\text{where } i = 1, \cdots, NP \text{ and } j = 1, \cdots, N \end{aligned} \tag{9}$$

5.2 Mutation

For every generation G and for every i-th candidate, three indices r_1, r_2 and r_3, are randomly chosen such that i, r_1, r_2 and r_3 are mutually exclusive. The corresponding candidates are used to generate the mutant vector $V_{i,G}$ as shown in Eq. (10). The parameter, F_S is called the scale factor, and is a randomly chosen real value in the range $[0, 2]$.

$$\begin{aligned} &V_{i,G} = X_{r_1,G} + F_S \times (X_{r_2,G} - X_{r_3,G}) \\ &\text{where } i = 1, \cdots, NP \end{aligned} \tag{10}$$

5.3 Recombination

For every next generation $(G+1)$, the i-th trial vector, $U_{i,G+1}$, is generated by picking the j-th element from i-th mutant vector or the i-th candidate vector depending on the crossover rate (CR) according to Eq. (11). Usually a high value in the range $[0, 1]$ is chosen for CR so that constituents of the mutant vector, $v_{ij,G}$, have higher selection chances. The creation of the trial vector is constrained such that a random element (I_{rand}) of the trial vector is always same as that from the mutant vector. This ensures that the generated trial vector is not the same as the candidate vector.

$$u_{ij,G+1} = \begin{cases} v_{ij,G}, & \text{if } rand_{ij} \leq CR \text{ or } j = I_{rand} \\ x_{ij,G}, & \text{if } rand_{ij} > CR \text{ and } j \neq I_{rand} \end{cases} \tag{11}$$
$$\text{where } i = 1, \cdots, NP \text{ and } j = 1, \cdots, N$$

5.4 Selection

In this stage, a choice is made between the candidate vector, $X_{i,G}$, and the trial vector, $U_{i,G+1}$ based on Pareto-dominance relation to yield the candidate vector for the next generation, $X_{i,G+1}$. This is shown in Eq. (12).

$$X_{i,G+1} = \begin{cases} X_{i,G}, & \text{if } X_{i,G} \succ U_{i,G+1} \\ U_{i,G+1}, & \text{otherwise} \end{cases} \tag{12}$$
$$\text{where } i = 1, \cdots, NP$$

6 Benchmark Problems

The proposed MaOO algorithm which uses objective reduction is compared with other state-of-the-art MaOO algorithms based on the performance on a few test problems from the DTLZ test suite [7]. The performance analysis is done on unimodal problems like DTLZ2 and DTLZ4, and also on multimodal problems like DTLZ1 and DTLZ3 [7, 8]. The number of objectives is varied upto 20. A brief description of the test problem is as follows:

6.1 DTLZ1 Problem

M-objectives of a MaOO problem of this kind is given by:

$$
\begin{aligned}
&\text{Minimize } f_1(X) = \frac{1}{2}x_1 x_2 \cdots x_{M-1}(1 + g(X_M))\\
&\text{Minimize } f_2(X) = \frac{1}{2}x_1 x_2 \cdots (1 - x_{M-1})(1 + g(X_M))\\
&\vdots\\
&\text{Minimize } f_{M-1}(X) = \frac{1}{2}x_1(1 - x_2)(1 + g(X_M))\\
&\text{Minimize } f_M(X) = \frac{1}{2}(1 - x_1)(1 + g(X_M))\\
&\text{subject to } 0 \leq x_i \leq 1, \text{ for } i = 1, 2, \cdots, N\\
&\text{where, } g(X_M) =\\
&100\left[|X_M| + \sum_{x_i \in X_M} \left\{(x_i - 0.5)^2 - \cos(20\pi(x_i - 0.5))\right\}\right]
\end{aligned}
$$

Pareto-Front of this problem is given by $x_i^* = 0.5$ where $x_i^* \in X_M$ and $\sum_{i=1}^{M} f_i(X) = 0.5$. Hence, the Pareto-Front is linear. From [7], the suggested cardinality of set X_M is $k = |X_M| = 5$ which also dictates the number of decision variables i.e. $N = M + k - 1 = M + 4$.

6.2 DTLZ2 Problem

M-objectives of a MaOO problem of this kind is given by:

$$
\begin{aligned}
&\text{Minimize } f_1(X) = (1 + g(X_M))\cos\left(x_1\frac{\pi}{2}\right)\cdots\cos\left(x_{M-1}\frac{\pi}{2}\right)\\
&\text{Minimize } f_2(X) = (1 + g(X_M))\cos\left(x_1\frac{\pi}{2}\right)\cdots\sin\left(x_{M-1}\frac{\pi}{2}\right)\\
&\vdots\\
&\text{Minimize } f_{M-1}(X) = (1 + g(X_M))\cos\left(x_1\frac{\pi}{2}\right)\sin\left(x_2\frac{\pi}{2}\right)\\
&\text{Minimize } f_M(X) = (1 + g(X_M))\sin\left(x_1\frac{\pi}{2}\right)\\
&\text{subject to } 0 \leq x_i \leq 1, \text{ for } i = 1, 2, \cdots, N\\
&\text{where, } g(X_M) = \sum_{x_i \in X_M} (x_i - 0.5)^2
\end{aligned}
$$

Pareto-Front of this problem is given by $x_i^* = 0.5$, where $x_i^* \in X_M$ and $\sum_{i=1}^{M} f_i^2(X) = 1$. Hence, the Pareto-Front is concave. From [7], the suggested cardinality of set X_M is $k = |X_M| = 10$ which also dictates the number of decision variables i.e. $N = M + k - 1 = M + 9$.

6.3 DTLZ3 Problem

M-objectives of a MaOO problem of this kind is similar to DTLZ2 except the $g(.)$ function of DTLZ1 is considered. Pareto-Front of this problem is given by $x_i^* = 0.5$ where $x_i^* \in X_M$ and $\sum_{i=1}^{M} f_i^2(X) = 1$. Hence, the Pareto-Front is concave. From [7], the suggested cardinality of set X_M is $k = |X_M| = 10$ which also dictates the number of decision variables i.e. $N = M + k - 1 = M + 9$.

6.4 DTLZ4 Problem

M-objectives of a MaOO problem of this kind is a similar to the DTLZ2 problem except a meta-variable mapping ($x_i \to x_i^\alpha$) is considered. Hence, all the objectives are functions of x_i^α instead of x_i. Pareto-Front of this problem is given by $x_i^* = 0.5$ where $x_i^* \in X_M$ and $\sum_{i=1}^{M} f_i^2(X) = 1$. Hence, the Pareto-Front is concave. From [7], the suggested cardinality of set X_M is $k = |X_M| = 10$ which also dictates the number of decision variables i.e. $N = M + k - 1 = M + 9$. Also, [7] suggests $\alpha = 100$.

7 Performance Metrics

As visualisation of M-dimensional objective space is not possible when $M > 3$, the analysis of the Pareto-optimal state (more specifically, the Pareto-Front) can only be made on the basis of performance metrics. Two major characteristics which assess the quality of the Pareto-Front are convergence [2, 6] and diversity [2, 1]. Convergence indicates the closeness of the approximation of the Pareto-Front with the true scenario. On the other hand, diversity measures how well spread is the points constituting the approximated Pareto-Front over the entire surface. There are several metrics available in the literature for quantifying these two features of a Pareto-Front.

This work uses Convergence Metric and Hypervolume Indicator as performance metrics. Convergence metric provides information regarding convergence only, whereas, hypervolume indicator provides information regarding both convergence and diversity. The decision made by considering these two metrics simultaneously can often be conflicting [10], yet this simultaneous analysis, in presence of conflict, is helpful for pointing out whether a MaOO algorithm has poor convergence or poor diversity which is otherwise impossible by using either kind of metric exclusively. Other popular performance measures are generational distance (GD) [2, 5], and R2 indicator [2, 5]. But literature survey reveals that calculation of GD is similar to convergence metric whereas R2 indicator and hypervolume indicator are positively correlated [2, 4]. All these measures either need information about true Pareto-front or are biased by the choice of reference points.

7.1 Convergence Metric

Convergence Metric [2, 6], indicates the convergence of the approximated Pareto-Front, PF, to true Pareto-Front. A sampled version of the true Pareto-Front is used for this purpose. A set of points, H, is randomly and uniformly picked up from the true Pareto-Front. The distance, d_i, is obtained as the minimum distance between the i-th point of the approximated Pareto-Front $F_i \in PF$ and all the points of H. Convergence Metric of PF (δ_{PF}) is quantified as the arithmetic mean of d_i over all the points in PF as shown in Eq. (13) where $D_E(.)$ refers to Euclidean distance. For two Pareto-Fronts PF_1 and PF_2, when $\delta_{PF_1} < \delta_{PF_2}$, $PF_1 \succ PF_2$.

$$d_i = \min_{X \in H}(D_E(F_i, X)) \text{ and } \delta_{PF} = \frac{1}{|PF|} \sum_{i=1}^{|PF|} d_i \tag{13}$$

7.2 Hypervolume Indicator

A more robust measure is Hypervolume Indicator [2, 1] as it yields a scalar value depicting the details of convergence as well as diversity without requiring any knowledge of the true Pareto-Front. However, its major drawback is the sensitivity to the location of reference point, required for its evaluation. A hyper-rectangle is defined between the origin and the user-defined reference point, as the diagonally opposite points, in the objective space. To represent the hyper-rectangle, a set, H, of randomly sampled points is considered where the sampling is done by Monte-Carlo simulation. Finally, the metric is given by the ratio of the number of points of H which are Pareto-dominated by the points of the approximated Pareto-Front, PF, to the total number of points in H. An attainment function which is expressed by Eq. (14), aids in the evaluation and results in 1 when any point $\vec{z}$ of H is Pareto-dominated by any point of PF, otherwise 0. The average of the return values of the attainment function over the number of points in H, gives the Hypervolume Indicator. Among two Pareto-Fronts, PF_1 and PF_2, $PF_1 \succ PF_2$ when $I_H(PF_1) > I_H(PF_2)$.

$$\alpha_{PF}(\vec{z}) = \begin{cases} 1, & \text{if} PF \succ \{z\} \text{ where } \vec{z} \in H \\ 0, & \text{otherwise} \end{cases} \tag{14}$$

References

[1] J. Bader and E. Zitzler. Hype: An algorithm for fast hypervolume-based many-objective optimization. *Evolutionary computation*, 19(1):45–76, 2011.

[2] S. Bandyopadhyay and A. Mukherjee. An algorithm for many-objective optimization with reduced objective computations: A study in differential evolution. *Evolutionary Computation, IEEE Transactions on*, 19(3):400–413, 2015.

[3] D. Brockhoff and E. Zitzler. Objective reduction in evolutionary multiobjective optimization: theory and applications. *Evolutionary Computation*, 17(2):135–166, 2009.

[4] Dimo Brockhoff, Tobias Wagner, and Heike Trautmann. On the properties of the r2 indicator. In *Proceedings of the 14th annual conference on Genetic and evolutionary computation*, pages 465–472. ACM, 2012.

[5] Shelvin Chand and Markus Wagner. Evolutionary many-objective optimization: a quick-start guide. *Surveys in Operations Research and Management Science*, 20(2):35–42, 2015.

[6] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan. A fast and elitist multiobjective genetic algorithm: Nsga-ii. *Evolutionary Computation, IEEE Transactions on*, 6(2):182–197, 2002.

[7] K. Deb, L. Thiele, M. Laumanns, and E. Zitzler. Scalable multi-objective optimization test problems. In *Proceedings of the Congress on Evolutionary Computation (CEC-2002),(Honolulu, USA)*, pages 825–830. Proceedings of the Congress on Evolutionary Computation (CEC-2002),(Honolulu, USA), 2002.

[8] S. Huband, P. Hingston, L. Barone, and L. While. A review of multiobjective test problems and a scalable test problem toolkit. *Evolutionary Computation, IEEE Transactions on*, 10(5):477–506, 2006.

[9] H. Ishibuchi, N. Tsukamoto, and Y. Nojima. Evolutionary many-objective optimization: A short review. In *Evolutionary Computation, 2008. CEC 2008. (IEEE World Congress on Computational Intelligence). IEEE Congress on*, pages 2419–2426, 2008.

[10] Monalisa Pal and Sanghamitra Bandyopadhyay. Reliability of convergence metric and hypervolume indicator for many-objective optimization. In *Control, Instrumentation, Energy & Communication (CIEC), 2016 2nd International Conference on*, pages 511–515. IEEE, Jan 2016.

[11] T. Robi and B. Filipi. Demo: Differential evolution for multiobjective optimization. In CarlosA. Coello Coello, Arturo Hernndez Aguirre, and Eckart Zitzler, editors, *Evolutionary Multi-Criterion Optimization*, volume 3410 of *Lecture Notes in Computer Science*, pages 520–533. Springer Berlin Heidelberg, 2005.