

Survey of Piecewise Convex Maximization and PCMP over Spherical Sets

Ider Tseveendorj and Dominique Fortin

Abstract The main investigation in this chapter is concerned with a piecewise convex function which can be defined by the pointwise minimum of convex functions, $F(x) = \min\{f_1(x), \dots, f_m(x)\}$. Such piecewise convex functions closely approximate nonconvex functions, that seems to us as a natural extension of the piecewise affine approximation from convex analysis. Maximizing $F(\cdot)$ over a convex domain have been investigated during the last decade by carrying tools based mostly on linearization and affine separation. In this chapter, we present a brief overview of optimality conditions, methods, and some attempts to solve this difficult nonconvex optimization problem. We also review how the line search paradigm leads to a radius search paradigm, in the sense that sphere separation which seems to us more appropriate than the affine separation. Some simple, but illustrative, examples showing the issues in searching for a global solution are given.

Keywords Piecewise convex • Nonconvex optimization • Nonsmooth optimization

Introduction

Convexity is a central concept in optimization. Solving optimization problems somehow leads to separate the constraint set and the set of points no worse than a given candidate. In the convex optimization case, both sets are convex which makes the separation affordable by a hyperplane. However, when one deals with nonconvex optimization problems, one needs more appropriate tools because both sets or at least one of them can be nonconvex. A decade-long effort for finding such tools is

I. Tseveendorj (✉)

University of Versailles, Université Paris-Saclay, 45 avenue des États-Unis,
78035 Versailles Cedex, France
e-mail: Ider.Tseveendorj@uvsq.fr

D. Fortin

INRIA, Domaine de Voluceau, Rocquencourt, B.P. 105, 78153 Le Chesnay Cedex, France
e-mail: Dominique.Fortin@inria.fr

presented briefly in the beginning of the chapter. The second part of the chapter is mainly devoted towards the ways of use spherical sets for the separation in nonconvex optimization instead of hyperplanes used in convex optimization.

We consider a nonconvex nonsmooth optimization problem:

$$\begin{cases} \text{maximize} & F(x) \\ \text{subject to} & x \in D, \end{cases} \quad (PCMP)$$

where D is a nonempty compact in \mathcal{R}^n and $F : \mathcal{R}^n \rightarrow \mathcal{R}$ is a piecewise convex function.

Definition 1. A function F is called a piecewise convex function iff it can be decomposed into the pointwise minimum of convex functions, namely:

$$F(x) = \min\{f_j(x) \mid j \in M\},$$

where all functions $f_j : \mathcal{R}^n \rightarrow \mathcal{R}$ are convex $j \in M := \{1, 2, \dots, m\}$.

Convex functions, concave functions are particular cases of piecewise convex functions, since clearly F is a convex when $m = 1$ and a concave whenever all functions f_j are affine.

For any real number $\alpha \in \mathcal{R}$ the Lebesgue set of a function f is defined like

$$\mathcal{L}_f(\alpha) = \{x \mid f(x) \leq \alpha\}.$$

A quasiconvex (in particular convex f_j) function has the property that its Lebesgue set (in particular $\mathcal{L}_{f_j}(\alpha)$) is a convex set. Piecewise convex functions have a nice geometrical interpretation that the Lebesgue set of such function is the union of a finite number of convex sets

$$\mathcal{L}_F(\alpha) = \bigcup_{j=1}^m \mathcal{L}_{f_j}(\alpha).$$

For a given $y \in \mathcal{R}^n$ the Lebesgue set $\mathcal{L}_F(F(y))$ defines also, in the sense of maximization, the set of points no better than y . The Lebesgue sets of piecewise convex function $\mathcal{L}_F(F(y))$ are generally nonconvex and can be disconnected or even discrete sets. As a result, the nonconvexity of the objective function $F(\cdot)$ poses the major difficulty for solving piecewise convex maximization problems since they generally have a large number of local optima which are not global optima.

Before solving an optimization problem it is useful to investigate the information about where global optima are attained? At some extreme points, on the boundary, or in the interior of the feasible set, etc. With this respect, it is well known [10, 11, 17, 22] that the global maximum occurs at an extreme point for convex maximization over a convex set while a solution to DC (difference of convex) optimization lies on the boundary of the feasible set. As regards (PCMP), in general,

its large number of local optima including the global maximum can lie anywhere in D . As such, it is computationally very difficult to solve, especially if one wishes to find the global optimum.

The space of piecewise convex functions has been studied in [8]; any continuous nonconvex function can be approximated by piecewise convex one. In addition virtually many optimization problems can, theoretically, be approximated by (*PCMP*). Indeed, the latter justifies the importance of this class of optimization problems as a powerful tool in nonconvex optimization.

The reader is referred to [2, 11, 13, 15, 17, 18, 22, 26] for finding out the close relationship between (*PCMP*) and the other nonconvex optimization problems like convex maximization, reverse convex minimization, DC, Lipschitz optimization, etc.

Despite the concerns mentioned above, (*PCMP*) does not seem to have been extensively studied.

The purpose of this chapter is twofold:

- to present a brief survey of some useful results, optimality conditions, methods, ideas for (*PCMP*);
- to propose a novel approach of solving (*PCMP*) based on nonlinear separation and consider piecewise convex maximization problems over spherical sets (balls, spheres) which play the key role for this new research direction.

A Survey of Studies on PCMP

During the last decade we have been focusing on tools for solving (*PCMP*). This section provides a brief survey of optimality conditions, methods, and some useful ideas for (*PCMP*).

Global Optimality Conditions

Let us quote the article [20] for the global optimality conditions. First, we define an active index set at any z by

$$I(z) = \{i \in M \mid f_i(z) = F(z)\}$$

and at given $k \in M, z \in D$ a special subset of D by

$$D_k(z) = \{x \in D \mid f_j(x) > F(z) \text{ for all } j \in M \setminus \{k\}\}.$$

The following results summarize our findings on optimality conditions so far:

Proposition 1 ([20]). *If $z \in D$ is a global maximum of (*PCMP*) then for all $k \in I(z)$*

$$\partial f_k(y) \cap N(D_k(z), y) \neq \emptyset \text{ for all } y \text{ such that } f_k(y) = F(z)$$

Theorem 1 ([20]). *Let $z \in D$ and assume there exist $k \in I(z)$ and $v \in \mathcal{R}^n$ such that $f_k(v) < f_k(z)$. Then a sufficient condition for z to be the global maximum for (PCMP) is*

$$\partial f_k(y) \cap N(\text{clco}(D_k(z)), y) \neq \emptyset \forall y \text{ such that } f_k(y) = F(z).$$

Here cl , co stand for a closure and a convex hull of a set, respectively.

These necessary and sufficient conditions show that solving (PCMP) leads to choose one function f_k and to maximize it over $D_k(\cdot)$ or over $\text{clco}(D_k(\cdot))$. This is the well known convex maximization problem

$$\begin{cases} \text{maximize} & f(x) \\ \text{subject to} & x \in D, \end{cases} \quad (CM)$$

its optimality conditions have been obtained in [19]

$$\partial f(y) \cap N(D, y) \neq \emptyset \text{ such that } y \text{ such that } f(y) = f(z).$$

In [20], one can find geometric interpretation of the optimality conditions along with their illustrations in some examples.

Methods

Linearization Oriented Algorithm

To our knowledge, the first algorithm for solving (PCMP) has been presented in [4]. The article provides an algorithm based on optimality conditions (Proposition 1 and Theorem 1.) presented in the previous subsection. For the sake of simplicity of presentation it is assumed that functions $f_j(\cdot), j \in M$ are strongly convex quadratic, the domain D a full dimensional polytope.

Quadratic convex maximization problems (the particular case of (PCMP) when $m = 1$) are normally classified as NP-hard. Furthermore, just finding a local maximum of a nonconvex quadratic programming problems is NP-hard too. Thus, even the local solution search for (PCMP) is not trivial. At the same time an efficient algorithm for finding local maxima may be the crucial factor in design of the global maximum search stage. Therefore, in [4] a local search algorithm for (PCMP) has attracted considerable attention. For this important issue of the local search, an algorithm derived from linearization of convex functions is proposed, and its convergence is examined carefully.

Assume that we are given a local solution $y \in D$. In order to improve the best known local solution, according to the optimality conditions, one should look for a

point in $D_k(y)$. Then a practical global search algorithm is provided which combines the local search algorithm with successive inner and outer approximations to $D_k(\cdot)$.

Computational experiments on small examples in $\mathcal{R}^2, \mathcal{R}^3$ are reported, which show the efficiency of the approach. One can find also therein some details of implementation as well as new notions like an intersection graph on the Lebesgue sets $\mathcal{L}_{f_j}(F(\cdot))$, the relationship of the optima of (PCMP) with Helly's theorem from discrete geometry.

Piece Adding TeCHnique (PATCH for Short)

For nonconvex optimization problems, many standard techniques rely on local search and the challenge still remains to escape from a local maximum area.

Among other things the question about how to escape from a local maximum area was investigated in article [8]. The authors first studied the space of piecewise convex functions and showed that this class is closed under operations like addition, positive scalar multiplication, operations "min", F^+, F^- . One can add missing operation "max" into the above list of operations, which has been observed recently.

The following one dimensional example given in [7] well illustrates the idea behind, so-called, the piece adding technique.

$$\begin{cases} \text{maximize } x^2 - 2x, \\ \text{subject to } 0 \leq x \leq 3. \end{cases}$$

Obviously, $x = 0$ is a local maximum, (but not the global !), an accumulation point of local search algorithms with a starting point $x^0 \in [0, 1]$ and there is no clear way to escape from its region, once one is therein. The main idea for escaping is, it makes sense first to add into the objective function a convex piece $x^2 - 4$ issued from the local solution $x = 0$, then solve the following problem

$$\begin{cases} \text{maximize } \min\{x^2 - 2x, x^2 - 4\}, \\ \text{subject to } 0 \leq x \leq 3. \end{cases}$$

Notice that the previous local solution $x = 0$ is a minimum of the new convex piece $x^2 - 4$. Now it is clear (Fig. 1) that any local search will not get back to $x = 0$ and easily finds the global solution $x = 3$ from any starting point.

Similarly for escaping from a local solution y of (PCMP), it is proposed to solve the following problem

$$\text{maximize } \min\{F(x) - F(y), p(x)\}, \text{ subject to } x \in D \quad (PP)$$

where $p(x)$ is a convex function, that will be specified hereafter.

Remark 1. Here we underline that by adding a convex piece into objective function

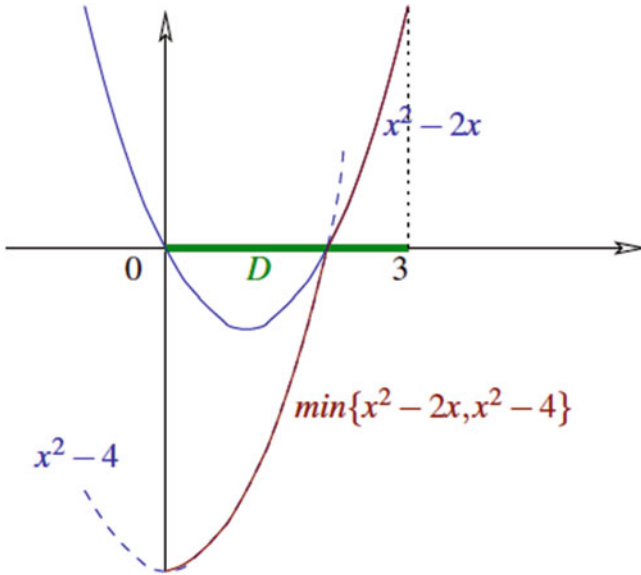


Fig. 1 Piecewise Adding Technique (PATCH) idea

- one cuts off *virtually* from D a subset of points no better than y ; (“*virtual cut*” in the sense that we add a convex piece in objectives rather than a reverse convex constraint $p(x) \geq 0$)
- the objective function remains piecewise convex;
- the space of the problem is unchanged.

The function $p(\cdot)$ may be defined in different ways among which we select two.

Definition 2. Let z be the global solution to (PCMP). A strictly convex function $p_y : \mathcal{R}^n \rightarrow \mathcal{R}$ is called a **patch** around a local solution y of the problem (PCMP) iff

- $p_y(y) < 0$;
- $p_y(x) \geq 0$ for all $x \in D$ such that $F(x) > F(y)$;
- moreover $p_y(z) > 0$

We notice that the conditions for patches are not easy to verify with respect to unknown global solution z , therefore we weaken those conditions and introduce a relaxed function with almost the same features.

Definition 3. A strictly convex function $p_v : \mathcal{R}^n \rightarrow \mathcal{R}$ is called a **pseudopatch** at $v \in D$ iff

- $p_v(v) < 0$ and
- there is $u \in D$ such that $F(u) > F(v)$ and $p_v(u) \geq 0$.

Unlike *the patch*, the *virtual cutting* $p_v(x) \geq 0$ defined by *the pseudopatch* could cut a feasible point v together with some better points (even the global solution!)

but it should leave at least one better point. For practical purposes, pseudopatches are added in the objective function temporarily then, after they are dropped at each improvement since the global solution z can be incidentally cut off by a pseudopatch.

Outline of algorithm

1. Compute a local solution y^k ($k = 1, 2, \dots$) obtained from a feasible point $u \in D$;
2. Construct a strictly convex function $p_k(\cdot)$ as a pseudopatch or a patch around y^k ;
3. Solve

$$\text{maximize } \Phi_k(x), \text{ subject to } x \in D, \quad (P_k)$$

where

$$\Phi_k(x) := \min\{F(x) - F(y^k), p_1(x), p_2(x), \dots, p_k(x)\}.$$

Let u be an optimal solution to (P_k) ;

4. If $p_k(\cdot)$ is a pseudopatch then drop it from $\Phi_k(x)$;
5. Repeat the sequence with $k = k + 1$

In a like manner, after one iteration we either obtain a better point due to pseudopatch or reduce *virtually* the domain by new patch $p_k(\cdot)$ around y^k .

For ease of presenting the main result, let consider two problems for a given local solution y :

$$\text{maximize } F(x) - F(y), \text{ subject to } x \in D, \quad (PCMP)$$

and

$$\text{maximize } \Phi(x) \text{ subject to } x \in D, \quad (PP)$$

where $\Phi(x) := \min\{F(x) - F(y), p_y(x)\}$, $p_y(\cdot)$ is a patch around y .

Assumption 1. Let us assume that $p_y(z) \geq F(z) - F(y)$ at the global solution z .

Proposition 2. If z is a global solution to (PP) and $p_y(\cdot)$ is a patch satisfying Assumption 1 then z solves globally $(PCMP)$ too.

This technique aims also at carrying piecewise affine approximation from convex optimization to piecewise convex approximation for the nonconvex case; since among others DC and Lipschitz functions have locally tight piecewise convex majorants, it shows the potential strength of this approach.

The key tool lies behind the *virtual* cutting function; we call it either a patch to avoid cycling through the same local solutions or a pseudo patch to early detect a better point.

Attractive Force Search Algorithm

Newton's law of universal gravitation states that *any two bodies in the universe attract each other with a force that is directly proportional to the product of their masses and inversely proportional to the square of the distance between them.*

$$\mathcal{F} = \mathcal{G} \frac{m_1 \times m_2}{r^2}$$

defines *the attractive force between two bodies that possess masses m_1, m_2 , respectively.*

Inspired by the well known law, the authors provided an algorithm in article [9] that calculates an improvement from a current local solution y of (PCMP) by using some analogy of Newton's attractive force.

How does one search for an improvement from a feasible solution x ? Up until now, there have been used a local solution search algorithm to find a local solution y with subsequent checking the inclusion $D \subset \mathcal{L}_F(F(y))$ for a possible further improvement. Of course, one can check also the inclusion $D \subset \mathcal{L}_F(F(x))$ directly at $x \in D$. On the other hand, since

$$\mathcal{L}_F(F(\cdot)) = \bigcup_{i=1}^m \mathcal{L}_{f_i}(F(\cdot))$$

an improvement can occur when

$$D \not\subset \bigcup_{i=1}^m \mathcal{L}_{f_i}(F(\cdot)).$$

Anyway, for an improvement we seek a point in D , but outside of all the Lebesgue sets $\mathcal{L}_{f_i}(F(\cdot))$.

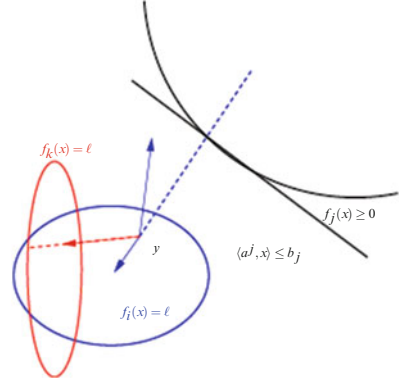
If we imagine each of these nonempty Lebesgue sets $\mathcal{L}_{f_i}(F(y))$ as a convex body that possesses mass, then a direction of improvement at a current point y could be calculated by analogy to Newton's attractive force.

In [9] a feasible set is assumed to be a full dimensional polytope defined by linear constraints

$$D = \{x \in \mathcal{R}^n \mid Ax \leq b\} = \{x \in \mathcal{R}^n \mid \langle a^j, x \rangle \leq b_j, j = 1, 2, \dots, k\}.$$

To deal with the polyhedral domain we consider implicitly convex functions that approximate linear functions: $f_j(x) \approx b_j - \langle a^j, x \rangle$. The linear constraints $b_j - \langle a^j, x \rangle \geq 0$ in the domain, may be seen as the linearization of reverse convex constraints $f_j(x) \geq 0$ for some strictly convex quadratic functions $f_j(\cdot)$; (a huge flat ellipsoid). This viewpoint looks strange at first sight, however, it turns all the Lebesgue sets of constraints, objective functions into convex bodies. The interaction between these convex bodies seems to be related to Newton's law, which gives the main idea of our

Fig. 2 Newton's gravitational force



algorithm. At a given point y , two kind of forces can be involved from each body (the Lebesgue set): attracting and repelling forces.

An example, given in Fig. 2, shows for y

- a repelling force from $\mathcal{L}_{f_i}(\ell)$ body,
- an attracting force from $\mathcal{L}_{f_k}(\ell)$ body

since $f_i(y) < \ell$ and $f_k(y) > \ell$, and in addition, a repelling force from the constraint. More precisely, we select respectively

- the gradient of the function at y multiplied by a positive (resp. negative) scalar in repelling (resp. attracting) case as an analogy for the direction of Newton's force;
- the distance from y to $\mathcal{L}_{f_i}(\ell)$ as an analogy for the distance between the masses.

Now at y , we are able to compute a gravitational force as the weighted sum of all attracting/repelling forces from each body similarly with Newton's attractive force model:

$$G(y) = \sum_{i \in \mathcal{M}} \frac{F(\text{Pr}_{f_i}(y)) - F(y)}{\|\text{Pr}_{f_i}(y) - y\|^2} \frac{\nabla f_i(y)}{\|\nabla f_i(y)\|},$$

where $\text{Pr}_{f_i}(y)$ stands for projection of y on the Lebesgue set $\text{Pr}_{\mathcal{L}_{f_i}(\ell)}(y)$ with ℓ , the best known level set value.

Remark 2. When i is active $i \in I(y)$, in other words $f_i(y) = \ell$ the projection is replaced by c_i the center of $\mathcal{L}_{f_i}(\ell)$ which is a minimum of $f_i(\cdot)$ and the gradient by the direction $x - c_i$ (the mass of the convex body is supposed at center). For active linear constraints, the center is assumed at infinity so that no attracting/repelling force contributes to $G(x)$.

Before describing the algorithm, let us introduce a set which appears very convenient in data structure.

Definition 4. The resolving border of $F(\cdot)$ at level set value ℓ is defined like the set of points:

$$\text{rb}(F, \ell) = \{x \in \mathcal{R}^n \mid \exists i, f_i(x) = \ell, F(x) \leq \ell\}.$$

For each i , we may distinguish three parts:

1. $f_i(x) = \ell, f_j(x) > \ell$ for all $j \neq i$
2. $f_i(x) = \ell, f_j(x) \geq \ell$ for all $j \neq i$
3. $f_i(x) = \ell, f_j(x) > \ell$ for some $j \neq i$.

We notice that cases (1.) and (2.) easily lead to a better point.

A resolving border data structure stores points according to active functions ordered by decreasing values of $F(\cdot)$.

For sake of consistency with previous versions of global search algorithm [4, 8] we still refer to **Newton's Attractive force Search Algorithm** as a local search, but it clearly outperforms a strict local searching since it cruises in the surroundings of the resolving border.

PCMP local Search (Newton's Attractive force Search Algorithm)

- $(D, M, w^j \in D)$
- Initialize RB=(key,sortedset) associating map
- $y^1 = \text{setAndBetter}(w^j, \text{RB})$
- **if** y^1 not null **then return** y^1
- $\text{dir} = G(w^j)$; Newton attraction at w^j
- $u = w^j + \alpha \text{dir}$; $\alpha > 0$ gives the nearest intersection with the resolving border
- $y^2 = \text{setAndBetter}(u, \text{RB})$
- **if** y^2 not null **then return** y^2
- **else return** findBetter(RB)

Relationship with Other Optimization Problems

- **Combinatorial Optimization** Many practical problems give rise to combinatorial optimization problems can be formulated by the binary constraint $x \in \{0, 1\}^n$, by the permutation constraint like the assignment problems. The following two articles [5, 6] are devoted to continuous approaches for combinatorial optimization problems. The hardness of these problems consists in nonconvex domains. The current subsection highlights a couple of ideas about how to solve some combinatorial optimization problems using (PCMP).

We consider the multiknapsack problem [5]

$$\begin{cases} \max & \langle c, x \rangle \\ \text{s.t.} & Ax \leq b \\ & x \in \{0, 1\}^n \end{cases} \quad (\text{MKP})$$

Since the binary constraint $x_i \in \{0, 1\}$ can be written like

$$x_i(x_i - 1) \geq 0, 0 \leq x_i \leq 1,$$

(MKP) has an equivalent continuous formulation

$$\begin{cases} \max & \langle c, x \rangle \\ \text{s.t.} & Ax \leq b \\ & x \in [0, e] \\ & x_i(x_i - 1) \geq 0, \forall i = 1, \dots, n, \end{cases}$$

where $e = (1, \dots, 1)^\top \in \mathcal{R}^n$. Introducing a function

$$\varphi(x) = \min\{x_i(x_i - 1) \mid i = 1, 2, \dots, n\},$$

we replace n constraints $x_i(x_i - 1) \geq 0$, $i = 1, \dots, n$ with a constraint $\varphi(x) \geq 0$ and obtain an equivalent to (MKP) problem

$$\begin{cases} \max & \varphi(x) \\ \text{s.t.} & Ax \leq b \\ & x \in [0, e] \\ & \langle c, x - y \rangle \geq 0. \end{cases}$$

for an admissible point y of (MKP). The latter is the piecewise convex maximization problem with n pieces.

Let assume that there is suitable index set's division J_1, \dots, J_m such that $\bigcup_{i=1}^m J_i = \{1 \dots n\}$ and $J_i \cap J_j = \emptyset$, $\forall i \neq j$. Then it holds also

$$x \in \{0, 1\}^n \Leftrightarrow f_1(x) \geq 0, \dots, f_m(x) \geq 0, 0 \leq x \leq e,$$

where $f_i(x)$ denotes $f_{J_i}(x)$ defined like

$$f_J(x) = \sum_{i \in J} \left(x_i - \frac{1}{2} \right)^2 - \frac{|J|}{4}$$

for any $J \subseteq \{1 \dots n\}$ of $|J|$ elements. In this way we obtain (PCMP) where number of pieces m much less than the dimension of the problem ($m < n$)

$$\begin{cases} \max & \min\{f_i(x) \mid i = 1 \dots m\} \\ \text{s.t.} & Ax \leq b \\ & x \in [0, e] \\ & \langle c, x - y \rangle \geq 0. \end{cases}$$

But, in practice, the nature of pieces as well as a number of pieces are crucial for solving (*PCMP*). One can find in [5] a practical algorithm with $m = 2$ of this approach along with computational results in contrast with the best known solutions found by heuristics from combinatorial optimization.

What concerns the permutation constrained problems investigated in [6], in order to retain our focus on common features of continuous optimization, first, we study shortly relationship between (*PCMP*) and DC optimization, then as a result (*PCMP*) formulation for the well known quadratic assignment problems (*QAP*) is given. We consider a problem of maximization of a difference of two convex functions f, g over a convex compact $D \subset \mathcal{R}^n$

$$\begin{aligned} & \max f(x) - g(x) \\ & \text{s.t. } x \in D. \end{aligned}$$

It is straightforward to turn it into (*PCMP*) of dimension $n + 1$ by introducing another variable $t = g(x)$ and splitting the equality constraint into a convex constraint $g(x) - t \leq 0$ while the converse inequality is *dualized* to add a new piece as $F(x, t) = \min\{f(x) - t, g(x) - t\}$. Then, it is equivalent to solving the following problem

$$\begin{aligned} & \max F(x, t) \\ & \text{s.t. } x \in D, t \in \mathcal{R} \\ & \quad g(x) - t \leq 0. \end{aligned}$$

- **Multicriteria Optimization**

We consider multicriteria optimization problem

$$\begin{cases} \text{minimize } \Omega(x), \\ \text{subject to } x \in D \end{cases} \quad (\text{MOP})$$

where $\Omega(\cdot)$ is a vector valued function from \mathcal{R}^n to \mathcal{R}^m whose components are the convex functions $f_i(\cdot), i \in M = \{1, 2, \dots, m\}$ namely :

$$\Omega(x) = (f_1(x), \dots, f_m(x))^\top \in \mathcal{R}^m.$$

Let us formally recall the definition of Pareto optimal solutions.

Definition 5. A solution $y \in D$ is called Pareto optimal, if there is no $x \in D$ such that $f_i(x) \leq f_i(y), i = 1, \dots, m$ and $f_j(x) < f_j(y)$ for some $j \in M$.

An interesting relationship between (*PCMP*) and multicriteria optimization is presented [21] and afterwards in [9].

We recall also some basic definitions and results from multicriteria optimization.

Definition 6.

– $y \in D$ is called weakly Pareto optimal if there is no $x \in D$ such that

$$f_i(x) < f_i(y), \forall i = 1, \dots, m$$

– $y \in D$ is called strictly Pareto optimal if there is no $x \in D, x \neq y$ such that

$$f_i(x) \leq f_i(y), \forall i = 1, \dots, m \text{ and } f_j(x) < f_j(y) \text{ for some } j \in M.$$

We define also so called *the level curve, the strict level set* of $f(\cdot)$ at α respectively

$$\mathcal{L}_f^=(\alpha) = \{x \mid f(x) = \alpha\}, \quad \mathcal{L}_f^<(\alpha) = \{x \mid f(x) < \alpha\}$$

Theorem 2 ([1], Chap. 2). *Let $y \in D$ then*

1. *y is strictly Pareto optimal if and only if*

$$\bigcap_{i=1}^m \mathcal{L}_{f_i}(f_i(y)) = \{y\}.$$

2. *y is Pareto optimal if and only if*

$$\bigcap_{i=1}^m \mathcal{L}_{f_i}(f_i(y)) = \bigcap_{i=1}^m \mathcal{L}_{f_i}^=(f_i(y)).$$

3. *y is weakly Pareto optimal if and only if*

$$\bigcap_{i=1}^m \mathcal{L}_{f_i}^<(f_i(y)) = \emptyset.$$

The following results summarize our findings so far:

Proposition 3. *If y is local maximum to (PCMP) such that $y \in \text{int}(D)$ then for some $\epsilon > 0$*

$$F(y) = \min_{x \in B(y, \epsilon)} \max\{f_i(x) \mid i \in I(y)\},$$

where $B(y, \epsilon)$ stands for the ball of radius ϵ , centered at y .

Let us denote $\Omega_N(\cdot) \in \mathcal{R}^{|N|}$ with corresponding components $f_i(\cdot)$ for $i \in N \subset M$. Proposition 3 together with Theorem 2 provide the following relationship.

Theorem 3. *If y is local maximum to (PCMP) such that $y \in \text{int}(D)$ then y is strictly Pareto optimal to the following multicriteria optimization problem:*

$$\begin{cases} \text{minimize } \Omega_{f(y)}(x), \\ \text{subject to } x \in D \end{cases}$$

Remark 3. Weakly Pareto optimality of y is implied by Proposition 6.5 from [1].

Abstract Nonlinear Covering Method

The remainder of the chapter is devoted to new approach of solving (*PCMP*).

For all maximization problems, in particular for (*PCMP*), clearly, z is the global maximum iff all points of the domain are no better than z , in other words:

$$D \subset \mathcal{L}_F(F(z)).$$

Since as we observed early $\mathcal{L}_F(F(y)) = \bigcup_{i=1}^m \mathcal{L}_{f_i}(F(y))$, the above inclusion means that

- for all x no better than z (i.e. $F(x) \leq F(z)$) there exists $j \in M$ such that

$$x \in \mathcal{L}_{f_j}(F(z)),$$

- if there is $u \in D$ better than z (i.e. $F(z) < F(u)$) then u does not belong to any Lebesgue set:

$$u \notin \mathcal{L}_{f_i}(F(z)) \text{ for all } i \in M.$$

In order to present the main idea of a new approach for solving (*PCMP*), we give a definition along with an abstract result on an equivalence of problems.

Definition 7. An open subset C satisfying conditions

$$C \subset \mathcal{L}_F(F(y)) \text{ and } C \neq \text{int}(\mathcal{L}_F(F(y)))$$

is called a **covering set at level $F(y)$** .

Proposition 4. Let y be a feasible point for (*PCMP*) such that

$$F(y) = \max\{F(x) \mid x \in D\} - \delta$$

for some $\delta > 0$. Let also C be a covering set at level $F(y)$. Then the following problem is equivalent to (*PCMP*):

$$\begin{cases} \text{maximize } F(x) \\ \text{subject to } x \in D \setminus C. \end{cases} \quad (CC)$$

The main algorithmic feature now looks like

- to cover the feasible set (the domain) by a union of covering sets.
- if the domain is covered by C totally, then stop and the global optimum is found.
- otherwise, solve problem (CC) for an improvement.

In other words, one have to construct an “(union of covering sets)” such that

$$D \subset (\text{union of covering sets}) \subset \mathcal{L}_F(F(z)).$$

Starting with an initial guess of covering sets, a method bootstraps its way up to ever more accurate “sandwich” approximations to answer “the global optimum” or “improvement”. What concerns the covering set, the first that comes to mind, is use balls (spherical set) as a simpler nonlinear shape.

(PCMP) over Spherical Sets

Let $S(c; \rho)$ and $B(c; \rho)$ be respectively a sphere and a ball of center c and radius ρ . This section is devoted to the following two (PCMP), solutions of them are going to be key tools in nonlinear separation as stated in previous section.

$$\left\{ \begin{array}{l} \text{maximize } F(x), \\ \text{subject to } x \in S(c; \rho), \end{array} \right. \quad \left\{ \begin{array}{l} \text{maximize } F(x), \\ \text{subject to } x \in B(c; \rho). \end{array} \right.$$

We turn our attention to the problem over a sphere, that is a problem of piecewise maximization over a nonconvex feasible set with an empty interior. Since all feasible points are degenerated even the local search for this problem is worth-while to study. We consider a convex function f and notice that the KKT optimality condition for an optimizer of f (either min or max) over $S(c, \rho)$ implies collinearity of gradient $\nabla f(u)$ and $u - c$. It is well known that the projected gradient method, from a reasonably good starting point, quickly converges to the optimal solution. If no meaningful guess is known, the gradient $\nabla f(c)$ is a good ray direction to find a starting point for maximizing f over the sphere; for minimizing we start from the antigradient $-\nabla f(c)$.

Since $F(x) = \min\{f_1(x), \dots, f_m(x)\}$ a piecewise convex function; using $2 \times m$ times the above optimization paradigm (for both $\bar{l}^j = \arg \min_S f_j(x)$ and $\bar{u}^j = \arg \max_S f_j(x)$) yields a *sparse* set of points which the actual values $F(\bar{l}^j)$ and $F(\bar{u}^j)$ may be computed for all $j = 1, \dots, m$. We call for the best value of $\max\{F(\bar{l}^i), F(\bar{u}^j) \mid i, j = 1, \dots, m\}$ as sparse piecewise convex maximization value over a sphere and denote by z the point of this value.

Now we borrow from [9] the resolving border heuristic that focuses on points on the level set in the vicinity of $D_k(\cdot)$ for some k that are likely to improve the easy sparse optimizers.

We use the following two examples as the main lead to illustrate the different steps in Piecewise Convex Maximizing:

Algorithm 1 PCMP-over-sphere: Resolving border for $\text{argloc max}_{x \in S} F(x)$

Input: $c, \rho, F = \{f_j \mid j = 1, m\}$
Output: $z = \text{argloc max}_{x \in S} F(x)$

```

pq=PriorityQueue(sparseoptimize(S, F)) // decreasing values of sparse argloc max
while pq ≠ ∅ do
  u = pop(pq)
  if F(u) better then
    z = u
  end if
  for m = 1, M do
    v = Prfj(x)=F(z)(u)
    uj = c + ρ  $\frac{v-c}{\|v-c\|}$  push(pq, F(uj), uj) //enqueue value for unprocessed points only
  end for
end while
return z

```

$$\begin{aligned}
 F_1^{2D} &= \min \left\{ \begin{aligned} &x_1^2 + (x_2 + 4)^2 - 36, \\ &(x_1 + 8)^2 + (x_2 - 3)^2 - 36, \\ &x_1^2 + (x_2 - 8)^2 - 16, \\ &(x_1 - 8)^2 + (x_2 - 3)^2 - 53, \\ &(x_1 - 10)^2 + (x_2 + 10)^2 - 4 \end{aligned} \right\}, \\
 F_2^{2D} &= \min \left\{ \begin{aligned} &x_1^2 + (x_2 + 2)^2 - 9, \\ &9(x_1 + 3)^2 + 4x_2^2 - 36, \\ &(x_1 + 1)^2 + (x_2 - 4)^2 - 4, \\ &\frac{1}{9}(x_1 - 3)^2 + \frac{1}{36}(x_2 - 4)^2 - 1, \\ &(x_1 - 5)^2 + (x_2 + 5)^2 - 1 \end{aligned} \right\}
 \end{aligned}$$

respectively in spherical domain (spheres/balls) with $(c, \rho) = ([0, 0], 4)$.

Despite there is no proof of global optimality, we experimented an effective behavior to find points on the sphere well ordered by the priority queue and likely to lay close to $\cap D_j(z)$ (the best z found is drawn as a blue dot on Fig. 3).

In the remainder we concentrate on elementary coding steps, starting on optimization over a sphere towards more involved recursive calls of (*PCMP*) over a ball.

There is a subtlety when $v = c$ for defining w ; we choose $w = \frac{z+c}{2}$ betting on a point in $\cap D_k(z)$ (early detection of a better point). In both cases Figs. 3 and 4 blue points were found quickly which localize the global solution areas.

Concluding Remarks

In this chapter we have presented an overview of studies on piecewise convex maximization problems: necessary and sufficient global optimality conditions, methods including linearization oriented algorithm, piece adding technique (patch for short), and Newton's attractive force search algorithm (nasa for short). A novel

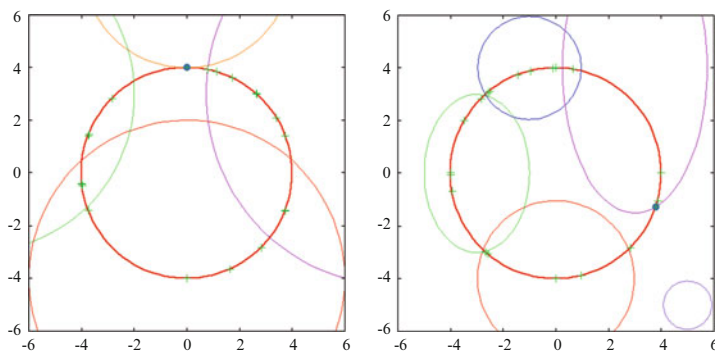


Fig. 3 Maximize F_i^{2D} over a sphere, $i = 1, 2$ (left, right)

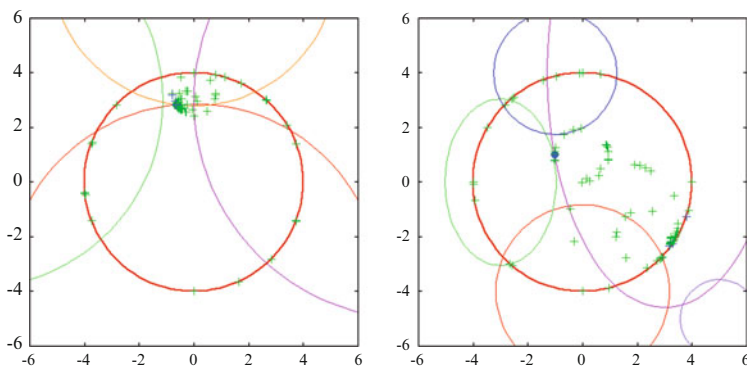


Fig. 4 Maximize F_i^{2D} over a ball, $i = 1, 2$ (left, right)

approach of solving (PCMP) based on a nonlinear separation (instead of the traditional affine separation) has been proposed. We call it a nonlinear covering method. As a rule the approach opens up a field which should be investigated more thoroughly: a choice of the covering sets, algorithms of solving subproblems over them, etc.

Our aim is now to construct an “(union of covering sets)” such that

$$D \subset (\text{union of covering sets}) \subset \mathcal{L}_F(F(z))$$

to check the global optimality of z for (PCMP).

As a preliminary we have considered piecewise convex maximization problems over spherical sets (balls, spheres) which play the key role for this new research direction. Algorithms along with some computational results with simple, but illustrative, examples have been reported.

Algorithm 2 PCMP-over-ball: Recursive radius search for $\arg\text{loc}\max_{x \in B} F(x)$

Input: $c, \rho, F = \{f_j \mid j = 1, \dots, m\}$
Output: $z = \arg\text{loc}\max_{x \in B} F(x)$
 $r = \rho$
 $z = c$
for $l = 1, \text{maxiterations}$ **do**
if $r \leq \varepsilon$ **then**
return z
end if
 $z = \text{PCMP-over-sphere}(S(c; r), F)$ // $z = \arg\text{loc}\max_S F$ State $\text{loc}z = z$
for $j = 1, m$ **do**
 $v = \text{Pr}_{f_j(x)=F(z)}(u)$ // projection on non active function

if $d(v, z) \leq \varepsilon$ or $d(v, z) > 2r$ **then**

continue // skip projection too close or too far

end if
if $d(v, c) > r$ **then**
 $w = c + \frac{r}{d(v, c)}(v - c)$ // outside $B(c; r) \Rightarrow w \in S(c; r)$
else
 $w = z + \frac{d(v, z)}{r}(c - z)$ // inside $B(c; r) \Rightarrow w \in B(c; r)$
end if
if w better $\text{loc}z$ **then**
 $\text{loc}z = w$
end if
end for
if $\text{loc}z$ better z **then**
 $\rho = \min\{\|z - \text{loc}z\|, \|c - \text{loc}z\|\}$
 $z = \text{PCMP-over-ball}(B(\text{loc}z, \rho), F)$
 $r = \|c - z\|$
else
 $r = \frac{r}{2}$ // halve radius

end if
end for
return z

Acknowledgements This research benefited from the support of the FMJH “Program Gaspard Monge for optimization and operations research”, and from the support from EDF.

The authors acknowledge use of the IBM ILOG CPLEX under the academic initiative license.

The authors would like to thank the anonymous referee for his/her useful comments on this chapter.

References

1. Ehrgott, M.: Multicriteria Optimization. Lecture Notes in Economics and Mathematical Systems, vol. 491. Springer, Berlin (2000)
2. Enkhbat, R.: Quasiconvex Programming. National University of Mongolia, Ulaanbaatar (2004)
3. Floudas, C.A., Pardalos, P.M., Adjiman, C.S., Esposito, W.R., Gümüş, Z.H., Harding, S.T., Klepeis, J.L., Meyer, C.A., Schweiger, C.A.: Handbook of Test Problems in Local and Global Optimization. Nonconvex Optimization and its Applications, vol. 33. Kluwer Academic Publishers, Dordrecht (1999)

4. Fortin, D., Tseveendorj, I.: Piecewise-convex maximization problems: algorithm and computational experiments. *J. Global Optim.* **24**(1), 61–77 (2002). doi:10.1023/A:1016221931922. <http://dx.doi.org/10.1023/A:1016221931922>
5. Fortin, D., Tseveendorj, I.: Piecewise convex maximization approach to multiknapsack. *Optimization* **58**(7), 883–895 (2009). doi:10.1080/02331930902945033. <http://dx.doi.org/10.1080/02331930902945033>
6. Fortin, D., Tseveendorj, I.: A trust branching path heuristic for permutation problems. *Int. J. Pure Appl. Math.* **56**(3), 329–343 (2009)
7. Fortin, D., Tseveendorj, I.: Piece adding technique for convex maximization problems. *J. Global Optim.* **48**(4), 583–593 (2010). doi:10.1007/s10898-009-9506-z. <http://dx.doi.org/10.1007/s10898-009-9506-z>
8. Fortin, D., Tseveendorj, I.: Piecewise convex maximization problems: piece adding technique. *J. Optim. Theory Appl.* **148**(3), 471–487 (2011). doi:10.1007/s10957-010-9763-5. <http://dx.doi.org/10.1007/s10957-010-9763-5>
9. Fortin, D., Tseveendorj, I.: Attractive force search algorithm for piecewise convex maximization problems. *Optim. Lett.* **6**(7), 1317–1333 (2012). doi:10.1007/s11590-011-0395-y. <http://dx.doi.org/10.1007/s11590-011-0395-y>
10. Horst, R., Tuy, H.: *Global Optimization: Deterministic Approaches*, 2nd edn. Springer, Berlin (1993). doi:10.1007/978-3-662-02947-3. <http://dx.doi.org/10.1007/978-3-662-02947-3>
11. Horst, R., Pardalos, P.M., Thoai, N.V.: *Introduction to Global Optimization. Nonconvex Optimization and its Applications*, vol. 48, 2nd edn. Kluwer Academic Publishers, Dordrecht (2000). doi:10.1007/978-1-4615-0015-5. <http://dx.doi.org/10.1007/978-1-4615-0015-5>
12. Mordukhovich, B.S.: *Variational Analysis and Generalized Differentiation. I. Basic Theory* *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*, vol. 330. Springer, Berlin (2006)
13. Paulavičius, R., Žilinskas, J.: *Simplicial Global Optimization. Springer Briefs in Optimization*. Springer, New York (2014). doi:10.1007/978-1-4614-9093-7. <http://dx.doi.org/10.1007/978-1-4614-9093-7>
14. Penot, J.P.: Duality for anticonvex programs. *J. Global Optim.* **19**(2), 163–182 (2001)
15. Sergeev, Y.D., Kvasov, D.: *Diagonal global optimization methods* (in Russian). Fizmatlit, Moscow (2008)
16. Strekalovskii, A.S.: On the problem of the global extremum. *Dokl. Akad. Nauk SSSR* **292**(5), 1062–1066 (1987)
17. Strekalovskii, A.S.: *Elements of Nonconvex Optimization* (in Russian). NAUKA, Novosibirsk (2003)
18. Törn, A., Žilinskas, A.: *Global Optimization. Lecture Notes in Computer Science*, vol. 350. Springer, Berlin (1989). doi:10.1007/3-540-50871-6. <http://dx.doi.org/10.1007/3-540-50871-6>
19. Tseveendorj, I.: On the conditions for global optimality. *J. Mong. Math. Soc.* (2), 58–61 (1998)
20. Tseveendorj, I.: Piecewise-convex maximization problems: global optimality conditions. *J. Global Optim.* **21**(1), 1–14 (2001). doi:10.1023/A:1017979506314. <http://dx.doi.org/10.1023/A:1017979506314>
21. Tseveendorj, I., Fortin, D.: Pareto optima and local optimality of piecewise convex maximization problems. In: *Proceedings of the 15-th Baikal International Conference on Optimization Methods and Their Applications, “Mathematical Programming”*, vol. 2, pp. 25–29. ISDCT SB of the Russian Academy of Sciences’ Publisher, Irkutsk (2011)
22. Tuy, H.: *Convex Analysis and Global Optimization. Nonconvex Optimization and Its Applications*, vol. 22. Kluwer Academic Publishers, Dordrecht (1998)
23. Zălinescu, C.: On the maximization of (not necessarily) convex functions on convex sets. *J. Global Optim.* **36**(3), 379–389 (2006)

24. Zhiglyavskii, A.A., Zhilinskas, A.G.: *Metody poiska globalnogo ekstremuma*. “Nauka”, Moscow (1991). With an English summary
25. Zhigljavsky, A., Žilinskas, A.: *Stochastic Global Optimization*. Springer Optimization and Its Applications, vol. 9. Springer, New York (2008)
26. Zhilinskas, A.: *Globalnaya optimizatsiya*. “Mokslas”, Vilnius (1986). *Aksiomatika statisticheskikh modelei, algoritmy, primeneniya*. [Axiomatics of statistical models, algorithms, and applications.] With Lithuanian and English summaries