Averaged Hausdorff Approximations of Pareto Fronts based on Multiobjective Estimation of Distribution Algorithms

Luis Martí
Department of Electrical Engineering
Pontifical Catholic University of Rio de Janeiro, Brazil
lmarti@ele.puc-rio.br

Christian Grimme
Information Systems and Statistics
University of Münster, Germany
christian.grimme@uni-muenster.de

Pascal Kerschke
Information Systems and Statistics
University of Münster, Germany
kerschke@uni-muenster.de

Heike Trautmann
Information Systems and Statistics
University of Münster, Germany
trautmann@uni-muenster.de

Günter Rudolph
Department of Computer Science
TU Dortmund University, Germany
guenter.rudolph@tu-dortmund.de

ABSTRACT

We propose a post-processing strategy which consists of applying the averaged Hausdorff indicator to the complete archive of solutions after optimization by multiobjective estimation of distribution algorithms (MEDAs) to select a uniformly distributed subset of non-dominated solutions.

Categories and Subject Descriptors

I.2.8 [**Artificial Intelligence**]: Problem Solving, Control Methods and Search; I.2.m [**Artificial Intelligence**]: Evolutionary Computing and Genetic Algorithms—*Multiobjective Evolutionary Algorithms*

Keywords

Multiobjective optimization, Averaged Hausdorff distance, Estimation of Distribution Algorithm

1. INTRODUCTION

A broad range of heuristics and metaheuristics has been used to address multiobjective problems (MOPs). Evolutionary multiobjective optimization algorithms (EMOAs) [2] have been found to be a competent approach in a wide variety of application domains. Alternatively, multiobjective estimation of distribution algorithms (MEDAs) [6] were introduced which aim at learning the problem structure and characteristics along the run and, hence, explore the search space in a more efficient manner [3]. MEDAs replace the application of evolutionary operators in the offspring generation process with the creation of a statistical model of the fittest elements of the population in a process known as model-building. This model is then sampled to produce new elements. Nevertheless, MEDAs have not lived up to their a priori expectations. This can be attributed to the fact that most MEDAs have limited themselves to transforming single-objective EDAs into a multiobjective formulation by including an existing multiobjective fitness assignment function. Additionally, the tendency of MEDAs loosing population diversity has been reported [1]. This situation is particularly dramatic in the multiobjective case, as diversity and homogeneity are among the desired features of the final non-dominated set. Here, we experimentally investigate how MEDAs perform compared to classical EMOAs and how especially MEDA results can be improved after the run by means of a post-processing approach in terms of equally spaced solutions on the non-dominated front.

GECCO '15 July 11-15, 2015, Madrid, Spain

ACM ISBN 978-1-4503-3488-4/15/07.
DOI: http://dx.doi.org/10.1145/2739482.2764631

2. POSTPROCESSING OF MEDA RESULTS

The *averaged Hausdorff distance* [7] can be used to assess whether a non-dominated front has a sufficiently good spread and a small distance to the true Pareto front.

DEFINITION 1. *Let $A, B \subset \mathbb{R}^M$ be non-empty finite sets. The value*

$$\Delta_p(A,B) = \max(\mathsf{GD}_p(A,B), \mathsf{IGD}_p(A,B)) \text{ with}$$

$$\mathsf{GD}_p(A,B) = \left(\frac{1}{|A|}\sum_{a\in A} d(a,B)^p\right)^{1/p} \text{ and}$$

$$\mathsf{IGD}_p(A,B) = \left(\frac{1}{|B|}\sum_{b\in B} d(b,A)^p\right)^{1/p}$$

for $p > 0$ is termed the averaged Hausdorff distance *between sets A and B where $d(u, A) := \inf\{\|u - v\| : v \in A\}$ for $u, v \in \mathbb{R}^M$ and some vector norm $\| \cdot \|$.*

Suppose that some MEDA has generated a non-dominated front for some MOP. Typically, the points on the front are not evenly distributed. Therefore, we propose the following *post-processing* approach:

1. Run your favorite MEDA with a tiny add-on: store each generated offspring in a file.
2. After termination of your favorite MEDA: construct an evenly spaced reference front from a non-dominated front (e.g. the last population); then feed each stored offspring into the Δ_p archive updater (see Alg. 1) sequentially; the final content of the archive A is the desired approximation.

For bi-objective problems the reference front R can be constructed as a linear interpolation from a given approximation of the Pareto front [4]. Alg. 1 shows a naive Δ_1 update operation.

Algorithm 1 Naive Δ_1-update

Require: archive set A, reference set R, new element $\boldsymbol{x}$
$A = \mathsf{ND}_f(A \cup \{\boldsymbol{x}\}, \preceq)$ // non-dom. filtering
if $|A| > N_R := |R|$ **then**
  for all $a \in A$ **do**
    $h(a) = \Delta_1(A \setminus \{a\}, R)$
  end for
  $A^* = \{a^* \in A : a^* = \mathsf{argmin}\{h(a) : a \in A\}\}$
  if $|A^*| > 1$ **then**
    $a^* = \mathsf{argmin}\{\mathsf{GD}_1(A \setminus \{a\}, R) : a \in A^*\}$
  end if
  $A = A \setminus \{a^*\}$
end if

For feeding the stored pairs $(\boldsymbol{x}, \boldsymbol{F}(\boldsymbol{x}))$ into the archive updater we use a 'forward update' in which individuals are added in the order of their generation. Often, they will pass the initial dominance check, so that subsequent Δ_p calculations are necessary. A more time saving 'backward update' adds stored pairs into the archive updater in inverted order. This way, most points from later stages of the inverted sequence will probably not pass the initial dominance check, which leads to less Δ_p calculations. Since the order of adding the points clearly affects the final outcome, we compare both approaches experimentally.

3. EXPERIMENTS AND RESULTS

For experimentation we used well known MOPs for benchmarking [4]: the sphere problem, DENT, ZDT3 and WFG1 which feature different convexity and concavity characteristics. Reference fronts covered by 1,000 uniformly spaced points were created based on a parametric form which allows exact calculation of the optimal fronts' length for all benchmark problems (except WFG1). Our evaluation is based on four state-of-the-art general purpose EMOAs and four MEDAs. As EMOAs, we employed NSGA2 and SMS-EMOA with standard parameters as well as PSEMOA and SCD-NSGA2 being special purpose methods that try to keep evenly spaced solutions. As MEDAs we use naive MIDEA and MO-CMA-ES as well as MONEDA and MARTEDA which are designed for diversity preservation. All test problems were optimized 20 times by all algorithms for 50,000 function evaluations each, population sizes $\mu \in \{10, 20, 100\}$, and $p \in \{1, 2\}$.

From a first MEDA/EMOA comparison regarding the overall algorithmic behavior it is notable that the MEDAs lead to very stable results over the repeated runs which is reflected by both hypervolume and Δ_p indicators while no MEDA is superior to the others. Thus, the MEDAs are at least competitive with the classical EMOAs on our test problems.

Applying the post-processing strategy to the results improves the final approximation quality of the MEDAs in terms of Δ_p for all population sizes. As the Pareto front of ZDT3 is disconnected, the performance of the considered post-processing differs from the remaining test problems, as the linear interpolation ignores the discontinuity in the interpolation and tends to place reference points also in unattainable regions.

Further, the direction of the archive update strategy has no influence on the final approximation quality. However, the backward strategy requires much less updating iterations until the selected subset becomes stable than the forward strategy. Therefore, we recommend to favor this approach over the forward update in general such that the post-processing will be much more computationally efficient.

An extended and more detailed paper containing the post-processing method, the experimental setup, and results can be found in [5].

4. REFERENCES

[1] C. W. Ahn and R. S. Ramakrishna. Multiobjective real-coded Bayesian optimization algorithm revisited: Diversity preservation. In *GECCO 2007*, pages 593–600, New York, NY, USA, 2007. ACM Press.

[2] C. A. Coello Coello, G. B. Lamont, and D. A. Van Veldhuizen. *Evolutionary Algorithms for Solving Multi-Objective Problems.* Genetic and Evolutionary Computation. Springer, New York, 2nd edition, 2007.

[3] D. W. Corne. Single objective = past, multiobjective = present, ??? = future. In Z. Michalewicz, editor, *2008 IEEE CEC*, Piscataway, New Jersey, 2008. IEEE Press.

[4] K. Gerstl, G. Rudolph, O. Schütze, and H. Trautmann. Finding evenly spaced fronts for multiobjective control via averaging Hausdorff-measure. In *Proceedings of 8th CCE*, pages 1–6. IEEE Press, 2011.

[5] L. Marti, C. Grimme, P. Kerschke, H. Trautmann, and G. Rudolph. Averaged Hausdorff approximations of pareto fronts based on multiobjective estimation of distribution algorithms. Technical report, arXiv.org, http://arxiv.org/abs/1503.07845, March 2015.

[6] M. Pelikan, K. Sastry, and D. E. Goldberg. Multiobjective estimation of distribution algorithms. In M. Pelikan, K. Sastry, and E. Cantú-Paz, editors, *Scalable Optimization via Probabilistic Modeling: From Algorithms to Applications*, Studies in Computational Intelligence, pages 223–248. Springer Verlag, 2006.

[7] O. Schütze, X. Esquivel, A. Lara, and C. A. Coello Coello. Using the averaged Hausdorff distance as a performance measure in evolutionary multiobjective optimization. *IEEE Transactions on Evolutionary Computation*, 16(4):504–522, 2012.