The Pareto -Box Problem for the Modelling of Evolutionary Multiobjective Optimization Algorithms

Chapter · January 2005

DOI: 10.1007/3-211-27389-1_46

Mario Köppen
Kyushu Institute of Technology

Raul Vicente Garcia
Fraunhofer Institute for Production Systems and Design Technology IPK

The PARETO-Box Problem for the Modelling of Evolutionary Multiobjective Optimization Algorithms

Mario Köppen, Raul Vicente-Garcia, Bertram Nickolay
Fraunhofer IPK, Pascalstr. 8-9, 10587 Berlin, Germany
{mario.koeppen|raul|nickolay}@ipk.fhg.de

Abstract

This paper presents the Pareto-Box problem for modelling evolutionary multi-objective search. The problem is to find the Pareto set of randomly selected points in the unit hypercube. While the Pareto set itself is only comprised of the point 0, this problem allows for a complete analysis of random search and demonstrates the fact that with increasing number of objectives, the probability of finding a dominated vector is decreasing exponentially. Since most nowadays evolutionary multi-objective optimization algorithms rely on the existence of dominated individuals, they show poor performance on this problem. However, the fuzzification of the Pareto-dominance is an example for an approach that does not need dominated individuals, thus it is able to solve the Pareto-Box problem even for a higher number of objectives.

1 Introduction

In multiobjective optimization, the optimization goal is given by more than one objective to be extreme. Formally, given a domain as subset of R^n, there are assigned m functions $f_1(x_1,\ldots,x_n),\ldots,f_m(x_1,\ldots,x_n)$. Usually, there is not a single optimum but rather the so-called PARETO set of *non-dominated* solutions.

Evolutionary Computation (EC) has been shown to be a powerful technique for multi-objective optimization [1][2][3] (EMO - Evolutionary Multi-Objective Optimization). This biologically inspired methodology offers both flexibility in goal specification and good performance in multimodal, nonlinear search spaces.

If we want to solve a highly complex multi-objective optimization problem, we might select one of the best ranked evolutionary approaches reviewed in the literature, like NSGA-II [4] or SPEA2 [5] and hopefully start reaching good results quickly. Sadly, that will be rarely the case when we face real-world problems with high number of objectives. Usually, the different algorithms are compared by measuring their performance indices in difficult test searches [1][2][3][6]. However, most of the benchmarks in the literature do not consider problems with high number of objectives. Moreover, very complex test functions should not be the only reference for the design of new approaches, as they prevent us from keeping track of the populations dynamics unambiguously (as already stated by Coello in [2]). In order to design a faster PARETO-dominance-based EC technique, we need an "easy" multi-objective test function that allows us to observe the search progress and that is yet easily scalable to higher number of objectives as well. The PARETO-Box Problem, which will be presented and studied in this paper, unifies these crucial properties. It will help us to know more about how the PARETO-front is searched for in EMO, and to measure the progress of the novel Fuzzy PARETO Dominance-Driven Genetic Algorithm (FDD-GA) approach in search problems with higher number of objectives. In the following section, we will introduce the PARETO-Box problem and its analysis for random search. These results will be used in an exemplary manner to study the dynamics of EMOs in section 3.

2 The Pareto-Box Problem

Given are m uniformly randomly selected n-dimensional points P_i in the n-dimensional unit hypercube ($1 \le i \le m$), with coordinates P_{ij} ($1 \le j \le n$). Thus, for each P_{ij} we have $0 \le P_{ij} \le 1$. The problem we state is:

PARETO-Box Problem: *What is the expectation value for the size of the* PARETO *set of these points?*

Here, we use the minimum version of PARETO dominance, so for two n-dimensional vectors $a=(a_i)$ and $b=(b_i)$ it is said that a dominates b (written as $a \prec b$) if and only if

$$\forall i:\ a_i \le b_i,\ \wedge\ \exists j:\ a_j < b_j\ \ (1 \le i,j \le n) \qquad (1)$$

For a set M of points, its PARETO set $P(M)$ is the subset for which none of its elements is dominated by any element of M. The PARETO set of the complete unit hypercube contains only one element, the point 0. The random sampling represents a random search in the unit hypercube, thus we are also going to answer the question if random search can find the PARETO set of the unit hypercube.

Obviously, the PARETO set of this problem is not hard to find, and there is also no conflict in the objectives. However, the following analysis will show that it is a hard problem for multi-objective optimization, once the dimension n of the problem is increased. Moreover, this problems allows for a precise analysis of the *progress* of algorithmic search, including the approach to the PARETO front and the entering of concave regions of the PARETO front.

In the following, $e_m(n)$ denotes the expectation value for the size of the PARETO set of m randomly selected points in the n-dimensional unit hypercube. Then, the following theorems hold:

Theorem 1. *Given are m randomly selected points in the n-dimensional hypercube. For the expectation value of the size of the* PARETO *set of these m points we have the recursive relation:*

$$\begin{aligned} e_1(n) &= 1 \\ e_m(1) &= 1 \\ e_m(n) &= e_{m-1}(n) + \frac{1}{m} e_m(n-1) \quad (n, m \geq 2) \end{aligned} \tag{2}$$

Theorem 2. *The expectation value for the size of the* PARETO *set of $m \geq 1$ randomly selected points in the n-dimensional hypercube ($n \geq 1$) is*

$$e_m(n) = \sum_{k=1}^{m} \frac{(-1)^{k+1}}{k^{n-1}} \binom{m}{k} \tag{3}$$

Due to space limitations, the proofs of these theorems can not be given here.

Theorems 1 and 2 allow for the specification of the limiting behaviour of the expectation values for increasing number of points and increasing dimensions. This is stated in the following central theorem.

Theorem 3. *For fixed dimension $n > 1$ and the number of points $m \to \infty$, the expectation value $e_m(n) \to \infty$, the ratio of the non-dominated points $e_m(n)/m \to 0$ and for fixed $m > 1$ and dimension $n \to \infty$ it holds $e_m(n) \to m$.*

Proof. We see that

$$e_m(2) = \sum_{k=1}^{m} \frac{1}{k} = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \ldots + \frac{1}{m} \tag{4}$$

which is the harmonic series and known to be divergent. Now, eq. (3) shows that for $n > 2$ always $e_m(n) \geq e_m(n-1) \geq \ldots \geq e_m(2)$, so for $m \to \infty$ $e_m(n) \to \infty$ as well. From the corresponding property of the harmonic series, $e_m(n)/m \to 0$ for $m \to \infty$ can be seen in a similar manner.

On the other hand, if $m > 1$ is fixed, all terms in eq. (3) but the one for $k = 1$ will go to 0 for $n \to \infty$, and the term for $k = 1$ itself computes to m. So, it is easy to see that $e_m(n) \to m$ for $n \to \infty$. □

We can express this result as follows: for increasing number of sample points in the hypercube, the number of non-dominated points will also increase, and never "shrink" to the PARETO set of the hypercube, which only contains the point 0. So, random search will not solve the problem to find the PARETO set of the hypercube in any dimension.

For increasing dimension, it will become more and more unlikely to find *any* dominated point in a population of random sampling points. In fact, the probability falls exponentially. The PARETO set of m points will contain nearly all m points.

We conclude this section by providing some special results:

$(m,2):\quad e_m(2) = \sum_{k=1}^{m} \frac{1}{k}$

$(2,n):\quad e_2(n) = 2 - \frac{1}{2^{n-1}}$

$(3,n):\quad e_3(n) = 3 - \frac{3}{2^{n-1}} + \frac{1}{3^{n-1}}$

$(4,n):\quad e_4(n) = 4 - \frac{6}{2^{n-1}} + \frac{4}{3^{n-1}} - \frac{1}{4^{n-1}}$

3 EMO Analysis

The remarkable point on the PARETO-Box problem is that it establishes the fact that the probability of finding dominated points in higher dimensions (i.e. increasing number of objectives) is falling exponentially with the dimension of the problem. Having a look on most prominent EMO algorithms like NSGA-II [5], SPEA2 [4] or PESA [7], it can be seen that they all need dominated points to perform their processing steps. For still yielding dominated points in the domain of higher number of objectives, these algorithms need an exponetially increasing *search effort*, be it by increasing the population size, or be it by increasing the number of generations.

Recently, the fuzzification of the PARETO dominance relation has been proposed [8], and a corresponding EMO has been presented. It will be shortly recalled here (see [9] for an alternative approach to introduce fuzzy logic in EMO). The underlying generic fuzzy ranking scheme for a set S of multivariate data (vectors) $\vec{a}_i$ with real-valued components a_{ij} and $1 \leq i \leq N$ is based on the provision of a *comparison function* $f_x(y) : R \times R \to [0,1]$ and a T-norm. Then, the following two steps are performed:

1. We compute the *comparison values* for any two vectors $\vec{a}_i = (a_{ik})$ and $\vec{a}_j = (a_{jk})$ by $c_{\vec{a}_i}(\vec{a}_j) = T(f_{a_{ik}}(a_{jk})|k = 1,\ldots,N)$ with N the number of components of each vector.

2. We compute the *ranking values* for any element $\vec{a}_i$ of S by $r_S(\vec{a}_i) = \max[c_{\vec{a}_i}(\vec{a}_j)|j \neq i]$.

Then, we consider vectors with lower numerical ranking values to be on a higher ranking position. For step 2, instead of max the min operator can be used as well, depending on the ranking to be favoured in increasing or decreasing order.

When using the comparison function bounded division and the algebraic (or product) norm as T-norm, the ranking scheme fulfills several useful properties like scale-independency in the data. The fuzzification of PARETO dominance relation can be written then as follows: It is said that vector $\vec{a}$ *dominates* vector $\vec{b}$ by degree μ_a with

$$\mu_a(\vec{a},\vec{b}) = \frac{\prod_i \min(a_i,b_i)}{\prod_i a_i} \qquad (5)$$

and that vector $\vec{a}$ *is dominated by* vector $\vec{b}$ at degree μ_p with

$$\mu_p(\vec{a},\vec{b}) = \frac{\prod_i \min(a_i,b_i)}{\prod_i b_i} \qquad (6)$$

For $\vec{a}$ PARETO-dominating $\vec{b}$, $\mu_a(\vec{a},\vec{b}) = 1$ and $\mu_p(\vec{b},\vec{a}) = 1$, but $\mu_p(\vec{a},\vec{b}) < 1$ and $\mu_a(\vec{b},\vec{a}) < 1$. Figure 1 gives a numerical example for the fuzzy PARETO dominance considered here.

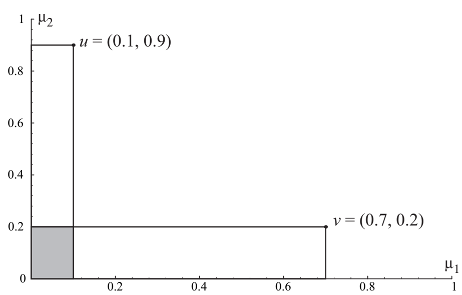

Fig. 1. Definition of Fuzzy-PARETO-Dominance. Here, u dominates v by degree $0.1 \cdot 0.2/0.1 \cdot 0.9 = 0.\bar{2}$ and is dominated by v by degree $0.1 \cdot 0.2/0.7 \cdot 0.2 \approx 0.143$.

The advantage of the FPD is that the problem of missing dominated points does not matter. This will be demonstrated by using the PARETO-Box problem. We also provide here a (Fuzzy-Dominance-Driven) FDD algorithm, a Genetic Algorithm (GA) variant that employs the FPD ordering of fitness values (represented as vectors in case of multiobjective optimization) for defining selection operators (see [10] for more details). As for nearly all newer EMOs, it also adds an *archive* for storing non-dominated individuals over the whole run of the algorithm.

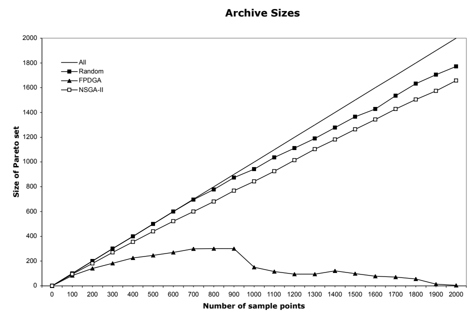

Fig. 2. Performances of NSGA-II and FDD-GA for the PARETO-Box problem.

Figure 2 compares the performance of a comparable set-up of NSGA-II and FDD-GA on the PARETO-Box problem for dimension $n = 20$. The NSGA-II implementation strictly followed [4]. For both algorithms, the population size was 10, and 200 generations were performed. Both algorithm used the same mutation probability and strength of 0.1. The selection scheme of FDD-GA was adapted due to having a known co-domain of the ranking values (aka fitness values). Tournament selection was performed using $\ln -r_i$ (with r_i the ranking values) as shared fitness values, and it was only selected among the non-dominated individuals. If all individuals got the same ranking values, it was randomly selected. The plot shows the size of the archive over the number of sample points (i.e. calls of the objective functions). Also given is the (numerically estimated) size of the PARETO-set for random sampling, and the total number of individuals (dominated and non-dominated). As established by Theorem 3, for random search the size of the PARETO set is close to the total number of points. However, also NSGA-II runs close to this curve, qualifying this search as more or less random as well. This is obvious from the fact that the probability of finding a dominated individual by applying randomized operators (mutation, crossover) is low.

For FDD-GA, we clearly see that even for dimension 20 it is able to find the single optimum of the PARETO-Box problem, and also stays strongly below the curve of random search all the time. To make this behaviour more clear, we considered the p.d.f. of the ranking values within a randomly created population (see fig. 3). This plot was obtained by 100 times creating a set of 20 random vector with 100 components from $[0,1]$. Then, among these 20 vectors the ranking values r_i were computed, and the intervall frequencies for $-\ln r_i$ were derived. Thus, we can model the handling of randomly selected points by the FPD ranking scheme (as it happens when applied to the PARETO-Box problem). The distribution has a tail at the sider of smaller ranking values, so roulette-wheel selection will acknowledge the fact that such individuals gradually perform better (with respect to PARETO-dominance). Such a behaviour can not be achieved when an EMO is relying on the presence of dominated individuals alone.

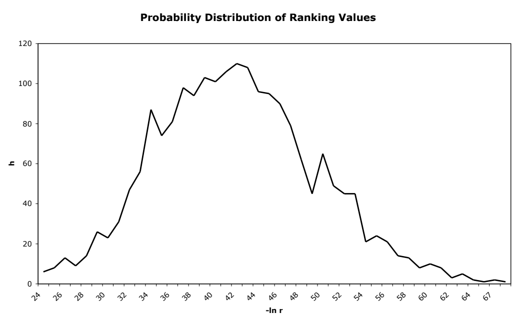

Fig. 3. Distribution of ranking values in FDD-GA algorithm.

It has to be noted (but can not be detailed here) that nevertheless NSGA-II, in this set-up, is also finding the optimum up to a problem dimension of 8. In low dimensions (2-3) the FDD-GA is also outperformed by NSGA-II.

Acknowledgment

The research presented in this paper was supported by German research project ”VisionIC,” funded by German Ministry of Research and Education (BMBF) as handled by ”Deutsche Gesellschaft für Luft- und Raumfahrt” (DLR). The authors also would like to express their thanks to Frank Hoffmann from University of Dortmund and Anna Ukovich from University of Trieste for their valuable comments and inspiring discussions about the topics of this work.

References

[1] Carlos A. Coello Coello, “A short tutorial on evolutionary multiobjective optimization,” in *First International Conference on Evolutionary Multi-Criterion Optimization*, Eckart Zitzler, Kalyanmoy Deb, Lothar Thiele, Carlos A. Coello Coello, and David Corne, Eds., vol. 1993 of *Lecture Notes in Computer Science*, pp. 21–40. Springer Verlag, 2001.

[2] Carlos A. Coello Coello, “An updated survey of GA-based multiobjective optimization techniques,” Technical Report RD-98-08, Laboratorio Nacional de Informática Avanzada (LANIA), Xalapa, Veracruz, México, December 1998.

[3] Carlos M. Fonseca and Peter J. Fleming, “An overview of evolutionary algorithms in multiobjective optimization,” *Evolutionary Computation*, vol. 3(1), no. 1, pp. 1–16, 1995.

[4] Kalayanmoy Deb, Samir Agrawal, Amrit Pratap, and T. Meyarivan, “A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization: NSGA-II,” in *Proceedings of the Parallel Problem Solving from Nature VI Conference*, Marc Schoenauer, Kalyanmoy Deb, Günter Rudolph, Xin Yao, Evelyne Lutton, J. J. Merelo, and Hans-Paul Schwefel, Eds., Paris, France, 2000, vol. 1917 of *Lecture Notes in Computer Science*, pp. 849–858, Springer.

[5] Eckart Zitzler, Marco Laumanns, and Lothar Thiele, “SPEA2: Improving the strength pareto evolutionary algorithm,” Tech. Rep. 103, Computer Engineering and Networks Laboratory (TIK), Swiss Federal Institute of Technology, Zurich, ETH Zentrum, Gloriastr 35, CH-8092 Zurich, Switzerland, September 2001.

[6] Tatsuya Okabe, Yaouchu Jin, and Bernhard Sendhoff, “A critical survey of performance indices for multi-objective optimization,” in *Proceedings of the 2003 Congress on Evolutionary Computation (CEC’2003)*, Canberra, Australia, December 2003, pp. 878–885, IEEE Press.

[7] David W. Corne, Joshua D. Knowles, and Martin J. Oates, “The pareto envelope-based selection algorithm for multiobjective optimization,” in *Proceedings of the Parallel Problem Solving from Nature VI Conference*. 2000, vol. 1917 of *Lecture Notes in Computer Science*, pp. 839–848, Springer.

[8] Mario Köppen and Raul Vicente Garcia, “A fuzzy scheme for the ranking of multivariate data and its application,” in *Proceedings of the 2004 Annual Meeting of the NAFIPS (CD-ROM)*, Banff, Alberta, Canada, 2004, NAFIPS, pp. 140–145.

[9] Marco Farina and Paolo Amato, “Fuzzy optimality and evolutionary multiobjective optimization,” in *EMO 2003, LNCS 2632*, C.M. Fonseca et al., Ed., 2003, pp. 58–72.

[10] Mario Köppen, Katrin Franke, and Bertram Nickolay, “Fuzzy-pareto-dominance driven multiobjective genetic algorithm,” in *Proceedings of the 10th IFSA World Congress*, 2003, pp. 450–453.