International Conference for Mathematical Modeling and Optimization in Mechanics (MMOM 2014)
S. Repin, P. Neittaanmäki, and T. Tuovinen (Eds.)

Proximal Bundle Method for Nonsmooth and Nonconvex Multiobjective Optimization

Marko M. Mäkelä, Napsu Karmitsa and Outi Wilppu

Abstract We present a proximal bundle method for finding weakly Pareto optimal solutions to constrained nonsmooth programming problems with multiple objectives. The method is a generalization of proximal bundle approach for single objective optimization. The multiple objective functions are treated individually without employing any scalarization. The method is globally convergent and capable of handling several nonconvex locally Lipschitz continuous objective functions subject to nonlinear (possibly nondifferentiable) constraints. Under some generalized convexity assumptions, we prove that the method finds globally weakly Pareto optimal solutions. Concluding, some numerical examples illustrate the properties and applicability of the method.

1 Introduction

Nonsmooth (nondifferentiable) optimization problems arise in very many fields of applications, for example, in optimal shape design (see, e.g., [2, 4, 11]), economics [19] and mechanics [17]. On the other hand, instead of one criterion the applications typically have several, often conflicting objectives. During the last three decades the rapid developement has been charasteristic to the areas of nonsmooth (see, e.g., [1, 3, 5, 8, 9, 10, 16, 20]) and multiobjective optimization (see, e.g., [14, 15, 18, 21], separately. Conversely the consideration of both of these approaches in the same framework, i.e. nonsmooth multiobjective optimization, is much less frequent. Thus there exists an increasing demand to be able to solve efficiently optimization problems with several, possible nonsmooth, objective functions.

In this paper we present a proximal bundle based method for constrained nonconvex nonsmooth programming problems with multiple objectives. The method generalizes the proximal bundle approach for single objective optimization [7] by

Marko M. Mäkelä
University of Turku, Department of Mathematics and Statistics, FI-20014 Turku, Finland e-mail: makela@utu.fi

Napsu Karmitsa
University of Turku, Department of Mathematics and Statistics, FI-20014 Turku, Finland e-mail: napsu@karmitsa.fi

Outi Wilppu
University of Turku, Department of Mathematics and Statistics, FI-20014 Turku, Finland e-mail: omwilp@utu.fi

employing the ideas presented in [6, 15, 22]. We can prove, that under some generalized convexity assumptions [13] the method can find globally weakly Pareto optimal solutions. Unlike the most multicriteria optimization methods the multiple objective functions are treated individually without employing any scalarization. The method is readily implementable and descent, i.e., the value of each objective function is expected to get an improvement at each iteration.

The paper is organized as follows. Chapter 2 contains some preliminary concepts and results of nonsmooth and multiobjective optimization theory. The algorithm of the multicriteria proximal bundle (MPB) method is described in Chapter 3. Some convergence results are presented in Chapter 4. Finally, Chapter 5 is devoted to some numerical examples illustrating the properties and applicability of the method.

2 Preliminaries

Let us consider a nonsmooth multiobjective optimization problem of the form

$$\begin{cases} \text{minimize} & \{f_1(x), \dots, f_k(x)\} \\ \text{subject to} & x \in S, \end{cases} \tag{1}$$

where

$$S = \{x \in \mathbb{R}^n \mid g_j(x) \le 0,\ j = 1, \dots, m\}.$$

The objective functions $f_i : \mathbb{R}^n \to \mathbb{R}$ and the constraint functions $g_j : \mathbb{R}^n \to \mathbb{R}$ are supposed to be *locally Lipschitz continuous* (not necessarily smooth nor convex). For a locally Lipschitz continuous function $f : \mathbb{R}^n \to \mathbb{R}$ the *Clarke generalized directional derivative* at x in the direction $d \in \mathbb{R}^n$ is defined by

$$f^\circ(x; d) = \limsup_{\substack{y \to x \\ t \downarrow 0}} \frac{f(y + td) - f(y)}{t}$$

and the *Clarke subdifferential* of f at x by

$$\partial f(x) = \{\xi \in \mathbb{R}^n \mid f^\circ(x; d) \ge \xi^T d \text{ for all } d \in \mathbb{R}^n\},$$

which is a nonempty, convex and compact subset of $\mathbb{R}^n$. Note, that if a locally Lipschitz continuous function attains its local minimum at x^*, then

$$0 \in \partial f(x^*). \tag{2}$$

For a finite maximum of locally Lipschitz continuous functions we have the following subderivation rule.

Theorem 1. *[11] Let $f_i : \mathbb{R}^n \to \mathbb{R}$ be locally Lipschitz continuous at x for all $i = 1, \dots, m$. Then the function*

$$f(x) = \max\,[f_i(x) \mid i = 1,\dots,m]$$

is locally Lipschitz continuous at x and

$$\partial f(x) \subseteq \operatorname{conv}\{\partial f_i(x) \mid f_i(x) = f(x),\ i = 1,\dots,m\}, \tag{3}$$

where conv *denotes the convex hull of a set.*

A function $f : \mathbb{R}^n \to \mathbb{R}$ is *weakly semismooth* if the classical directional derivative

$$f'(x,d) = \lim_{t \downarrow 0} \frac{f(x+td) - f(x)}{t}$$

exists for all x and d, and

$$f'(x,d) = \lim_{t \downarrow 0} \xi(x+td)^T d,$$

where $\xi(x+td) \in \partial f(x+td)$.

A function $f : \mathbb{R}^n \to \mathbb{R}$ is f°*-pseudoconvex*, if it is locally Lipschitz continuous and for all $x, y \in \mathbb{R}^n$

$$f(y) < f(x) \quad \text{implies} \quad f^\circ(x; y-x) < 0$$

and f°*-quasiconvex*, if

$$f(y) \le f(x) \quad \text{implies} \quad f^\circ(x; y-x) \le 0.$$

Note, that a convex function is always f°-pseudoconvex, which again is f°-quasiconvex [13]. Next we present two important properties of f°-pseudoconvex functions.

Theorem 2. *[13] An f°-pseudoconvex function f attains its global minimum at x^*, if and only if*

$$0 \in \partial f(x^*).$$

Theorem 3. *[1] Let $f_i : \mathbb{R}^n \to \mathbb{R}$ be f°-pseudoconvex for all $i = 1,\dots,m$. Then the function*

$$f(x) = \max\,[f_i(x) \mid i = 1,\dots,m]$$

is also f°-pseudoconvex.

Note, that for an f°-quasiconvex function f the level set $\operatorname{lev}_\alpha f := \{x \in \mathbb{R}^n \mid f(x) \le \alpha\}$ is a convex set for all $\alpha \in \mathbb{R}$ [13].

A vector x^* is said to be a *global Pareto optimum* of (1), if there does not exist $x \in S$ such, that

$$f_i(x) \le f_i(x^*) \ \text{ for all } \ i = 1,\dots,k \quad \text{and} \quad f_j(x) < f_j(x^*) \ \text{ for some } \ j.$$

Vector x^* is said to be a *global weak Pareto optimum* of (1), if there does not exist $x \in S$ such, that

$$f_i(x) < f_i(x^*) \quad \text{for all} \quad i = 1, \dots, k.$$

Vector x^* is a *local (weak) Pareto optimum* of (1), if there exists $\delta > 0$ such, that x^* is a global (weak) Pareto optimum on $B(x^*;\delta) \cap S$. Trivially every Pareto optimal point is weakly Pareto optimal.

The *contingent cone* and *polar cone* of set $S \in \mathbb{R}^n$ at point x are defined respectively as

$$\begin{aligned} K_S(x) &= \{d \in \mathbb{R}^n \mid \text{there exist } t_i \downarrow 0 \text{ and } d_i \to d \text{ with } x + t_i d_i \in S\} \\ S^{\leq} &= \{d \in \mathbb{R}^n \mid s^T d \leq 0, \text{ for all } s \in S\}. \end{aligned}$$

The closure of a set S is denoted by $\operatorname{cl} S$. A set $C \subset \mathbb{R}^n$ is a *cone* if $\lambda x \in C$ for all $\lambda \geq 0$ and $x \in C$. We also denote

$$\operatorname{ray} S = \{\lambda x \mid \lambda \geq 0, x \in S\} \quad \text{and} \quad \operatorname{cone} S = \operatorname{ray} \operatorname{conv} S.$$

In other words $\operatorname{ray} S$ is the smallest cone containing S and the conic hull $\operatorname{cone} S$ the smallest convex cone containing S. Furthermore, let

$$F(x) = \bigcup_{i=1}^{k} \partial f_i(x)$$

and

$$G(x) = \bigcup_{j \in J(x)} \partial g_j(x), \text{ where } J(x) = \{j \mid g_j(x) = 0\}.$$

For the optimality condition we pose the following *constraint qualification*

$$G^{\leq}(x) \subseteq K_S(x). \tag{4}$$

Now we can present the following generalized KKT optimality conditions.

Theorem 4. *[13] If x^* is a local weak Pareto optimum of* (1) *and the constraint qualification* (4) *is valid, then*

$$0 \in \operatorname{conv} F(x^*) + \operatorname{cl} \operatorname{cone} G(x^*). \tag{5}$$

Moreover, if f_i are f°-pseudoconvex for all $i = 1, \dots, k$ and g_j are f°-quasiconvex for all $j = 1, \dots, m$, then the condition (5) *is sufficient for x^* to be a global weak Pareto optimum of* (1).

A feasible point $x^* \in S$ is called a *substationary point* for problem (1), if it satisfies the necessary optimality condition (5).

3 Multiobjective Proximal Bundle Method

In this section we develop the MPB (Multiobjective Proximal Bundle) method. The original proximal bundle method of [7] for nonsmooth convex and unconstrained single objective optimization was generalized to handle nonconvex and constrained problems in [11]. The MPB method is a further extension into a multiobjective case. The strategy of handling several objective functions is based on the ideas presented in [6, 15, 22]. The idea, in brief, is to move into a direction where the values of all the objective functions improve simultaneously.

3.1 Direction finding

The MPB method is not directly based on employing any scalarizing function. Some kind of scalarization is, however, needed in deriving the minimization method for all the objective functions. Theoretically, we utilize the *improvement function* H : $\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ defined by

$$H(x,y) = \max\,[f_i(x) - f_i(y),\ g_j(x) \mid i = 1,\dots,k,\ j = 1,\dots,m].$$

Now we obtain the following connection between the improvement function and the problem (1).

Theorem 5. *A necessary condition for $x^* \in \mathbb{R}^n$ to be a global weak Pareto optimum of* (1) *is that*

$$x^* = \operatorname{argmin}_{x\in\mathbb{R}^n} H(x,x^*). \tag{6}$$

Moreover, if f_i are f°-pseudoconvex for all $i = 1,\dots,k$, g_j are f°-quasiconvex for all $j = 1,\dots,m$ and the constraint qualification (4) *is valid, then the condition* (6) *is sufficient for x^* to be a global weak Pareto optimum of* (1).

Proof. Suppose first, that $x^* \in \mathbb{R}^n$ is a global weak Pareto optimum of (1). Since $x^* \in S$ we have $g_j(x^*) \le 0$ for all $j = 1,\dots,m$, thus $H(x^*,x^*) = 0$. If x^* would not be a global minimizer of $H(\cdot,x^*)$, there would exist $y^* \in \mathbb{R}^n$ such that

$$H(y^*,x^*) < H(x^*,x^*) = 0.$$

Then we have $g_j(y^*) < 0$ for all $j = 1,\dots,m$, in other words, $y^* \in S$. Furthermore, $f_i(y^*) < f_i(x^*)$ for all $i = 1,\dots,k$, which contradicts the global weak Pareto optimality of x^*.

Suppose next that (6) holds true. Suppose also that f_i are f°-pseudoconvex for all $i = 1,\dots,k$, g_j are f°-quasiconvex for all $j = 1,\dots,m$ and (4) is valid. Since x^* is a global minimizer of $H(\cdot,x^*)$ by (2), Theorem 3 and Lemma 2.10 of [12] we have

$$
\begin{aligned}
0 \in \partial H(x^*,x^*) &= \operatorname{conv}\left\{ \bigcup_{i=1}^{k} \partial f_i(x^*) \cup \bigcup_{j\in J(x^*)} \partial g_j(x^*) \right\} \\
&= \operatorname{conv}\{F(x^*) \cup G(x^*)\} \\
&\subseteq \operatorname{conv}\{\operatorname{conv} F(x^*) \cup \operatorname{conv} G(x^*)\} \\
&= \lambda \operatorname{conv} F(x^*) + (1-\lambda)\operatorname{conv} G(x^*), \quad \text{where } \lambda \in [0,1].
\end{aligned}
$$

Then for $\lambda \in (0,1]$ we have

$$
\begin{aligned}
0 &\in \operatorname{conv} F(x^*) + \frac{1-\lambda}{\lambda}\operatorname{conv} G(x^*) \\
&\subseteq \operatorname{conv} F(x^*) + \operatorname{ray}\operatorname{conv} G(x^*) \\
&= \operatorname{conv} F(x^*) + \operatorname{cone} G(x^*) \\
&\subseteq \operatorname{conv} F(x^*) + \operatorname{cl}\operatorname{cone} G(x^*).
\end{aligned}
$$

Thus, Theorem 4 implies that x^* is a global weak Pareto optimum of (1). □

Let x^h be the current approximation to the solution of (1) at the iteration h. Then, by Theorem 5, we seek for the search direction d^h as a solution of

$$
\begin{cases}
\text{minimize} & H(x^h+d,x^h) \\
\text{subject to} & d \in \mathbb{R}^n.
\end{cases} \tag{7}
$$

Since (7) still is a nonsmooth problem, we must approximate it somehow. Let us assume for a moment that the problem (1) is convex. We suppose that, at the iteration h besides the current iteration point x^h, we have some auxiliary points $y^j \in \mathbb{R}^n$ from the past iterations and subgradients $\xi_{f_i}^j \in \partial f_i(y^j)$ for $j \in J^h = \{1,\dots,h\}$, $i = 1,\dots,k$, and $\xi_{g_l}^j \in \partial g_l(y^j)$ for $j \in J^h$, $l = 1,\dots,m$. We linearize the objective and the constraint functions at the point y^j by

$$
\begin{aligned}
\bar{f}_{i,j}(x) &= f_i(y^j) + (\xi_{f_i}^j)^T(x-y^j) \quad \text{for all} \quad i=1,\dots,k,\ j \in J^h, \quad \text{and} \\
\bar{g}_{l,j}(x) &= g_l(y^j) + (\xi_{g_l}^j)^T(x-y^j) \quad \text{for all} \quad l=1,\dots,m,\ j \in J^h.
\end{aligned}
$$

Now we can define a convex piecewise linear approximation to the improvement function by

$$
\hat{H}^h(x) = \max\left[\bar{f}_{i,j}(x) - f_i(x^h),\ \bar{g}_{l,j}(x) \mid i=1,\dots,k,\ l=1,\dots,m,\ j \in J^h\right]
$$

and we get an approximation to (7) by

$$
\begin{cases}
\text{minimize} & \hat{H}^h(x^h+d) + \frac{1}{2}u^h\|d\|^2 \\
\text{subject to} & d \in \mathbb{R}^n,
\end{cases} \tag{8}
$$

where $u^h > 0$ is some weighting parameter. The penalty term $\frac{1}{2}u^h\|d\|^2$ is added to guarantee the existence and uniqueness of a solution to (8) and also to keep the approximation local enough. Notice that (8) still is a nonsmooth problem, but due to its minmax-nature it is equivalent to the following (smooth) quadratic problem

$$\begin{cases} \text{minimize} & v + \frac{1}{2}u^h\|d\|^2 \\ \text{subject to} & -\alpha^h_{f_i,j} + (\xi^j_{f_i})^T d \leq v,\ i = 1,\ldots,k,\ j \in J^h \\ & -\alpha^h_{g_l,j} + (\xi^j_{g_l})^T d \leq v,\ l = 1,\ldots,m,\ j \in J^h, \end{cases} \tag{9}$$

where

$$\alpha^h_{f_i,j} := f_i(x^h) - \bar{f}_{i,j}(x^h), \quad i = 1,\ldots,k,\ j \in J^h, \quad \text{and}$$
$$\alpha^h_{g_l,j} := -\bar{g}_{l,j}(x^h), \quad l = 1,\ldots,m,\ j \in J^h,$$

are so-called *linearization errors*.

In the nonconvex case, we replace the linearization errors by *subgradient locality measures*

$$\beta^h_{f_i,j} := \max\,[|\alpha^h_{f_i,j}|\,,\gamma_{f_i}\|x^h - y^j\|^2]$$
$$\beta^h_{g_l,j} := \max\,[|\alpha^h_{g_l,j}|\,,\gamma_{g_l}\|x^h - y^j\|^2],$$

where $\gamma_{f_i} \geq 0$ for $i = 1,\ldots,k$ and $\gamma_{g_l} \geq 0$ for $l = 1,\ldots,m$, ($\gamma_{f_i} = 0$ if f_i is convex and $\gamma_{g_l} = 0$ if g_l is convex).

3.2 Line search

Let (d^h, v^h) be a solution of (9). We perform the following two-point line search strategy, which will detect discontinuities in the gradients of the objective functions. We assume that $m_L \in (0, \frac{1}{2})$, $m_R \in (m_L, 1)$ and $\bar{t} \in (0, 1]$ are some fixed line search parameters. First, we search for the largest number $t^h_L \in [0, 1]$ such that

$$\max\,[f_i(x^h + t^h_L d^h) - f_i(x^h) \mid i = 1,\ldots,k] \leq m_L t^h_L v^h, \quad \text{and}$$
$$\max\,[g_l(x^h + t^h_L d^h) \mid l = 1,\ldots,m] \leq 0.$$

If $t^h_L \geq \bar{t}$, we take a *long serious step*:

$$x^{h+1} = x^h + t^h_L d^h \quad \text{and} \quad y^{h+1} = x^{h+1},$$

if $0 < t^h_L < \bar{t}$, then we take a *short serious step*:

$$x^{h+1} = x^h + t^h_L d^h \quad \text{and} \quad y^{h+1} = x^h + t^h_R d^h$$

and if $t_L^h = 0$, we take a *null step*:

$$x^{h+1} = x^h \quad \text{and} \quad y^{h+1} = x^h + t_R^h d^h,$$

where $t_R^h > t_L^h$ is such that

$$-\beta_{f_i,h+1}^{h+1} + (\xi_{f_i}^{h+1})^T d^h \geq m_R v^h.$$

The iteration is terminated when

$$-\tfrac{1}{2} v^h < \varepsilon_s,$$

where $\varepsilon_s > 0$ is an accuracy parameter supplied by the user.

3.3 Algorithm

Next we aggregate the previous subsections and present the algorithm of the multiobjective proximal bundle method .

Algorithm (MPB)

1. *(Initialization)* Select a feasible starting point $x^1 \in S$, a final accuracy tolerance $\varepsilon_s > 0$, an initial weight $u^1 > 0$, line search parameters $m_L \in (0, \frac{1}{2})$, $m_R \in (m_L, 1)$ and $\bar{t} \in (0, 1]$. Choose the distance measure parameters $\gamma_{f_i} \geq 0$ for $i = 1, \dots, k$ and $\gamma_{g_l} \geq 0$ for $l = 1, \dots, m$, ($\gamma_{f_i} = 0$ if f_i is convex and $\gamma_{g_l} = 0$ if g_l is convex). Set $h := 1$, $y^1 := x^1$ and calculate $\xi_{f_i}^1 \in \partial f_i(y^1)$ for $i = 1, \dots, k$ and $\xi_{g_l}^1 \in \partial g_l(y^1)$ for $l = 1, \dots, m$.
2. *(Direction finding)* Solve the problem (9) in order to get the solution (d^h, v^h).
3. *(Stopping criterion)* If $-\frac{1}{2} v^h < \varepsilon_s$, then STOP.
4. *(Line search)* Find the step sizes $t_L^h \in [0, 1]$ and $t_R^h \in [t_L^h, 1]$. Set

$$x^{h+1} = x^h + t_L^h d^h \quad \text{and} \quad y^{h+1} = x^h + t_R^h d^h.$$

5. *(Updating)* Set $h := h + 1$, calculate $\xi_{f_i}^h \in \partial f_i(y^h)$ for $i = 1, \dots, k$ and $\xi_{g_l}^h \in \partial g_l(y^h)$ for $l = 1, \dots, m$. Choose $J^h \subseteq \{1, \dots, h\}$ and update the weight u^h. Go to Step 2.

The subgradient aggregation strategy due to [5] is used to bound the storage requirements (i.e., the size of the index set J^h). We use the line search algorithm of [11] to produce the step-sizes t_L^h and t_R^h in Step 4, and a modification of the weight updating algorithm of [7] is used to update the weight u^h in Step 5.

4 Convergence Analysis

Next we give two important convergence results. First we prove, that for f°-pseudoconvex functions the algorithm produces a global weak Pareto optimum of the problem (1), while in more general case it end ups with a substationary point.

Theorem 6. *Let f_i and g_j be f°-pseudoconvex and weakly semismooth functions for all $i = 1,\dots,k$ and $j = 1,\dots,m$, and the constraint qualification* (4) *be valid. If the MPB algorithm stops with a finite number of iterations, then the solution is a global weak Pareto optimum of* (1). *On the other hand, any accumulation point of an infinite sequence of solutions generated by the MPB algorithm is global weak Pareto optimum of* (1).

Proof. Due to Theorem 3 the improvement function H is f°-pseudoconvex. The formulation of the MPB algorithm implies, that it is equivalent to the proximal bundle algorithm applied to unconstraint single objective optimization of H. According to the convergence analysis of the standard proximal bundle algorithm (see, e.g., [7, 20]) if it stops with a finite number of iterations, then the solution x^h is a substationary point of a weakly semismooth H, in other words $0 \in \partial H(x^h, x^h)$. Then by Theorem 2 function H attains its global minimum at x^h. Since every f°-pseudoconvex function is also f°-quasiconvex, the first assertion follows from Theorem 5. The proof of the case, when MPB generates an infinite sequence of solutions, goes similarly. □

Note, that in order to guarantee the f°-pseudoconvexity of the improvement function H also the constraint functions g_j are supposed to be f°-pseudoconvex in Theorem 6 although only the f°-quasiconvexity was required in Theorem 5.

Finally we show, that in more general case the algorithm produces substationary points of the problem (1).

Theorem 7. *Let the functions of* (1) *be weakly semismooth. If the MPB algorithm stops with a finite number of iterations, then the solution is a substationary point (i.e. satisfies the necessary optimality condition* (5)*). On the other hand, any accumulation point of an infinite sequence of solutions generated by the MPB algorithm is a substationary point.*

Proof. The proof is analogous to that of Theorem 6. □

5 Numerical Experiments

The efficiency and the reliability of the method is shown by some numerical experiments. The MPB algorithm was implemented in Fortran 77. The test runs have been performed on an Intel® Core™ 2 Duo CPU E8400 (3.00GHz, 2.99GHz) PC computer.

5.1 General tests

We wanted to test the method in different functions classes. First we formulated several f°-pseudoconvex objective functions. Next we combined f°-pseudoconvex functions with classical convex test examples from [1]. Finally some nonconvex test examples [1] being not f°-pseudoconvex nor f°-quasiconvex were solved. Furthermore, some f°-quasiconvex constraint functions were used in all the test examples. Thus the used function classes were

1. f°-pseudoconvex objective functions
2. f°-pseudoconvex + convex objective functions
3. Non(generalized)convex objective functions.

In all the test cases the number of variables n varied from 2 to 4, the number of objective functions k from 2 to 4, and the number of constraint functions m from 0 to 2. The numerical results are presented in Table 1, where the first column refers to the above mentioned test classes, the second column tells the number of the test problems in the class. Finally, the last two columns are devoted to the average of the used iterations and function evaluations, respectively. The last line summarizes the overall average numbers. The parameters of MPB were tuned as follows: $\varepsilon_s = 10^{-5}$, $m_L = 0.01$, $m_R = 0.5$, $\bar{t} = 0.01$, $\gamma_{f_i} = 0.5$ (0 for convex objectives) for $i = 1, \dots, k$ and $\gamma_{g_l} = 0.5$ for $l = 1, \dots, m$. The initial weight was chosen by

$$u^1 = \frac{1}{k} \sum_{i=1}^{k} \|\xi_{f_i}^1\|.$$

Table 1 Computational results

Test class	# Problems	Iterations	Func. calls
1	36	5.1	6.7
2	70	10.4	15.4
3	6	8.7	13.2
All	112	8.6	12.5

In order to summarize the numerical results reported in Table 1 we can state that MPB method seems to be reliable and efficient in all the test classes. The reason why it needed more resources in class 2 with convex problems is the complexity of some single test problems.

5.2 Numerical example

In order to illustrate the functioning of MPB in more details we consider the following problem

$$\begin{cases} \text{minimize} & f_1(x) = \sqrt{\|x\| + 2} \\ & f_2(x) = \max\left\{ -x_1 - x_2,\ -x_1 - x_2 + x_1^2 + x_2^2 - 1 \right\} \\ \text{subject to} & g(x) = \max\left\{ x_1^2 + x_2^2 - 10, 3x_1 + x_2 + 1.5 \right\} \le 0, \end{cases}$$

where f_1 is clearly f°-pseudoconvex (see [1]) , f_2 is convex and g convex and thus f°-quasiconvex. We used first the starting point $x^1 = (-0.5, -0.5)$ and the solution iteration by iteration is reported in Table 2.

Table 2 Results of the numerical example

h	x^h	$(f_1(x), f_2(x))$	Accuracy
0	$(-0.5000000, -0.5000000)$	$(1.645329, 1.000000)$	0.5181928
1	$(-0.4153649, -0.3124033)$	$(1.587367, 0.7277682)$	$0.9826704 \cdot 10^{-2}$
2	$(-0.4360219, -0.2067399)$	$(1.575612, 0.6427618)$	$0.3053751 \cdot 10^{-2}$
3	$(-0.4641460, -0.1123331)$	$(1.574022, 0.5764790)$	$0.4805499 \cdot 10^{-3}$
4	$(-0.4622420, -0.1137555)$	$(1.573542, 0.5759975)$	$0.4842027 \cdot 10^{-4}$
5	$(-0.4620497, -0.1138994)$	$(1.573493, 0.5759491)$	$0.4867036 \cdot 10^{-5}$

The numerical results are depicted in Figure 1, where red and blue colors refer to f_1 and f_2, respectively. Together with the iteration points (black, final solution violet) there can be seen also the unconstrained and constrained optima and contour lines of the objectives. Note, that the Pareto optimal solutions lie on the line segment between red and blue points.

In Figure 2 we illustrate the functioning of MPB by starting the optimization from several starting points. According to the character of the method, MPB projects those points to the Pareto optimal set by using the Chebyshev metric.

6 Conclusions

We have derived a multiobjective version of the proximal bundle method for nonsmooth and nonconvex optimization. The objective functions are treated individually without employing any scalarization. The method is globally convergent and descent, and under some generalized convexity assumptions it can be proved to find globally weakly Pareto optimal solutions. This kind of method is needed in many application areas. Especially, it can be used as a part of interactive multiobjective

Fig. 1 Iteration points of the numerical example

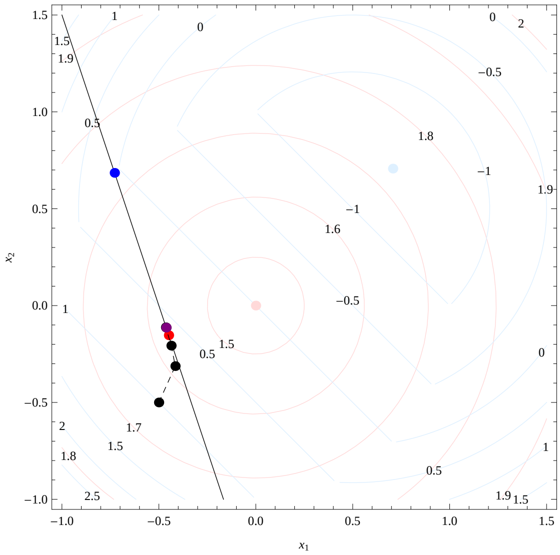
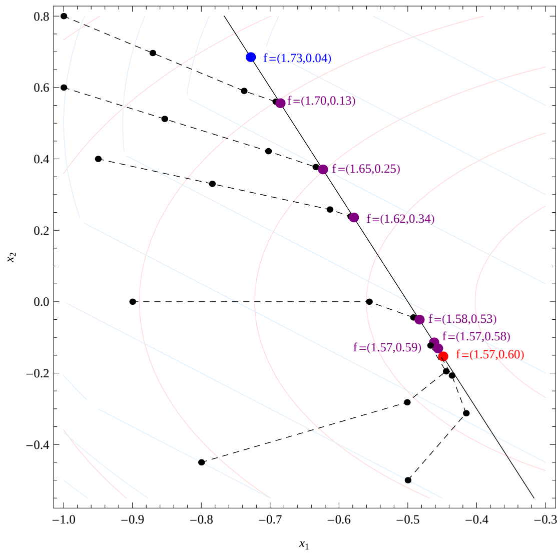


Fig. 2 Results with several starting points



optimization methods producing efficiently (weakly) Pareto optimal counterparts of nonoptimal solutions [14, 15, 18].

Acknowledgements This work is financially supported by the University of Turku and the Vilho, Yrjö and Kalle Väisälä foundation.

References

1. A. Bagirov, N. Karmitsa and M. M. Mäkelä. *Introduction to Nonsmooth Optimization: Theory, Practice and Software*. Springer, 2014.
2. N. V. Banichuk and P. J. Neittaanmäki. *Structural Optimization with Uncertainties*. Solid Mechanics and Its Applications, 162. Springer, Berlin, 2010.
3. M. Gaudioso and M. F. Monaco. Quadratic approximations in convex nondifferentiable optimization. *SIAM Journal on Control and Optimization*, 29:58–70, 1991.
4. J. Haslinger and P. Neittaanmäki. *Finite Element Approximation for Optimal Shape, Material and Topology Design*. John Wiley & Sons, Chichester, 1996.
5. K. C. Kiwiel. *Methods of Descent for Nondifferentiable Optimization*. Lecture Notes in Mathematics, 1133. Springer-Verlag, Berlin, 1985.
6. K. C. Kiwiel. A descent method for nonsmooth convex multiobjective minimization. *Large Scale Systems*, 8(2):119–129, 1985.
7. K. C. Kiwiel. Proximity control in bundle methods for convex nondifferentiable Optimization. *Mathematical Programming*, 46:105–122, 1990.
8. C. Lemaréchal. Nondifferentiable Optimization. In G. L. Nemhauser, A. H. G.Rinnooy Kan and M. J. Todd, editors, *Optimization*, pages 529–572, North-Holland, Amsterdam, 1989.
9. L. Lukšan and J. Vlček. Globally convergent variable metric method for convex nonsmooth unconstrained minimization. *Journal of Optimization Theory and Applications*, 102(3):593–613, 1999.
10. M. M. Mäkelä. Survey of bundle methods for nonsmooth optimization. *Optimization Methods and Software*, 17(1):1–29, 2002.
11. M. M. Mäkelä and P. Neittaanmäki. *Nonsmooth Optimization: Analysis and Algorithms with Applications to Optimal Control*. World Scientific Publishing Co., Singapore, 1992.
12. M. M. Mäkelä, V.-P. Eronen and N. Karmitsa. *On Nonsmooth Optimality Conditions with Generalized Convexities*. TUCS Technical Reports 1056, Turku Centre for Computer Science, Turku, 2012.
13. M. M. Mäkelä, V.-P. Eronen and N. Karmitsa. On Nonsmooth Multiobjective Optimality Conditions with Generalized Convexities. In Th. M. Rassias, C. A. Floudas, and S. Butenko, editors, *Optimization in Science and Engineering*, Springer Verlag, (to appear).
14. K. Miettinen. *Nonlinear Multiobjective Optimization*. Kluwer Academic Publishers, Boston, 1999.
15. K. Miettinen and M. M. Mäkelä. Interactive bundle-based method for nondifferentiable multiobjective optimization: NIMBUS. *Optimization*, 34:231–246, 1995.
16. R. Miflin. A Modification and an extension of Lemarechal's algorithm for nonsmooth minimization. *Mathematical Programming Study*, 176:77–90, 1982.
17. J. J. Moreau, P. D. Panagiotopoulos and G. Strang (Eds.). *Topics in Nonsmooth Mechanics*. Birkhäuser, Basel, 1988.
18. H. Mukai. Algorithms for multicriterion optimization. *IEEE Transactions on Automatic Control*, ac-25(2):177–186, 1980.
19. J. Outrata, M. Kočvara and J. Zowe. *Nonsmooth Approach to Optimization Problems with Equilibrium Constraints*. Theory, Applications and Numerical Results. Kluwer Academic Publishers, Dordrecht, 1998.
20. H. Schramm and J. Zowe. A Version of the bundle idea for minimizing a nonsmooth functions: Conceptual idea, convergence analysis, numerical results. *SIAM Journal on Optimization*, 2:121–152, 1992.
21. R. E. Steuer. *Multiple Criteria Optimization: Theory, Computation, and Applications*. John Wiley & Sons, New York, 1986.
22. S. Wang. Algorithms for Multiobjective and Nonsmooth Optimization. In P. Kleinschmidt, F. J. Radermacher, W. Schweitzer and H. Wildermann, editors, *Methods of Operations Research, 58*, pages 131–142, Athenäum Verlag, Frankfurt am Main, 1989.