TUCS

Outi Wilppu | Marko M. Mäkelä | Yury Nikulin

Two-Slope Parameterized Achievement Scalarizing Functions for Multiobjective Optimization

Turku Centre for Computer Science

TUCS Technical Report
No 1114, June 2014

Two-Slope Parameterized Achievement Scalarizing Functions for Multiobjective Optimization

Outi Wilppu
University of Turku, Department of Mathematics and Statistics
FI-20014 Turku, Finland
outi.wilppu@utu.fi
Marko M. Mäkelä
University of Turku, Department of Mathematics and Statistics
FI-20014 Turku, Finland
makela@utu.fi
Yury Nikulin
University of Turku, Department of Mathematics and Statistics
FI-20014 Turku, Finland
yurnik@utu.fi

TUCS Technical Report
No 1114, June 2014

Abstract

Most of the methods for multiobjective optimization utilize some scalarization technique where several goals of the original multiobjective problem are converted to one single-objective problem. One common scalarization technique is a use of achievement scalarizing functions. In this paper, we introduce a new family of two-slope parameterized achievement scalarizing functions for multiobjective optimization. With these two-slope parameterized ASFs we can guarantee the (weak) Pareto optimality of the solutions produced and every (weakly) Pareto optimal solution can be obtained. Parameterization of this kind gives a systematic way to produce different solutions from the same preference information. With two weighting vectors depending on the achievability of the reference point there is no need for any assumptions about the reference point. In addition to theory, we give the graphical illustrations of two-slope parameterized ASFs and analyze the quality of the solutions produced in convex and nonconvex testproblems.

Keywords: Multiobjective optimization, achievement scalarizing functions, Pareto optimal solutions

TUCS Laboratory
Turku Optimization Group (TOpGroup)

1 Introduction

In many applications, the aim is to optimize several objectives and to find a solution which is as good as possible for every objective at the same time. Usually, these objectives are conflicting and due to that it is not possible to find a solution being optimal for every objective simultaneously. That is why compromises between these conflicting objectives are needed. The compromise is optimal if any objective cannot be improved without impairing at least one of the other objectives. The problem of this kind is called a multiobjective optimization problem and its optimal solution is called a Pareto optimal solution.

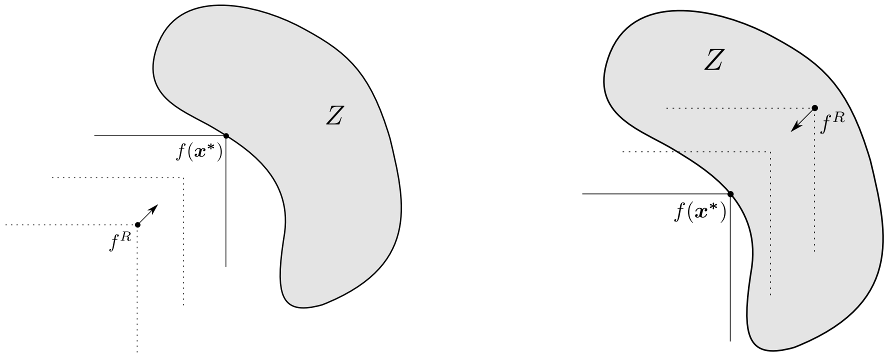
Usually there are several mathematically equally good Pareto optimal solutions and someone needs to choose the best solution for a particular problem. This person is called a decision maker who has an insight into the problem. It is also possible to obtain some additional information of the problem from the decision maker.

As it was said, the problem setting of multiobjective optimization problem differs a lot from the single-objective optimization and by solving only one objective of the multiobjective optimization problem with a single-objective method can lead to an arbitrary bad solution with respect to other objectives for the original multiobjective problem. Thus different methods in order to solve multiobjective problems are needed. Several methods are described in [1, 7, 13] and references therein. Most of the methods for multiobjective optimization utilize scalarization. In scalarization at first several goals of the original multiobjective problem are converted to one single-objective problem and then applied some suitable single-objective method. Several scalarization techniques are introduced and compared in [8].

One of the most common scalarization technique is a use of achievement scalarizing functions [7, 14, 15]. In this approach, a reference point is, for instance, asked from the decision maker and after that an achievement scalarizing function is optimized in order to find a solution being the closest to the reference point.

Chebyshev type achievement scalarizing function [14] is one of the most popular achievement scalarizing functions. If the reference point is unachievable, the distance from the reference point to the feasible region is minimized. In graphical illustration in Figure 1a right-angled contours are increasing from the unachievable reference point towards the feasible region. The optimal solution is the first point from the feasible region touching the contour. On the other hand, if the reference point is achievable, the maximum value of the negative difference between the reference point and the nondominated set (i.e. the set of Pareto optimal solutions in objective space) is minimized. In graphical illustration in Figure 1b the optimal solution for scalarizied problem is the nondominated point touching the contour last.

The wide usage of Chebyshev type achievement scalarizing function is due to



(a) Unachievable reference point (b) Achievable reference point

Figure 1: Graphical illustration of Chebyshev type achievement scalarizing function

its good mathematical properties. With this L_∞ metric any weakly Pareto optimal solution can be obtained. Also other type of metrics can be used, for example linear L_1 metric but unlike Chebyshev metric with L_1 metric not every weakly Pareto optimal solution are necessarily obtained in nonconvex case since there might exist nonsupported solutions. To overcome this drawback in [12] there is presented L_1 based metric ensuring that every weakly Pareto optimal solution can be obtained.

In this paper, we propose a new family of two-slope parameterized achievement scalarizing functions (TSPASF). These functions base on the parameterized achievement scalarizing functions (PASF) introduced in [11]. By using parameterization metrics varying from Chebyshev metric to linear metric are possible to utilize. We generalize the PASF by utilizing the idea of two different weighting vectors depending on the achievability of the reference point described in [3]. The advantage of this new TSPASF is that any Pareto optimal solution can be found by moving the reference point or changing the weighting vectors. Another advantage compared with PASF is that we need neither to assume anything about the reference point nor to test whether the reference point is achievable or not. This occurs since the formulation of the problem guarantees that the right weighting vector is used in every case.

This paper is organized as follows: In Section 2 we recall some basic results of multiobjective optimization and describe the ideas of achievement scalarizing functions and parameterized achievement scalarizing functions. Section 3 is dedicated to a new two-slope parameterized achievement function and a special case of three objectives is analyzed in Section 4. In Section 5, we give some numerical examples and in Section 6 some final remarks.

2 Preliminaries

We consider a multiobjective optimization problem where all the objectives are minimized simultaneously. This problem is of the form

$$\begin{aligned} \min \quad & f(\boldsymbol{x}) = (f_1(\boldsymbol{x}), \ldots, f_m(\boldsymbol{x})) \\ \mathrm{s.t.} \quad & \boldsymbol{x} \in X, \end{aligned} \tag{1}$$

where the partial objective functions are defined $f_i : X \to \mathbb{R}$, $i \in \mathbb{N}_m = \{1, \ldots, m\}$ and they are assumed to be lower semicontinuous. A set $X \subset \mathbb{R}^n$ is a non-empty compact set of feasible solutions. The image of this set X is called a feasible objective region $Z = f(X)$. The objective functions are also assumed to be conflicting and thus it is impossible to have a solution being minimal for every objective function.

We recall some basic results from multiobjective optimization. For more details we refer to [1, 7]. In the following we use notation $\boldsymbol{x} < \boldsymbol{y}$ if $x_i < y_i$ for all $i \in \mathbb{N}_n$ and notation $\boldsymbol{x} \leq \boldsymbol{y}$ if $x_i \leq y_i$ for all $i \in \mathbb{N}_n$.

An optimal solution of problem (1) is called Pareto optimal if any objective cannot be improved without deteriorating some other objective at the same time. Formally we can define that a solution $\boldsymbol{x}^* \in \mathbb{R}^n$ of the problem (1) is *Pareto optimal* if there does not exist another point $\boldsymbol{x} \in \mathbb{R}^n$ such that $f_i(\boldsymbol{x}) \leq f_i(\boldsymbol{x}^*)$ for all $i \in \mathbb{N}_m$ and $f_j(\boldsymbol{x}) < f_j(\boldsymbol{x}^*)$ for at least one index $j \in \mathbb{N}_m$. Under the assumptions of problem (1), Pareto optimal solutions exist [13]. Usually, there exist several mathematically equally good Pareto optimal solutions and a set of these Pareto optimal solutions is called the *Pareto set*.

We can also define a generalized concept where a solution $\boldsymbol{x}^* \in \mathbb{R}^n$ is called *weakly Pareto optimal* if there does not exist another point $\boldsymbol{x} \in \mathbb{R}^n$ such that $f_i(\boldsymbol{x}) < f_i(\boldsymbol{x}^*)$ for all $i \in \mathbb{N}_m$. In this case there exists no other solution such that all objectives have a better value. Note that the set of Pareto optimal solutions is a subset of the set of weakly Pareto optimal solutions.

To get some information about Pareto optimal solutions an ideal and a nadir vector, f^I and f^N, can be calculated giving a lower and an upper bound for the range of Pareto optimal solutions, respectively. The components of the *ideal vector* are obtained by minimizing every objective separately. Thus the i:th component of the ideal vector can be defined by solving the problem $\min_{\boldsymbol{x} \in X} f_i(\boldsymbol{x})$. The ideal vector tells how good solutions can be found but normally ideal vector is not a feasible solution. If the ideal vector is a feasible solution, then it would clearly be also an optimal solution of problem (1).

The *nadir vector* relates the upper bound for Pareto optimal solutions representing the worst solution. The components of the nadir vector can be calculated by maximizing objectives over the set of Pareto optimal solutions. Due to this optimization over the Pareto set it is usually difficult to obtain the nadir vector but it can be approximated for example with the pay-off matrix [1, 7].

A *utopian vector* gives the strictly better solution than any of the Pareto optimal solution and even better than the ideal vector. The components of the utopian vector are of the form $f_i^U = f_i^I - \varepsilon_i$ where $\varepsilon_i > 0$ is a sufficient small constant.

A point which consists of desirable values for objective functions is called a *reference point* $f^R = (f_i^R, \dots, f_m^R)$. These desirable values have been provided by the decision maker who tells what (s)he wishes to achieve. The reference point is said to be *achievable* if $f^R \in Z + \mathbb{R}_+^m$ where $\mathbb{R}_+^m = \{\boldsymbol{y} \in \mathbb{R}^m \mid y_i \geq 0 \text{ for } i \in \mathbb{N}_m\}$. Otherwise the reference point is said to be *unachievable*.

In this paper, we are focusing on achievement scalarizing functions (ASF) [14, 15] in order to scalarize multiobjective problem (1). This scalarized problem is of the form

$$\min_{\boldsymbol{x} \in X} s_R(f(\boldsymbol{x}), \boldsymbol{\lambda}). \tag{2}$$

One example of achievement scalarizing functions is Chebyshev type

$$s_R(f(\boldsymbol{x}), \boldsymbol{\lambda}) = \max_{i \in \mathbb{N}_m} \left\{ \lambda_i (f_i(\boldsymbol{x}) - f_i^R) \right\}, \tag{3}$$

where the vector f^R is a reference point and the value $\lambda_i > 0$ is a weighting coefficient for the objective function f_i specifying the direction of the projection from the reference point to the Pareto frontier.

If the reference point is unachievable, then the ASF is minimizing the distance from the reference point to the feasible set. On the other hand, if the reference point is achievable, we are minimizing the maximum value of the negative difference between the reference point and the nondominated set. By moving the reference point or manipulating $\boldsymbol{\lambda}$, any (weakly) Pareto optimal solution can be obtained [7].

In order to guarantee that problem (2) generates Pareto optimal solutions the following properties of the ASFs can be described.

Definition 2.1. [17] An achievement scalarizing function $s_R : \mathbb{R}^m \times \mathbb{R}_+^m \to \mathbb{R}$ is said to be

1. *increasing* if for any $\boldsymbol{y}_1$, $\boldsymbol{y}_2 \in \mathbb{R}^m$, $\boldsymbol{y}_1 \leq \boldsymbol{y}_2$, then $s_R(\boldsymbol{y}_1, \boldsymbol{\lambda}) \leq s_R(\boldsymbol{y}_2, \boldsymbol{\lambda})$.
2. *strictly increasing* if for any $\boldsymbol{y}_1$, $\boldsymbol{y}_2 \in \mathbb{R}^m$, $\boldsymbol{y}_1 < \boldsymbol{y}_2$, then $s_R(\boldsymbol{y}_1, \boldsymbol{\lambda}) < s_R(\boldsymbol{y}_2, \boldsymbol{\lambda})$.
3. *strongly increasing* if for any $\boldsymbol{y}_1$, $\boldsymbol{y}_2 \in \mathbb{R}^m$, $\boldsymbol{y}_1 \leq \boldsymbol{y}_2$ and $\boldsymbol{y}_1 \neq \boldsymbol{y}_2$, then $s_R(\boldsymbol{y}_1, \boldsymbol{\lambda}) < s_R(\boldsymbol{y}_2, \boldsymbol{\lambda})$.

Note that any strongly increasing ASF is also strictly increasing and any strictly increasing ASF is also increasing. For example, a function of Chebyshev type (3) is strictly increasing.

The following two theorems specifies necessary and sufficient conditions to (weak) Pareto optimality:

Theorem 2.2. [16, 17] *The following two statements are true:*

1. *Let s_R be strongly increasing. If $\boldsymbol{x}^* \in X$ is an optimal solution of problem* (2)*, then $\boldsymbol{x}^*$ is (weakly) Pareto optimal for problem* (1)*.*

2. *If s_R is increasing and the solution $\boldsymbol{x}^* \in X$ of problem* (2) *is unique, then $\boldsymbol{x}^*$ is Pareto optimal for problem* (1)*.*

Theorem 2.3. [7] *If s_R is strictly increasing and $\boldsymbol{x}^* \in X$ is weakly Pareto optimal solution for problem* (1)*, then it is a solution of problem* (2) *with $f^R = f(\boldsymbol{x}^*)$ and the optimal value of s_R is zero.*

A starting point for this paper is a parameterized achievement scalarizing function developed in [11] which extends the ideas of an additive achievement scalarizing function introduced in [12]. Let I^q be a subset of $\mathbb{N}_m$ of cardinality q. A parameterized achievement scalarizing function (PASF) [11] is of the form

$$\tilde{s}_R^q(f(\boldsymbol{x}), \boldsymbol{\lambda}) = \max_{I^q \subseteq \mathbb{N}_m : |I^q| = q} \left\{ \sum_{i \in I^q} \max[\lambda_i (f_i(\boldsymbol{x}) - f_i^R), 0] \right\}$$

where $q \in \mathbb{N}_m$ and $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_m)$, $\lambda_i > 0$, $i \in \mathbb{N}_m$. The problem to be solved is then

$$\min_{\boldsymbol{x} \in X} \tilde{s}_R^q(f(\boldsymbol{x}), \boldsymbol{\lambda}). \tag{4}$$

Due to the formation of PASF the value of $\tilde{s}_R^q$ is always nonnegative. With different values of the parameter q different metrics varying between L_1 to L_∞ are obtained. Extreme cases are L_1 metric with $q = m$, where m is the number of objectives, and L_∞ metric with $q = 1$.

The following two properties were proven for $\tilde{s}_R^q$ in [11].

Theorem 2.4. [11] *Given problem* (4)*, let f^R be a reference point such that there exists no feasible solution whose image strictly dominates f^R and $\lambda_i > 0$ for all $i \in \mathbb{N}_m$. Then any optimal solution of problem* (4) *is a weakly Pareto optimal solution for problem* (1)*.*

Theorem 2.5. [11] *Given problem* (4)*, let f^R be any reference point and $\lambda_i > 0$ for all $i \in \mathbb{N}_m$. Then, among the optimal solutions of problem* (4) *there exists at least one Pareto optimal solution for problem* (1)*.*

Theorem 2.5 implies that if $\boldsymbol{x}^*$ is a unique solution of problem (4), then it is a Pareto optimal solution for problem (1).

With the PASF, several Pareto optimal solutions can be found by moving the reference point or by manipulating the weighting coefficients and the reference point stays fixed. A limitation of the PASF is that the reference point should not be strictly dominated by some feasible point. With the two-slope parameterized ASF described in the next section this limitation can be forgot.

3 Two-slope parameterized achievement scalarizing functions

Next we introduce an extended parameterized ASF which takes achievability of the reference point into account as in [3]. A two-slope parameterized ASF (TSPASF) is defined as follows:

$$\hat{s}^q_R(f(\boldsymbol{x}), \boldsymbol{\lambda}^U, \boldsymbol{\lambda}^A) = \max_{I^q \in \mathbb{N}_m : |I^q| = q} \Bigg\{ \sum_{i \in I^q} \Big[\max\left\{\lambda_i^U (f_i(\boldsymbol{x}) - f_i^R), 0\right\} + \min\left\{\lambda_i^A (f_i(\boldsymbol{x}) - f_i^R), 0\right\} \Big] \Bigg\}, \tag{5}$$

where $q \in \mathbb{N}_m$, $\boldsymbol{\lambda}^U = (\lambda_1^U, \ldots, \lambda_m^U)$ and $\boldsymbol{\lambda}^A = (\lambda_1^A, \ldots, \lambda_m^A)$, $\lambda_i^U, \lambda_i^A > 0$, $i \in \mathbb{N}_m$. Either of these two different weighting vectors $\boldsymbol{\lambda}^U$, $\boldsymbol{\lambda}^A$ are used depending on whether the reference point is achievable or unachievable, respectively. The problem to be solved is then

$$\min_{\boldsymbol{x} \in X} \hat{s}^q_R(f(\boldsymbol{x}), \boldsymbol{\lambda}^U, \boldsymbol{\lambda}^A). \tag{6}$$

Notice that if $q = 1$, then $\hat{s}^q_R$ has the same formation than two-slope ASF proposed in [3].

For the TSPASF is possible to prove similar result as Theorem 2.4. Note that the assumption of nonexisting feasible solution which image strictly dominates f^R is not needed.

Theorem 3.1. *Given problem* (6) *and let* $\lambda_i^U, \lambda_i^A > 0$ *for all* $i \in \mathbb{N}_m$. *Then any optimal solution of problem* (6) *is a weakly Pareto optimal solution for problem* (1).

Proof. Let $\boldsymbol{x}^*$ be an optimal solution of problem (6). Assume that $\boldsymbol{x}^*$ is not weakly Pareto optimal. Then there exists a feasible solution $\boldsymbol{x}' \in X$ such that $f_i(\boldsymbol{x}') < f_i(\boldsymbol{x}^*)$ for all $i \in \mathbb{N}_m$.

For any $\boldsymbol{x} \in X$, denote $I_{\boldsymbol{x}} = \left\{i \in \mathbb{N}_m \mid f_i^R \le f_i(\boldsymbol{x})\right\}$ and $J_{\boldsymbol{x}} = \left\{i \in \mathbb{N}_m \mid f_i^R > f_i(\boldsymbol{x})\right\}$. Since $I_{\boldsymbol{x}'} \subseteq I_{\boldsymbol{x}^*}$ and $J_{\boldsymbol{x}'} \supseteq J_{\boldsymbol{x}^*}$ we obtain

$$\begin{aligned}
&\hat{s}^q_R(f(\boldsymbol{x}'), \boldsymbol{\lambda}^U, \boldsymbol{\lambda}^A)\\
&= \max_{I^q \in \mathbb{N}_m : |I^q| = q} \left\{ \sum_{i \in I^q} \left[\max\left\{\lambda_i^U (f_i(\boldsymbol{x}') - f_i^R), 0\right\} + \min\left\{\lambda_i^A (f_i(\boldsymbol{x}') - f_i^R), 0\right\}\right] \right\}\\
&= \max_{I^q \in \mathbb{N}_m : |I^q| = q} \left\{ \sum_{i \in I^q \bigcap I_{\boldsymbol{x}'}} \lambda_i^U (f_i(\boldsymbol{x}') - f_i^R) + \sum_{i \in I^q \bigcap J_{\boldsymbol{x}'}} \lambda_i^A (f_i(\boldsymbol{x}') - f_i^R) \right\}\\
&< \max_{I^q \in \mathbb{N}_m : |I^q| = q} \left\{ \sum_{i \in I^q \bigcap I_{\boldsymbol{x}'}} \lambda_i^U (f_i(\boldsymbol{x}^*) - f_i^R) + \sum_{i \in I^q \bigcap J_{\boldsymbol{x}'}} \lambda_i^A (f_i(\boldsymbol{x}^*) - f_i^R) \right\}
\end{aligned}$$

$$\le \max_{I^q \in \mathbb{N}_m : |I^q| = q} \left\{ \sum_{i \in I^q \bigcap I_{\boldsymbol{x}^*}} \lambda_i^U (f_i(\boldsymbol{x}^*) - f_i^R) + \sum_{i \in I^q \bigcap J_{\boldsymbol{x}^*}} \lambda_i^A (f_i(\boldsymbol{x}^*) - f_i^R) \right\}$$
$$= \max_{I^q \in \mathbb{N}_m : |I^q| = q} \left\{ \sum_{i \in I^q} \left[\max \left\{ \lambda_i^U (f_i(\boldsymbol{x}^*) - f_i^R), 0 \right\} + \min \left\{ \lambda_i^A (f_i(\boldsymbol{x}^*) - f_i^R), 0 \right\} \right] \right\}$$
$$= \hat{s}_R^q(f(\boldsymbol{x}^*), \boldsymbol{\lambda}^U, \boldsymbol{\lambda}^A).$$

Inequality $\hat{s}_R^q(f(\boldsymbol{x}'), \boldsymbol{\lambda}^U, \boldsymbol{\lambda}^A) < \hat{s}_R^q(f(\boldsymbol{x}^*), \boldsymbol{\lambda}^U, \boldsymbol{\lambda}^A)$ contradicts the assumption of $\boldsymbol{x}^*$ being an optimal solution of problem (6). This implies that $\boldsymbol{x}^*$ is weakly Pareto optimal. □

Also the following result similar to Theorem 2.5 can be proven for the TSPASF.

Theorem 3.2. *Given problem* (6) *and let* $\lambda_i^U, \lambda_i^A > 0$ *for all* $i \in \mathbb{N}_m$*. Then among the optimal solutions of problem* (6) *there exists at least one Pareto optimal solution for problem* (1)*.*

Proof. Let $\boldsymbol{x}^*$ be an optimal solution of problem (6) but not Pareto optimal solution. Then according to the definition of Pareto optimality there exists $\boldsymbol{x}' \in X$ such that $f_i(\boldsymbol{x}') \le f_i(\boldsymbol{x}^*)$ for all $i \in \mathbb{N}_m$ and $f_j(\boldsymbol{x}') < f_j(\boldsymbol{x}^*)$ for at least on index $j \in \mathbb{N}_m$. Now

$$\hat{s}_R^q(f(\boldsymbol{x}'), \boldsymbol{\lambda}^U, \boldsymbol{\lambda}^A)$$
$$= \max_{I^q \in \mathbb{N}_m : |I^q| = q} \left\{ \sum_{i \in I^q} \left[\max \left\{ \lambda_i^U (f_i(\boldsymbol{x}') - f_i^R), 0 \right\} + \min \left\{ \lambda_i^A (f_i(\boldsymbol{x}') - f_i^R), 0 \right\} \right] \right\}$$
$$= \max_{I^q \in \mathbb{N}_m : |I^q| = q} \left\{ \sum_{i \in I^q \bigcap I_{\boldsymbol{x}'}} \lambda_i^U (f_i(\boldsymbol{x}') - f_i^R) + \sum_{i \in I^q \bigcap J_{\boldsymbol{x}'}} \lambda_i^A (f_i(\boldsymbol{x}') - f_i^R) \right\}$$
$$\le \max_{I^q \in \mathbb{N}_m : |I^q| = q} \left\{ \sum_{i \in I^q \bigcap I_{\boldsymbol{x}^*}} \lambda_i^U (f_i(\boldsymbol{x}^*) - f_i^R) + \sum_{i \in I^q \bigcap J_{\boldsymbol{x}^*}} \lambda_i^A (f_i(\boldsymbol{x}^*) - f_i^R) \right\} \tag{7}$$
$$= \hat{s}_R^q(f(\boldsymbol{x}^*), \boldsymbol{\lambda}^U, \boldsymbol{\lambda}^A).$$

This completes the proof since if the inequality (7) is strict this contradicts the assumption that $\boldsymbol{x}^*$ is an optimal solution for problem (6). If equality (7) holds, then $\boldsymbol{x}'$ is an optimal solution for problem (6) and Pareto optimal for problem (1). □

Theorem 3.2 implies the following corollary

Corollary 3.3. *If an optimal solution of problem* (6) *is unique, then it is a Pareto optimal solution for problem* (1)*.*

Corollary 3.3 can be proven also in a different way. According to Theorem 2.2, if $\hat{s}_R^q$ is increasing and the solution $\boldsymbol{x}^* \in X$ of problem (6) is unique, then $\boldsymbol{x}^*$ is Pareto optimal.

Now $\hat{s}_R^q$ is increasing according to Definition 2.1, since take $\boldsymbol{x}_1 \in X$ and $\boldsymbol{x}_2 \in X$ with $f_i(\boldsymbol{x}_1) < f_i(\boldsymbol{x}_2)$ for all $i \in \mathbb{N}_m$ and $\lambda_i^U, \lambda_i^A > 0$ for all $i \in \mathbb{N}_m$. Then

$$\begin{aligned}
&\hat{s}_R^q(f(\boldsymbol{x}_1), \boldsymbol{\lambda}^U, \boldsymbol{\lambda}^A) \\
&\quad = \max_{I^q \in \mathbb{N}_m : |I^q| = q} \left\{ \sum_{i \in I^q} \left[\max\left\{ \lambda_i^U (f_i(\boldsymbol{x}_1) - f_i^R), 0 \right\} + \min\left\{ \lambda_i^A (f_i(\boldsymbol{x}_1) - f_i^R), 0 \right\} \right] \right\} \\
&\quad \leq \max_{I^q \in \mathbb{N}_m : |I^q| = q} \left\{ \sum_{i \in I^q} \left[\max\left\{ \lambda_i^U (f_i(\boldsymbol{x}_2) - f_i^R), 0 \right\} + \min\left\{ \lambda_i^A (f_i(\boldsymbol{x}_2) - f_i^R), 0 \right\} \right] \right\} \\
&\quad = \hat{s}_R^q(f(\boldsymbol{x}_2), \boldsymbol{\lambda}^U, \boldsymbol{\lambda}^A).
\end{aligned}$$

The following result guarantees that with TSPASFs it is possible to obtain every weakly Pareto optimal solution.

Theorem 3.4. *If $\boldsymbol{x}^*$ is weakly Pareto optimal for problem* (1)*, then it is a solution of problem* (6) *with $f^R = f(\boldsymbol{x}^*)$ and optimal value is zero.*

Proof. Theorem 2.3 implies this theorem if $\hat{s}_R^q$ is strictly increasing. Now we prove that $\hat{s}_R^q$ indeed is strictly increasing.

Take $\boldsymbol{x}_1 \in X$ and $\boldsymbol{x}_2 \in X$ with $f_i(\boldsymbol{x}_1) < f_i(\boldsymbol{x}_2)$ for all $i \in \mathbb{N}_m$. Since $\lambda_i^U, \lambda_i^A > 0$ for all $i \in \mathbb{N}_m$, $I_{\boldsymbol{x}_1} \subseteq I_{\boldsymbol{x}_2}$ and $J_{\boldsymbol{x}_1} \supseteq J_{\boldsymbol{x}_2}$, we obtain

$$\begin{aligned}
&\hat{s}_R^q(f(\boldsymbol{x}_1), \boldsymbol{\lambda}^U, \boldsymbol{\lambda}^A) \\
&\quad = \max_{I^q \in \mathbb{N}_m : |I^q| = q} \left\{ \sum_{i \in I^q} \left[\max\left\{ \lambda_i^U (f_i(\boldsymbol{x}_1) - f_i^R), 0 \right\} + \min\left\{ \lambda_i^A (f_i(\boldsymbol{x}_1) - f_i^R), 0 \right\} \right] \right\} \\
&\quad = \max_{I^q \in \mathbb{N}_m : |I^q| = q} \left\{ \sum_{i \in I^q \bigcap I_{\boldsymbol{x}_1}} \lambda_i^U (f_i(\boldsymbol{x}_1) - f_i^R) + \sum_{i \in I^q \bigcap J_{\boldsymbol{x}_1}} \lambda_i^A (f_i(\boldsymbol{x}_1) - f_i^R) \right\} \\
&\quad < \max_{I^q \in \mathbb{N}_m : |I^q| = q} \left\{ \sum_{i \in I^q \bigcap I_{\boldsymbol{x}_2}} \lambda_i^U (f_i(\boldsymbol{x}_2) - f_i^R) + \sum_{i \in I^q \bigcap J_{\boldsymbol{x}_2}} \lambda_i^A (f_i(\boldsymbol{x}_2) - f_i^R) \right\} \\
&\quad = \max_{I^q \in \mathbb{N}_m : |I^q| = q} \left\{ \sum_{i \in I^q} \left[\max\left\{ \lambda_i^U (f_i(\boldsymbol{x}_2) - f_i^R), 0 \right\} + \min\left\{ \lambda_i^A (f_i(\boldsymbol{x}_2) - f_i^R), 0 \right\} \right] \right\} \\
&\quad = \hat{s}_R^q(f(\boldsymbol{x}_2), \boldsymbol{\lambda}^U, \boldsymbol{\lambda}^A).
\end{aligned}$$

□

It is also possible to prove that the convexity of the original objective functions of problem (1) preserves also to $\hat{s}_R^q$. However, this proof necessitates the following lemma.

Lemma 3.5. *Let functions f_i be the objective functions of problem* (1) *and sets $I_{\boldsymbol{x}}$ and $J_{\boldsymbol{x}}$ are defined by $I_{\boldsymbol{x}} = \left\{i \in \mathbb{N}_m \mid f_i^R \leq f_i(\boldsymbol{x})\right\}$ and $J_{\boldsymbol{x}} = \left\{i \in \mathbb{N}_m \mid f_i^R > f_i(\boldsymbol{x})\right\}$. If all the functions f_i are convex and $\boldsymbol{x} \in X$, where X is a convex set, then*

1. $I_{\boldsymbol{x}} \subset (I_{\boldsymbol{x}_1} \bigcup I_{\boldsymbol{x}_2})$
2. $J_{\boldsymbol{x}} \supset (J_{\boldsymbol{x}_1} \bigcap J_{\boldsymbol{x}_2})$.

Proof. Let $\boldsymbol{x} = \theta \boldsymbol{x}_1 + (1-\theta)\boldsymbol{x}_2 \in X$, $\theta \in [0,1]$ be a convex combination of $\boldsymbol{x}_1 \in X$ and $\boldsymbol{x}_2 \in X$, where X is convex. Since f_i assumed to be convex the following holds

$$f_i(\boldsymbol{x}) \leq \theta f_i(\boldsymbol{x}_1) + (1-\theta) f_i(\boldsymbol{x}_2) \quad \text{for all } i. \tag{8}$$

In order to proof the first case consider an index i such that $i \in I_{\boldsymbol{x}}$ and $i \notin (I_{\boldsymbol{x}_1} \bigcup I_{\boldsymbol{x}_2})$. Since $i \in I_{\boldsymbol{x}}$ then $f_i^R \leq f_i(\boldsymbol{x})$ and since $i \notin (I_{\boldsymbol{x}_1} \bigcup I_{\boldsymbol{x}_2})$ then $f_i^R > f_i(\boldsymbol{x}_1)$ and $f_i^R > f_i(\boldsymbol{x}_2)$. The latter property implies the following

$$f_i^R = \theta f_i^R + (1-\theta) f_i^R > \theta f_i(\boldsymbol{x}_1) + (1-\theta) f_i(\boldsymbol{x}_2). \tag{9}$$

Since f_i is convex inequality (8) holds and from this, inequality (9) and the assumption that $i \in I_{\boldsymbol{x}}$ follows

$$f_i^R \leq f_i(\boldsymbol{x}) \leq \theta f_i(\boldsymbol{x}_1) + (1-\theta) f_i(\boldsymbol{x}_2) < f_i^R$$

which contradicts the assumption that $i \notin (I_{\boldsymbol{x}_1} \bigcup I_{\boldsymbol{x}_2})$ and thus $i \in (I_{\boldsymbol{x}_1} \bigcup I_{\boldsymbol{x}_2})$. Same holds for every index of $i \in I_{\boldsymbol{x}}$ and thus $I_{\boldsymbol{x}} \subset (I_{\boldsymbol{x}_1} \bigcup I_{\boldsymbol{x}_2})$.

In the second case assume that an index $i \in (J_{\boldsymbol{x}_1} \bigcap J_{\boldsymbol{x}_2})$ and thus $f_i^R > f_i(\boldsymbol{x}_1)$ and $f_i^R > f_i(\boldsymbol{x}_2)$. This property and the convexity of f_i implies

$$f_i^R = \theta f_i^R + (1-\theta) f_i^R > \theta f_i(\boldsymbol{x}_1) + (1-\theta) f_i(\boldsymbol{x}_2) \geq f_i(\boldsymbol{x}).$$

Now $f_i^R > f_i(\boldsymbol{x})$ and thus $i \in J_x$. Same holds for every index of $i \in (J_{\boldsymbol{x}_1} \bigcap J_{\boldsymbol{x}_2})$ and thus $J_{\boldsymbol{x}} \supset (J_{\boldsymbol{x}_1} \bigcap J_{\boldsymbol{x}_2})$. ☐

Theorem 3.6. *Let functions f_i be the objective functions of problem* (1). *If all the functions f_i are convex then $\hat{s}_R^q(f(\boldsymbol{x}), \boldsymbol{\lambda}^U, \boldsymbol{\lambda}^A)$ is also convex when $\boldsymbol{x} \in X$, where X is a convex set.*

Proof. Let $\boldsymbol{x} = \theta \boldsymbol{x}_1 + (1-\theta)\boldsymbol{x}_2 \in X$, $\theta \in [0,1]$ be a convex combination of $\boldsymbol{x}_1 \in X$ and $\boldsymbol{x}_2 \in X$, where X is convex. Since f_i assumed to be convex, inequality (8) holds.

Now according to Lemma 3.5 $I_{\boldsymbol{x}} \subset (I_{\boldsymbol{x}_1} \bigcup I_{\boldsymbol{x}_2})$ and $J_{\boldsymbol{x}} \supset (J_{\boldsymbol{x}_1} \bigcap J_{\boldsymbol{x}_2})$ and we have

$$\begin{aligned}
&\hat{s}_R^q(f(\boldsymbol{x}), \boldsymbol{\lambda}^U, \boldsymbol{\lambda}^A) \\
&= \max_{I^q \in \mathbb{N}_m : |I^q| = q} \left\{ \sum_{i \in I^q} \left[\max\left\{\lambda_i^U (f_i(\boldsymbol{x}) - f_i^R), 0\right\} + \min\left\{\lambda_i^A (f_i(\boldsymbol{x}) - f_i^R), 0\right\} \right] \right\}
\end{aligned}$$

$$
\begin{aligned}
&= \max_{I^q \in \mathbb{N}_m:\, |I^q| = q} \left\{ \sum_{i \in I^q \bigcap I_{\boldsymbol{x}}} \lambda_i^U (f_i(\boldsymbol{x}) - f_i^R) + \sum_{i \in I^q \bigcap J_{\boldsymbol{x}}} \lambda_i^A (f_i(\boldsymbol{x}) - f_i^R) \right\} \\
&\leq \max_{I^q \in \mathbb{N}_m:\, |I^q| = q} \left\{ \sum_{i \in I^q \bigcap (I_{\boldsymbol{x}_1} \bigcup I_{\boldsymbol{x}_2})} \lambda_i^U (\theta f_i(\boldsymbol{x}_1) + (1-\theta) f_i(\boldsymbol{x}_2) - f_i^R) \right. \\
&\qquad \left. + \sum_{i \in I^q \bigcap (J_{\boldsymbol{x}_1} \bigcap J_{\boldsymbol{x}_2})} \lambda_i^A (\theta f_i(\boldsymbol{x}_1) + (1-\theta) f_i(\boldsymbol{x}_2) - f_i^R) \right\} \\
&= \max_{I^q \in \mathbb{N}_m:\, |I^q| = q} \left\{ \sum_{i \in I^q \bigcap (I_{\boldsymbol{x}_1} \bigcup I_{\boldsymbol{x}_2})} \lambda_i^U (\theta (f_i(\boldsymbol{x}_1) - f_i^R) + (1-\theta)(f_i(\boldsymbol{x}_2) - f_i^R)) \right. \\
&\qquad \left. + \sum_{i \in I^q \bigcap (J_{\boldsymbol{x}_1} \bigcap J_{\boldsymbol{x}_2})} \lambda_i^A (\theta (f_i(\boldsymbol{x}_1) - f_i^R) + (1-\theta)(f_i(\boldsymbol{x}_2) - f_i^R)) \right\} \\
&\leq \theta \max_{I^q \in \mathbb{N}_m:\, |I^q| = q} \sum_{i \in I^q \bigcap I_{\boldsymbol{x}_1}} \lambda_i^U (f_i(\boldsymbol{x}_1) - f_i^R) \\
&\quad + (1-\theta) \max_{I^q \in \mathbb{N}_m:\, |I^q| = q} \sum_{i \in I^q \bigcap I_{\boldsymbol{x}_2}} \lambda_i^U (f_i(\boldsymbol{x}_2) - f_i^R) \\
&\quad + \theta \max_{I^q \in \mathbb{N}_m:\, |I^q| = q} \sum_{i \in I^q \bigcap J_{\boldsymbol{x}_1}} \lambda_i^A (f_i(\boldsymbol{x}_1) - f_i^R) \\
&\quad + (1-\theta) \max_{I^q \in \mathbb{N}_m:\, |I^q| = q} \sum_{i \in I^q \bigcap J_{\boldsymbol{x}_2}} \lambda_i^A (f_i(\boldsymbol{x}_2) - f_i^R) \\
&= \theta \max_{I^q \in \mathbb{N}_m:\, |I^q| = q} \left\{ \sum_{i \in I^q \bigcap I_{\boldsymbol{x}_1}} \lambda_i^U (f_i(\boldsymbol{x}_1) - f_i^R) + \sum_{i \in I^q \bigcap J_{\boldsymbol{x}_1}} \lambda_i^A (f_i(\boldsymbol{x}_1) - f_i^R) \right\} \\
&\quad + (1-\theta) \max_{I^q \in \mathbb{N}_m:\, |I^q| = q} \left\{ \sum_{i \in I^q \bigcap I_{\boldsymbol{x}_2}} \lambda_i^U (f_i(\boldsymbol{x}_2) - f_i^R) + \sum_{i \in I^q \bigcap J_{\boldsymbol{x}_2}} \lambda_i^A (f_i(\boldsymbol{x}_2) - f_i^R) \right\} \\
&= \theta \ \hat{s}_R^q (f(\boldsymbol{x}_1), \boldsymbol{\lambda}^U, \boldsymbol{\lambda}^A) + (1-\theta) \ \hat{s}_R^q (f(\boldsymbol{x}_2), \boldsymbol{\lambda}^U, \boldsymbol{\lambda}^A).
\end{aligned}
$$

Now $\hat{s}_R^q(f(\boldsymbol{x}), \boldsymbol{\lambda}^U, \boldsymbol{\lambda}^A) \leq \theta \, \hat{s}_R^q(f(\boldsymbol{x}_1), \boldsymbol{\lambda}^U, \boldsymbol{\lambda}^A) + (1-\theta) \, \hat{s}_R^q(f(\boldsymbol{x}_2), \boldsymbol{\lambda}^U, \boldsymbol{\lambda}^A)$ and thus convexity is preserved. □

Note that in order to guarantee Pareto optimality of a solution produced we can add an augmented term [7] to (5) and the following augmented form

$$
\hat{s}_R^q + \rho \sum_{i \in \mathbb{N}_m} \lambda_i (f_i(\boldsymbol{x}) - f_i^R), \quad \rho > 0
$$

is used in practice.

The advantages of TSPASF are that we always find at least a weakly Pareto optimal solution and the different solutions may be obtained by changing the reference point, weighting vectors or the value of the parameter q. Additionally, compared with PASF there is no restrictions for the location of the reference point and there is no need for any tests of achievability of the reference point since the formulation (5) uses always the right weighting coefficient.

The parameterization used in TSPASF and PASF gives a systematic way to produce possible different (weakly) Pareto optimal solutions from the same preference information with different metrics. The systematic way of this kind may be useful in some interactive methods [7], for example synchronous NIMBUS [10], using several ASFs basing on the same preference information. In order to find different (weakly) Pareto optimal solutions problem (6) can be solved with all values or just some values of the parameter q.

In problem (6) exists a min-max term and thus the problem is nonsmooth even if the objective functions of problem (1) are differentiable. Nonsmooth problems can be solved efficiently with bundle methods [4]. Problem (6) can also be turned into differentiable MINLP form as follows

$$
\begin{aligned}
\min \quad & \alpha && (10)\\
\text{s.t.} \quad & \alpha \geq \sum_{i \in I_s^q} \left[\lambda_i^U (1 - z_i^s)(f_i(\boldsymbol{x}) - f_i^R) + z_i^s \lambda_i^A (f_i(\boldsymbol{x}) - f_i^R) \right], \; s = 1, \ldots, \binom{m}{q} \\
& f_i^R - f_i(\boldsymbol{x}) \leq z_i^s M, \; i \in I^q, \, s = 1, \ldots, \binom{m}{q} \\
& f_i^R - f_i(\boldsymbol{x}) \geq (z_i^s - 1) M, \; i \in I^q, \, s = 1, \ldots, \binom{m}{q} \\
& \boldsymbol{x} \in X, \, z_i^s \in \{0, 1\}, \, i \in \mathbb{N}_m, \, s = 1, \ldots, \binom{m}{q},
\end{aligned}
$$

where s enumerates a q-element subsets I_s^q of an m-element set $\mathbb{N}_m$, z_i^s is a binary variable and M is a sufficiently large number to ensure that $z_i^s = 1$ if and only if $f_i^R - f_i(\boldsymbol{x}) > 0$ and $z_i^s = 0$ if and only if $f_i^R - f_i(\boldsymbol{x}) \leq 0$. Due to the binary variable z_i^s some mixed-integer programming solver, for example generalized αECP algorithm [2], is needed.

According to Theorem 3.6 in the case there all the objectives f_i are convex also $\hat{s}_R^q$ is convex and thus the global optimum can be found. In general, if the objectives f_i are nonconvex, then problem (6) can be solved with bundle method and problem (10) with αECP algorithm but only the local optimum can be guaranteed. If the objectives are assumed to be f°-pseudoconvex then also global optimum can be guaranteed with bundle [6] and αECP [2] method.

4 Case of three objectives

In order to illustrate the functioning of TSPASF, let us consider a special case where $m = 3$ which is $\hat{s}_R^q$ in case of three objectives. Now (5) has the form

$$\hat{s}_R^q(f(\boldsymbol{x}), \boldsymbol{\lambda}^U, \boldsymbol{\lambda}^A) = \max_{I^q \subseteq \{1,2,3\}:|I^q|=q} \left\{ \sum_{i \in I^q} \left[\max\left\{\lambda_i^U (f_i(\boldsymbol{x}) - f_i^R), 0\right\} + \min\left\{\lambda_i^A (f_i(\boldsymbol{x}) - f_i^R), 0\right\} \right] \right\},$$

where $q = 1, 2, 3$, $\boldsymbol{\lambda}^U = (\lambda_1^U, \lambda_2^U, \lambda_3^U)$ and $\boldsymbol{\lambda}^A = (\lambda_1^A, \lambda_2^A, \lambda_3^A)$, $\lambda_i^U, \lambda_i^A > 0$, $i \in \mathbb{N}_3$. Now
for $q = 1$:

$$\begin{aligned}\hat{s}_R^1(f(\boldsymbol{x}), \boldsymbol{\lambda}^U, \boldsymbol{\lambda}^A) = \max\Big\{ & \max\left\{\lambda_1^U (f_1(\boldsymbol{x}) - f_1^R), 0\right\} + \min\left\{\lambda_1^A (f_1(\boldsymbol{x}) - f_1^R), 0\right\}; \\ & \max\left\{\lambda_2^U (f_2(\boldsymbol{x}) - f_2^R), 0\right\} + \min\left\{\lambda_2^A (f_2(\boldsymbol{x}) - f_2^R), 0\right\}; \\ & \max\left\{\lambda_3^U (f_3(\boldsymbol{x}) - f_3^R), 0\right\} + \min\left\{\lambda_3^A (f_3(\boldsymbol{x}) - f_3^R), 0\right\} \Big\}\end{aligned}$$

for $q = 2$:

$$\begin{aligned}\hat{s}_R^2(f(\boldsymbol{x}), \boldsymbol{\lambda}^U, \boldsymbol{\lambda}^A) = \max\Big\{ & \max\left\{\lambda_1^U (f_1(\boldsymbol{x}) - f_1^R), 0\right\} + \min\left\{\lambda_1^A (f_1(\boldsymbol{x}) - f_1^R), 0\right\} \\ & + \max\left\{\lambda_2^U (f_2(\boldsymbol{x}) - f_2^R), 0\right\} + \min\left\{\lambda_2^A (f_2(\boldsymbol{x}) - f_2^R), 0\right\}; \\ & \max\left\{\lambda_1^U (f_1(\boldsymbol{x}) - f_1^R), 0\right\} + \min\left\{\lambda_1^A (f_1(\boldsymbol{x}) - f_1^R), 0\right\} \\ & + \max\left\{\lambda_3^U (f_3(\boldsymbol{x}) - f_3^R), 0\right\} + \min\left\{\lambda_3^A (f_3(\boldsymbol{x}) - f_3^R), 0\right\}; \\ & \max\left\{\lambda_2^U (f_2(\boldsymbol{x}) - f_2^R), 0\right\} + \min\left\{\lambda_2^A (f_2(\boldsymbol{x}) - f_2^R), 0\right\} \\ & + \max\left\{\lambda_3^U (f_3(\boldsymbol{x}) - f_3^R), 0\right\} + \min\left\{\lambda_3^A (f_3(\boldsymbol{x}) - f_3^R), 0\right\} \Big\}\end{aligned}$$

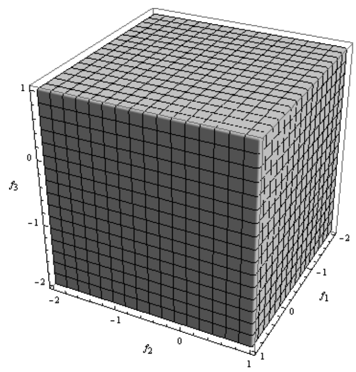
for $q = 3$:

$$\begin{aligned}\hat{s}_R^3(f(\boldsymbol{x}), \boldsymbol{\lambda}^U, \boldsymbol{\lambda}^A) = & \max\left\{\lambda_1^U (f_1(\boldsymbol{x}) - f_1^R), 0\right\} + \min\left\{\lambda_1^A (f_1(\boldsymbol{x}) - f_1^R), 0\right\} \\ & + \max\left\{\lambda_2^U (f_2(\boldsymbol{x}) - f_2^R), 0\right\} + \min\left\{\lambda_2^A (f_2(\boldsymbol{x}) - f_2^R), 0\right\} \\ & + \max\left\{\lambda_3^U (f_3(\boldsymbol{x}) - f_3^R), 0\right\} + \min\left\{\lambda_3^A (f_3(\boldsymbol{x}) - f_3^R), 0\right\}.\end{aligned}$$

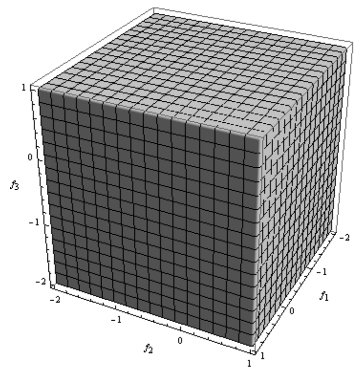
Next we give some graphical illustrations of 1-level sets in 3-dimensional space for both parameterized ASF $\tilde{s}_R^q$ and two-slope parameterized ASF $\hat{s}_R^q$ to see the difference between them. The algebraic form of $\tilde{s}_R^3$ is given in [11]. The view is restricted within a rectangular $\{f = (f_1, f_2, f_3)^T : -2 \leq f_i \leq 1, i \in N_3\}$ and the reference point is assumed to be $f^R = (0, 0, 0)^T$. All the objective functions are assumed to be identity mappings $f_i(\boldsymbol{x}) = \boldsymbol{x}$ and all weighting coefficients are equal to one, $\lambda_1^U, \lambda_2^U, \lambda_3^U, \lambda_1^A, \lambda_2^A, \lambda_3^A = 1$. Figures 2a, 3a and 4a show 1-level set of $\tilde{s}_R^1$, $\tilde{s}_R^2$ and $\tilde{s}_R^3$ respectively and Figures 2b, 3b and 4b show 1-level set of $\hat{s}_R^1$, $\hat{s}_R^2$ and $\hat{s}_R^3$ respectively.

Notice that a choice of the parameter q affects to the shape of R-levels. These R-levels may vary from sharp to flat. Those cases where faces are parallel to the faces $f_1 f_2$, $f_1 f_3$ or $f_2 f_3$ correspond to situation then one of the maxima equals to one and other two are less than one or zero. If sum of two maxima equals to one and the third is less than one or zero, it corresponds the case where faces are sloped and parallel to the coordinate rays. If the all three maxima are positive and sum of them equals to one we have either a flat face (see Figure 4b) or a triangle pyramid with a top vertex $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$ in Figure 3b. These top vertices corresponds to those cases when all three maxima are participating.

Note that if the resulting optimal value of R-level set is positive in case of TSPASF, then it corresponds to the case of unachievable reference point. A negative value signals about reference point achievability. In case of PASF negative value of R-level set is not possible.

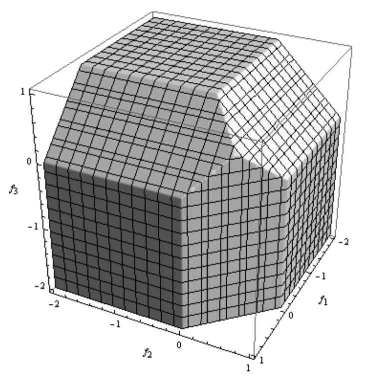


(a) 1-level set for $\tilde{s}_R^1(f(\boldsymbol{x}), \boldsymbol{\lambda})$

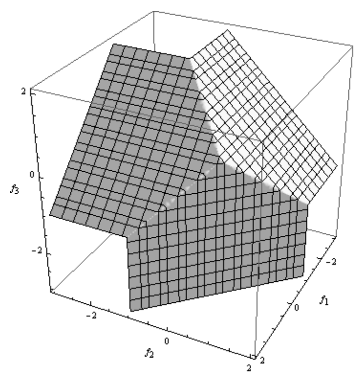


(b) 1-level set for $\hat{s}_R^1(f(\boldsymbol{x}), \boldsymbol{\lambda}^U, \boldsymbol{\lambda}^A)$

Figure 2: 1-level sets for PASF and TSPASF with $q = 1$

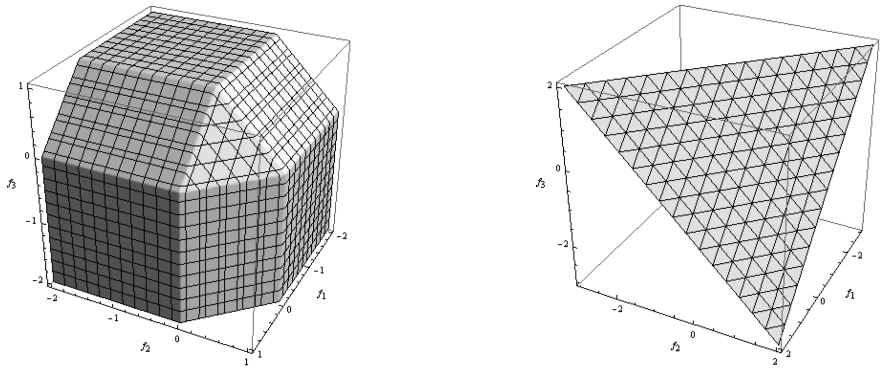


(a) 1-level set for $\tilde{s}_R^2(f(\boldsymbol{x}), \boldsymbol{\lambda})$



(b) 1-level set for $\hat{s}_R^2(f(\boldsymbol{x}), \boldsymbol{\lambda}^U, \boldsymbol{\lambda}^A)$

Figure 3: 1-level sets for PASF and TSPASF with $q = 2$



(a) 1-level set for $\tilde{s}^3_R(f(\boldsymbol{x}), \boldsymbol{\lambda})$ (b) 1-level set for $\hat{s}^3_R(f(\boldsymbol{x}), \boldsymbol{\lambda}^U, \boldsymbol{\lambda}^A)$

Figure 4: 1-level sets for PASF and TSPASF with $q = 3$

5 Computational experiments

In order to explore the behavior of TSPASF several test problems described in [9] are used. The computational calculations are carried out by applying multiobjective proximal bundle method [5]. This method is designed for nonconvex and constrained problems with possibly several objective functions.

Computational tests are divided into two groups based on their convexity. In both cases twenty reference points between the ideal and nadir point are randomly generated and the used weighting vectors $\boldsymbol{\lambda}^U$, $\boldsymbol{\lambda}^A$ are of the form

$$\boldsymbol{\lambda}^U = \frac{1}{f^N - f^R}, \quad \boldsymbol{\lambda}^A = \frac{1}{f^R - f^I}$$

as suggested in [3].

In these computational tests we have concentrated on studying three aspects. The first one is to guarantee that by changing the value of the parameter q we can indeed obtain different solutions, not only in theory but also in practice. Another interesting issue is a quality of the solutions produced or how much solutions differ. The last aspect, which has been singled out and studied, is the computational time.

According to our results at the most of the time the solutions obtained by varying the value of the parameter q differ and the difference is significant both in convex and nonconvex test problems. In the convex case the solutions produced with different values of the parameter q are the same order whereas in the nonconvex test problem $q = 1$ turns out to be clearly the most time-consuming value of the parameter q. Thus with TSPASF by varying metrics between Chebyshev and linear metric, different solutions with good quality are obtained without growing computational efforts compared with the two-slope ASF [3] equaling the case $q = 1$ in TSPASF.

In the following convex and nonconvex test problems are analyzed closer.

5.1 Convex case

We consider closer the following convex Chankong-Haimes test problem [9]

$$\begin{aligned} \min \quad & f(\boldsymbol{x}) = ((x_1-1)^2 + (x_2-1)^2,\, (x_1-2)^2 + (x_2-3)^2,\, (x_1-4)^2 + (x_2-2)^2) \\ \text{s.t.} \quad & x_1 + 2x_2 \leq 10 \\ & 0 \leq x_1 \leq 10 \\ & 0 \leq x_2 \leq 4, \end{aligned}$$

with three objectives and two variables.

In the following, we refer to cases where for one reference point a solution is calculated solution with every value of the parameter $q \in \mathbb{N}_3$ and thus twenty cases are considered due to the number of the generated reference points. Among these randomly generated reference points, there are 40$\%$ of achievable reference points and 60% of unachievable reference points.

In Table 1 the differences of solutions and also the quality of these solutions are analyzed. Here solutions with linear metric ($q = 3$) and Chebyshev metric ($q = 1$) are compared with the case where $q = 2$ and the used metric is something between Chebyshev and linear metric.

The first row in Table 1 presents the percentage value of cases where both values $q = 1$ and $q = 3$ give the same solution as the solution obtained with value $q = 2$. As Table 1 says with values $q = 3$ and $q = 2$ we never obtain the same solution and with values $q = 1$ and $q = 2$ only 15% of cases solutions were the same. Thus by varying the value of the parameter q mostly different solutions are obtained.

In the second row the quality of the solutions by calculating the (relative) distances between solutions is considered. Table 1 reports the averages of these distance calculations. In comparison of distances between solutions obtained with $q = 1$ and $q = 2$ the average distance in objective space is 0.52065. By excluding the cases where the same solution was obtained every distance belongs to the interval from 0.069274 to 1.59813 in objective space. The average distance with values $q = 2$ and $q = 3$ is 0.31203 in objective space and each of these solutions belongs to the interval from 0.11194 to 0.70154 in objective space. Based on these calculations it can be said that the differences between the solutions are significant.

Since the solutions obtained by varying the value of the parameter q actually are various solutions it is also interesting to know what the price of the different solutions is in terms of number of iterations. To explore this aspect, in Table 2 we describe the average number of iterations and function calls needed when calculations are carried out with multiobjective proximal bundle method. In these calculations, both average number of iterations and function calls are approximately on the same order regardless of the value of the parameter q.

Table 1: Differences of solutions

Convex case	$q = 1$	$q = 3$
Cases where $\boldsymbol{x}^*$ with $q = 2$ equals to $\boldsymbol{x}^*$ with	15%	0%
Average relative distance between $f(\boldsymbol{x}^*)$ with $q = 2$ and	0.52065	0.31203
Nonconvex case	$q = 1$	$q = 3$
Cases where $\boldsymbol{x}^*$ with $q = 2$ equals to $\boldsymbol{x}^*$ with	0%	10%
Average relative distance between $f(\boldsymbol{x}^*)$ with $q = 2$ and	201203.24	0.15565

Table 2: Computational times

Convex case	$q = 1$	$q = 2$	$q = 3$
Average number of iterations	12.20	16.95	6.00
Average number of function calls	15.10	22.80	8.10
Nonconvex case	$q = 1$	$q = 2$	$q = 3$
Average number of iterations	30.9	7.45	5.85
Average number of function calls	53.3	15.45	13.6

5.2 Nonconvex case

We scrutinize the following nonconvex Water resources planning test problem [9]

$$\begin{array}{ll} \min & f(\boldsymbol{x}) = (e^{0.001x_1}x_1^{0.02}x_2^2,\ 0.5x_2^2,\ -e^{0.005x_1}x_1^{0.001}x_2^2) \\ \mathrm{s.t.} & 0.01 \leq x_1 \leq 1.3 \\ & 0.01 \leq x_2 \leq 10, \end{array}$$

with are three objectives and two variables as in the example problem in the convex case. Also in this nonconvex case 20 different reference points are generated randomly. There are now 35% of achievable and 65% of unachievable reference points.

In Table 1, the quality of solutions is analyzed also in the nonconvex case. As Table 1 shows with values $q = 1$ and $q = 2$ the same solution was never obtained and with values $q = 2$ and $q = 3$ only 10% of cases the solutions produced are the same. Thus different solutions are obtained by varying the parameter q also in the nonconvex case. By comparing this with the convex case, we see that the total number of the same solutions is now smaller and the same solutions are produced only with values $q = 2$ and $q = 3$ whereas in the convex case the same solutions are produced with values $q = 1$ and $q = 2$.

When we consider the (relative) distances between the solutions produced described in Table 1, we see that in the nonconvex case the distances are significant as was also in the convex example. The average distance between solutions obtained with $q = 1$ and $q = 2$ is 201203.24 in objective space and every distance belongs to the interval from 0.0075685 to 762114 The distances between solutions obtained with $q = 2$ and $q = 3$ is 0.15565 in objective space. By excluding the cases where the same solution was obtained every distance belongs to the interval from 0.00067393 to 0.99999 in objective space.

As said, in Table 2 is described computational times and in the convex case the value of the parameter q do not affect computational time significantly. In the nonconvex case the most time-consuming value of the parameter q is 1 representing Chebyshev type function. In this case the number of iterations and function calls are both significantly larger than corresponding values for other two values of the parameter q.

6 Conclusion

In this paper, we have presented a new family of achievement scalarizing functions based on a parameterization utilizing two different weighting vectors depending on whether the reference point is achievable or not. We have proven that we always find at least weakly Pareto optimal solution and if the solution is unique it is Pareto optimal. We have also proven that every weakly Pareto optimal solution can be produced. Furthermore, we can find different solutions by changing the value of the parameter q, the reference point or weighting vectors. Additionally, to use TSPASFs there is no need for any assumptions about the reference point or for the test of achievability of the reference point.

We have also illustrated the shapes of different R-levels and the computational tests have been performed for both convex and nonconvex problems. These results have shown that the quality of solutions produced is good and the computational time does not grow with different values of the parameter q in the convex case but in nonconvex case Chebyshev type function turn out the most time-consuming value of the parameter q.

The presented TSPASF gives a systematic way to produce different (weakly) Pareto optimal solutions from the same preference information with different metrics. The property of this kind could be used for example in some interactive methods. Different solutions can be calculated with all values of the parameter q or just some of them. Thus it is interesting to know more about how the value q affects the shape of R-level.

Acknowledgements. The research has been financially supported by the Finnish Academy of Science and Letters (the Vilho, Yrjö and Kalle Väisälä Foundation).

References

[1] Ehrgott M.: *Multicriteria Optimization*. Second edition. Springer, Berlin, 2005.

[2] Eronen V.-P., Mäkelä M. M., Westerlund T.: *Extended Cutting Plane Method for a Class of Nonsmooth Nonconvex MINLP problems*. Optimization: A Journal of Mathematical Programming and Operations Research, 2013.

[3] Luque M., Miettinen K., Ruiz A. B., Ruiz F.: *A Two-Slope Achievement Scalarizing Function for Interactive Multiobjective Optimization*. Computers & Operations Research 39, 2012, pp. 1673–1681.

[4] Mäkelä M. M.: *Survey Of Bundle Methods for Nonsmooth Optimization*. Optimization Methods and Software, Vol. 17(1), 2002, pp. 1–29.

[5] Mäkelä M. M.: *Multiobjective Proximal Bundle Method for Nonconvex Nonsmooth Optimization: Fortran Subroutine MPBNGC 2.0*. Reports of the Department of Mathematical Information Technology, Series B, Scientific computing, No. B 13/2003, University of Jyväskylä, Jyväskylä, 2003.

[6] Mäkelä M. M., Karmitsa N., Eronen V.-P.: *On Generalized Pseudo- and Quasiconvexities for Nonsmooth Functions*. TUCS Technical Report, No. 989, Turku Centre of Computer Science, Turku, 2010.

[7] Miettinen K.: *Nonlinear Multiobjective Optimization*. Kluwer Academic Publishers, Boston, 1999.

[8] Miettinen K., Mäkelä M.M.: *On Scalarizing Functions in Multiobjective Optimization*. OR Spectrum 24, 2002, pp. 193–213

[9] Miettinen K., Mäkelä M. M., Kaario K.: *Experiments with Classification-based Scalarizing Functions in Interactive Multiobjective Optimization*. European Journal of Operational Research, 175, 2006, pp. 931–947.

[10] Miettinen K., Mäkelä M. M.: *Synchronous Approach in Interactive Multiobjective Optimization*. European Journal of Operation Research 170, 2006, pp. 909–922.

[11] Nikulin Y., Miettinen K., Mäkelä M. M.: *A New Achievement Scalarizing Function Based on Parameterization in Multiobjective Optimization*. OR Spectrum 34, 2012, pp. 69–87.

[12] Ruiz F., Luque M., Miguel F., del Mar Muñoz M.: *An Additive Achievement Scalarizing Function for Multiobjective Programming Problems*. European Journal of Operational Research 188(3), 2008, pp. 683–694.

[13] Sawaragi Y., Nakayama H., Tanino T.: *Theory of Multiobjective Optimization*. Academic Press, Orlando, 1985.

[14] Wierzbicki, A. P.: *Basic Properties of Scalarizing Functionals for Multiobjective Optimization*. Optimization 8, 1977, pp. 55–60.

[15] Wierzbicki, A. P.: *The Use of Reference Objectives in Multiobjective Optimization*. In: Fandel G., Gal T. (eds) Multiple Criteria Decision Making Theory and Applications. MCDM Theory and Applications Proceedings. Lecture Notes in Economics and Mathematical Systems, vol 177. Springer, Berlin, 1980, pp. 468–486.

[16] Wierzbicki, A. P.: *A Methodological Approach to Comparing Parametric Characterizations of Efficient Solutions*. In Fandel G. et al (eds) Large-scale Modelling and Interactive Decision Analysis. Lecture Notes in Economics and Mathematical Systems, vol 273, Springer, Berlin, 1986, pp. 27–45.

[17] Wierzbicki, A. P.: *On the Completeness and Constructiveness of Parametric Characterizations to Vector Optimization Problems*. OR Spectrum 8, 1986, pp. 73–87.

Turku
Centre *for*
Computer
Science

Joukahaisenkatu 3-5 A, 20520 TURKU, Finland | www.tucs.fi

University of Turku
Faculty of Mathematics and Natural Sciences
- Department of Information Technology
- Department of Mathematics

Turku School of Economics
- Institute of Information Systems Sciences

Åbo Akademi University
- Department of Computer Science
- Institute for Advanced Management Systems Research

ISBN 978-952-12-3084-4
ISSN 1239-1891