Dimensionality Reduction in Multiobjective Optimization: The Minimum Objective Subset Problem

Dimo Brockhoff and Eckart Zitzler {brockhoff,zitzler}@tik.ee.ethz.ch

Computer Engineering and Networks Laboratory, ETH Zurich, 8092 Zurich, Switzerland

Summary. The number of objectives in a multiobjective optimization problem strongly influences both the performance of generating methods and the decision making process in general. On the one hand, with more objectives, more incomparable solutions can arise, the number of which affects the generating method's performance. On the other hand, the more objectives are involved the more complex is the choice of an appropriate solution for a (human) decision maker. In this context, the question arises whether all objectives are actually necessary and whether some of the objectives may be omitted; this question in turn is closely linked to the fundamental issue of conflicting and non-conflicting optimization criteria. Besides a general definition of conflicts between objective sets, we here introduce the $\mathcal{NP}$-hard problem of computing a minimum subset of objectives without losing information (`MOSS`). Furthermore, we present for `MOSS` both an approximation algorithm with optimum approximation ratio and an exact algorithm which works well for small input instances. We conclude with experimental results for a random problem and the multiobjective 0/1-knapsack problem.

1 Motivation

It is indisputable, that a high number of criteria in a multiobjective optimization problem [5, 10] can cause additional difficulties compared to a low-dimensional problem. The number of incomparable solutions can raise [7], the (human) decision making process becomes harder and generating methods may require substantially more computational resources[1] when more criteria are involved in the problem formulation. Consequently, the question arises whether it is possible to omit some of the objectives without changing the characteristics of the underlying problem. Furthermore, one may ask under which conditions such an objective reduction is feasible and how a minimum set of objectives can be computed.

These questions have gained only little attention in the literature so far. Transferred to the multiobjective optimization setting, closely related research topics as principal component analysis [9] or dimension theory [14] aim at determining a (minimum) set of *arbitrary* objective functions that preserves (most of) the problem

[1] For example, when based on the $\mathcal{S}$-metric [15].

characteristics; however, here we are interested in determining a minimum subset of *original* objectives that maintains the order on the search space. Known attempts dealing with the latter do not preserve the dominance structure [4] or are not suitable for black-box scenarios [1]. Furthermore, there are a few studies that investigate the relationships between objectives in terms of conflicting and nonconflicting optimization criteria. The definitions of conflicting objectives are based on trade-offs within the Pareto optimal set [1, 3, 8] or on the number of incomparable solutions [11, 13]. Conflicts are either defined between objective pairs [1, 8, 11, 13] or as a property of the entire objective set [3]. However, these definitions are not sufficient to indicate whether objectives can be omitted or not; examples can be constructed, cf. [2], where all objectives are conflicting according to [3, 11, 13], but one among the given objectives can be removed while preserving the search space order.

Assuming that the decision making process is postponed after the search process, this paper addresses two open issues for a given set of (trade-off) solutions: (i) deriving general conditions under which certain objectives may be omitted and (ii) computing a minimum subset of objectives needed to preserve the problem structure between the given solutions. In particular, we

- propose a generalized notion of objective conflicts which comprises the definitions of Deb [3], Tan et al. [13], and Purshouse and Fleming [11],
- specify on this basis a necessary and sufficient condition under which objectives can be omitted,
- introduce the $\mathcal{NP}$-hard problem of minimum objective subsets (MOSS),
- provide both an approximation algorithm and an exact algorithm for MOSS,
- validate our approach in additional experiments by comparing the algorithms and investigating the influence of the number of objectives and the search space size.

2 A Notion of Objective Conflicts

Without loss of generality, in this paper we consider a minimization problem with k objective functions $f_i : X \to \mathbb{R}$, $1 \leq i \leq k$, where the vector function $f := (f_1, \dots, f_k)$ maps each solution $\mathbf{x} \in X$ to an objective vector $f(\mathbf{x}) \in \mathbb{R}^k$. Furthermore, we assume that the underlying dominance structure is given by the weak Pareto dominance relation[2] which is defined as follows: $\preceq_{\mathcal{F}'} := \{(\mathbf{x}, \mathbf{y}) \mid \mathbf{x}, \mathbf{y} \in X \wedge \forall f_i \in \mathcal{F}' : f_i(\mathbf{x}) \leq f_i(\mathbf{y})\}$, where $\mathcal{F}'$ is a set of objectives with $\mathcal{F}' \subseteq \mathcal{F} := \{f_1, \dots, f_k\}$. For better readability, we will sometimes only consider the objective functions' indices, e.g., $\mathcal{F}' = \{1, 2, 3\}$ instead of $\mathcal{F}' = \{f_1, f_2, f_3\}$. The Pareto (optimal) set is given as $\{\mathbf{x} \in X \mid \nexists \mathbf{y} \in X \setminus \{\mathbf{x}\} : \mathbf{y} \preceq_{\mathcal{F}} \mathbf{x} \wedge \mathbf{x} \not\preceq_{\mathcal{F}} \mathbf{y}\}$.

In the following, we introduce a general concept of conflicts between sets of objectives and formalize the notion of redundant objective sets. On this basis, we

[2] Any other preorder rel like the ε-dominance relation can be used as well if $rel_{\mathcal{F}} = \bigcap_{1 \leq i \leq k} rel_i$ holds for any set $\mathcal{F} = \{f_1, \dots, f_k\}$ of k objective functions. The proof of the equivalence is simple for $\preceq_{\mathcal{F}}$ and can be found in [2].

then propose two algorithms to exactly resp. approximately compute a minimum set of objectives, which induces the same preorder on X as the whole set of objectives.

Definition 1. *Let $\mathcal{F}_1, \mathcal{F}_2 \subseteq \mathcal{F}$ be two sets of objectives. We call*

- $\mathcal{F}_1$ nonconflicting *with $\mathcal{F}_2$ iff* $\preceq_{\mathcal{F}_1} \subseteq \preceq_{\mathcal{F}_2} \wedge \preceq_{\mathcal{F}_2} \subseteq \preceq_{\mathcal{F}_1}$
- $\mathcal{F}_1$ weakly conflicting *with $\mathcal{F}_2$ iff* $(\preceq_{\mathcal{F}_1} \subseteq \preceq_{\mathcal{F}_2} \wedge \preceq_{\mathcal{F}_2} \not\subseteq \preceq_{\mathcal{F}_1})$ *or* $(\preceq_{\mathcal{F}_2} \subseteq \preceq_{\mathcal{F}_1} \wedge \preceq_{\mathcal{F}_1} \not\subseteq \preceq_{\mathcal{F}_2})$
- $\mathcal{F}_1$ strongly conflicting *with $\mathcal{F}_2$ iff* $\preceq_{\mathcal{F}_1} \not\subseteq \preceq_{\mathcal{F}_2} \wedge \preceq_{\mathcal{F}_2} \not\subseteq \preceq_{\mathcal{F}_1}$

Two sets of objectives $\mathcal{F}_1, \mathcal{F}_2$ are called nonconflicting if and only if the corresponding relations $\preceq_{\mathcal{F}_1}$ and $\preceq_{\mathcal{F}_2}$ are identical but not necessarily $\mathcal{F}_1 = \mathcal{F}_2$. If $\mathcal{F}_1 \subset \mathcal{F}_2$ and $\mathcal{F}_1$ is nonconflicting with $\mathcal{F}_2$ we can simply omit all objectives in $\mathcal{F}_2 \setminus \mathcal{F}_1$ without influencing the preorder on X. This formulation of conflicting objectives can be regarded as a generalization of the conflict definition in [11]; moreover, it also comprises the notions in [3] and [13], see [2] for details.

A *minimal objective set w. r. t.* $\mathcal{F}$ is a subset $\mathcal{F}'$ of the original objectives $\mathcal{F}$ that cannot be further reduced without changing the associated preorder, i.e., $\preceq_{\mathcal{F}'} = \preceq_{\mathcal{F}}$ but $\nexists \mathcal{F}'' \subset \mathcal{F}' : \preceq_{\mathcal{F}''} = \preceq_{\mathcal{F}'}$. A *minimum objective set w. r. t.* $\mathcal{F}$ is the smallest possible subset of the original objectives $\mathcal{F}$ that preserves the original order on the search space. By definition, every minimum objective set is minimal, but not all minimal sets are at the same time minimum.

Definition 2. *A set $\mathcal{F}$ of objectives is called* redundant *if and only if there exists $\mathcal{F}' \subset \mathcal{F}$ that is minimal w. r. t. $\mathcal{F}$.*

This definition of redundancy represents a necessary and sufficient condition for the omission of objectives.

3 The Minimum Objective Subset Problem

Given a multiobjective optimization problem with the set $\mathcal{F}$ of objectives, the question arises whether objectives can be omitted without changing the order on the search space. If there is an objective subset $\mathcal{F}' \subseteq \mathcal{F}$ such that $\mathbf{x} \preceq_{\mathcal{F}'} \mathbf{y}$ holds for all solutions $\mathbf{x}, \mathbf{y} \in X$ if and only if $\mathbf{x} \preceq_{\mathcal{F}} \mathbf{y}$, we can omit all objectives in $\mathcal{F} \setminus \mathcal{F}'$ while preserving the preorder on X. Concerning the last section, we are interested in identifying a minimum objective subset with respect to $\mathcal{F}$, yielding a slighter representation of the same multiobjective optimization problem. Formally, this problem, the $\mathcal{NP}$-hardness of which can easily be shown [2], is stated as follows.

Definition 3. *Given a multiobjective optimization problem with objective set $\mathcal{F} = \{f_1, \ldots, f_k\}$, the problem* `MINIMUM OBJECTIVE SUBSET` *(*`MOSS`*) is defined as follows.*

Instance: *A set $A \subseteq X$ of solutions, the generalized weak Pareto dominance relation $\preceq_{\mathcal{F}}$ and for all objective functions $f_i \in \mathcal{F}$ the single relations $\preceq_i$ where $\bigcap_{1 \leq i \leq k} \preceq_i = \preceq_{\mathcal{F}}$.*

Task: *Compute an objective subset $\mathcal{F}' \subseteq \mathcal{F}$ of minimum size with $\preceq_{\mathcal{F}'} = \preceq_{\mathcal{F}}$.*

Algorithm 1 A greedy algorithm for MOSS

$E := \preceq_{\mathcal{F}}^{C}$ where $\preceq_{\mathcal{F}}^{C} := (A \times A) \setminus \preceq_{\mathcal{F}}$
$I := \emptyset$
while $E \neq \emptyset$ **do**
  choose an $i \in (\{1,\dots,k\} \setminus I)$ such that $|\preceq_i^C \cap E|$ is maximal
  $E := E \setminus \preceq_i^C$
  $I := I \cup \{i\}$
end while

Algorithm 2 An exact algorithm for MOSS

$S := \emptyset$
for each pair $\mathbf{x}, \mathbf{y} \in A$ of solutions **do**
  $S_x := \{\, \{i\} \,|\, i \in \{1,\dots,k\} \wedge \mathbf{x} \preceq_i \mathbf{y} \wedge \mathbf{y} \not\preceq_i \mathbf{x}\}$
  $S_y := \{\, \{i\} \,|\, i \in \{1,\dots,k\} \wedge \mathbf{y} \preceq_i \mathbf{x} \wedge \mathbf{x} \not\preceq_i \mathbf{y}\}$
  $S_{xy} := S_x \sqcup S_y$ where
    $S_1 \sqcup S_2 := \{s_1 \cup s_2 \,|\, s_1 \in S_1 \wedge s_2 \in S_2 \wedge (\nexists p_1 \in S_1, p_2 \in S_2 : p_1 \cup p_2 \subset s_1 \cup s_2)\}$
  if $S_{xy} = \emptyset$ **then** $S_{xy} := \{1,\dots,k\}$
  $S := S \sqcup S_{xy}$
end for
Output a smallest set $s_{\min}$ in S

Note, that the dominance relations in Def. 3 have to be known only on the set $A \subseteq X$ of solutions, which, in practice, will mostly be rather a Pareto set approximation than the entire search space.

With Algorithm 1, we present an approximation algorithm with polynomial running time for the MOSS problem, the approximation ratio of which is optimal. Known results for the set cover problem [6, 12] can be used to prove the algorithm's $\Theta(\log |A|)$ approximation ratio [2][3]. As an exact algorithm for MOSS, we present Algorithm 2, the running time of which is polynomial in the size of A but exponential in the number of objectives[4]. That no correct predication is possible whether a set of objectives is redundant, by observing only relations between objective subsets of constant size, can be likewise derived from the $\mathcal{NP}$-hardness of the MOSS problem. Thus, Algorithm 2 is forced to examine the type of conflict between all possible objective subsets. Note that Algorithm 2 computes all minimal objective subsets for a given MOSS instance. For further details on both algorithms, we refer to [2].

4 Experiments

The following experiments serve two goals: (i) to investigate the size of a minimum objective subset depending on the size of the search space and the number of original objective functions, and (ii) to compare the two proposed algorithms with respect to the size of the generated objective subsets and their running times. These issues have

[3] For the proof of correctness and the $O(k \cdot |A|^2)$ runtime, we also refer to [2].
[4] The runtime bounds of $O(|A|^2 \cdot k \cdot 2^k)$, resp. $\Omega(|A|^2 \cdot 2^{k/3})$, are proved in [2].

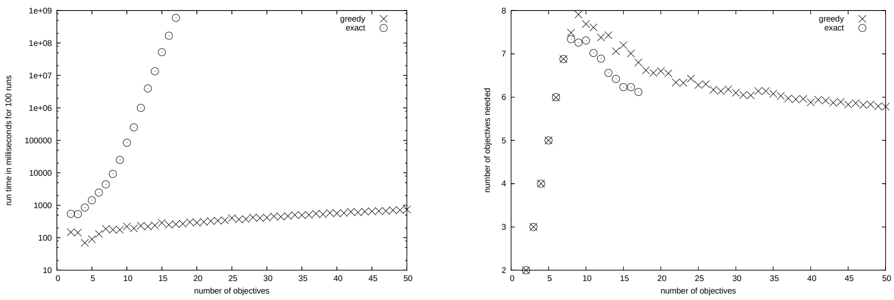

Fig. 1. Comparison between the greedy and the exact algorithm for the random problem with 32 solutions on a linux computer (SunFireV60x with 3060 MHz) over 100 runs: summed running times (*left*); averaged sizes of the computed minimum / minimal sets (*right*)

been considered both on the whole search space X for a random problem and on Pareto set approximations $A \subseteq X$ of the multiobjective 0/1-knapsack problem. For details on the experiment settings, we refer to [2] due to space limitations.

In a first experiment, we generated the objective values for a set of solutions X at random, where for each combination of search space size $|X|$ and number of objectives k, 100 independent samples were considered. Figure 2 shows the sizes of the minimum objective subsets, computed with the exact algorithm, Fig. 1 the results for both algorithms on a fixed problem size. For a second experiment, both the exact and the approximation algorithm were applied to Pareto set approximations A for the knapsack problem, generated by an evolutionary algorithm. Figure 3 shows the results together with the algorithms' runtimes.

Two main observations can be made. First, the experiments, based both on the entire decision space and on Pareto set approximations, show that the omission of objectives without information loss is possible. Furthermore, the minimum number of objectives decreases the more objectives are involved in the problem formulation, and the larger the search space the more objectives are in a minimum objective set. Second, in comparison to the exact algorithm, the greedy algorithm shows nearly the same output quality regarding the computed objective set size but is much faster. Due to the sizes of the computed subsets which are—in all of our experiments—less than one objective away from the optimum, the greedy algorithm seems to be applicable for more complex problems, particularly by virtue of its small running time.

5 Conclusions

From a practical point of view, the present study provides a first step towards dimensionality reduction of the objective space in multiple criteria optimization scenarios. We have introduced the minimum objective subset problem (`MOSS`) and proposed both an exact and a greedy algorithm for `MOSS`. The presented algorithms can be particularly useful to analyze Pareto set approximations generated by search procedures. How to integrate dimensionality reduction methods into search heuristics and the question whether this makes sense in general, has to be disputed in future work.

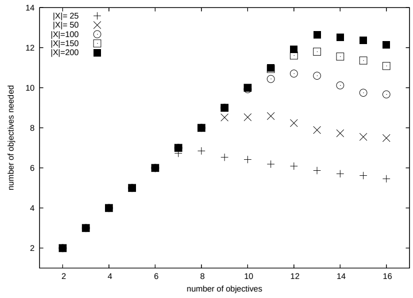

Fig. 2. Random model: The size of a minimum subset, computed with the exact algorithm, versus the number of objectives in the problem formulation

	objective set size		runtime [ms]	
k	greedy	exact	greedy	exact
5	4	4	47	196
10	5	5	46	2,271
15	8	8	67	87,113
20	13	13	88	90,524
25	16	16	78	$\approx 2.5 \cdot 10^6$
30	14	13	87	$\approx 15 \cdot 10^6$

Fig. 3. Results for both algorithms on a Pareto set approximation of the knapsack problem with different number k of objectives

References

1. P. J. Agrell. On redundancy in multi criteria decision making. *European Journal of Operational Research*, 98(3):571–586, 1997.
2. D. Brockhoff and E. Zitzler. On Objective Conflicts and Objective Reduction in Multiple Criteria Optimization. TIK Report 243, February 2006.
3. K. Deb. *Multi-objective optimization using evolutionary algorithms*. Wiley, Chichester, UK, 2001.
4. K. Deb and D. K. Saxena. On finding pareto-optimal solutions through dimensionality reduction for certain large-dimensional multi-objective optimization problems. Kangal report no. 2005011, December 2005.
5. M. Ehrgott. *Multicriteria Optimization*. Springer Berlin Heidelberg, 2005.
6. U. Feige. A threshold of ln n for approximating set cover. *J. ACM*, 45(4):634–652, 1998.
7. C. M. Fonseca and P. J. Fleming. An overview of evolutionary algorithms in multiobjective optimization. *Evolutionary Computation*, 3(1):1–16, 1995.
8. T. Gal and H. Leberling. Redundant objective functions in linear vector maximum problems and their determination. *European Journal of Operational Research*, 1(3):176–184, 1977.
9. I. T. Jolliffe. *Principal component analysis*. Springer, 2002.
10. K. M. Miettinen. *Nonlinear Multiobjective Optimization*. Kluwer Academic Publishers, 1999.
11. R. C. Purshouse and P. J. Fleming. Conflict, harmony, and independence: Relationships in evolutionary multi-criterion optimisation. In *EMO 2003 Proceedings*, pages 16–30. Springer, Berlin, 2003.
12. P. Slavík. A tight analysis of the greedy algorithm for set cover. In *STOC '96: Proceedings of the twenty-eighth annual ACM symposium on Theory of computing*, pages 435–441, New York, NY, USA, 1996. ACM Press.
13. K. C. Tan, E. F. Khor, and T. H. Lee. *Multiobjective Evolutionary Algorithms and Applications*. Springer, London, 2005.
14. W. T. Trotter. *Combinatorics and Partially Ordered Sets: Dimension Theory*. The Johns Hopkins University Press, Baltimore and London, 1992.
15. L. While. A new analysis of the lebmeasure algorithm for calculating hypervolume. In *EMO 2005 Proceedings*, pages 326–340. Springer, 2005.