European Journal of Operational Research 214 (2011) 485–492

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Continuous Optimization

An interior proximal method in vector optimization

Kely D.V. Villacorta [1], P. Roberto Oliveira *,[2]

Federal University of Rio de Janeiro, COPPE/UFRJ – Systems Engineering and Computer Science Program, P.O. Box 68511, 21945-972 Rio de Janeiro, RJ, Brazil

ARTICLE INFO

Article history:
Received 17 June 2009
Accepted 4 May 2011
Available online 10 May 2011

Keywords:
Interior point methods
Vector optimization
C-convex
Positively lower semicontinuous

ABSTRACT

This paper studies the vector optimization problem of finding weakly efficient points for maps from $\mathbb{R}^n$ to $\mathbb{R}^m$, with respect to the partial order induced by a closed, convex, and pointed cone $C \subset \mathbb{R}^m$, with nonempty interior. We develop for this problem an extension of the proximal point method for scalar-valued convex optimization problem with a modified *convergence sensing condition* that allows us to construct an interior proximal method for solving VOP on nonpolyhedral set.

1. Introduction

Let C be a closed, convex, and pointed cone in $\mathbb{R}^m$, with $\text{int}(C) \neq \emptyset$, where $\text{int}(C)$ denotes the interior of set C. Then, C induces a partial order $\preceq$ in $\mathbb{R}^m$, given by $y \preceq y'$ if, and only if, $y' - y \in C$, with its associated relation $\prec$ given by $y \prec y'$ if, and only if, $y' - y \in \text{int}(C)$. Our goal is to analyze methods to find a weakly efficient solution of the following problem

$$(VOP) \quad C - \min\{F(x) : x \in \Omega\},$$

where $F : \mathbb{R}^n \to \mathbb{R}^m \cup \{+\infty_C\}$ is a proper positively lower semicontinuous C-convex map and Ω is a nonempty convex closed subset of $\text{dom}(F)$, with nonempty interior.

Recently, (Huang and Yang, 2004; Sylva and Crema, 2004; Bonnel et al., 2005; Graña Drummond and Svaiter, 2005; Antczak, 2006; Jeyakumar et al., 2006; Ceng and Yao, 2007; Graña Drummond et al., 2008; Fliege et al., 2009; Gregório and Oliveira, 2010; Gutiérrez et al., 2010), among others, have worked in multiobjective and/or vector optimization, and gotten extensions of several theoretical results and numerical methods, well-known in the literature for scalar optimization. In those extensions, they define the iterates in the vector-valued case by considering the order $\preceq$ in Y, where Y is a real Banach space, or particularly Euclidean spaces, mimicking, whenever possible, the role of the usual order in $\mathbb{R}$ to the corresponding algorithm for scalar-valued optimization. In the meantime, we admit the possibility that F takes value $+\infty_C$.

The decade has seen considerable progress in the theory of proximal point methods for scalar-valued problems, several of them are based on generalized distances, see, e.g., (Auslender and Teboulle, 2006; Kaplan and Tichatschke, 2004, 2007a,b). Now, we give a brief description of this kind of methods. Consider the following convex minimization problem:

$$\inf\{f(x) : x \in \Omega\}, \tag{1.1}$$

where $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is a proper, lower semicontinuous and convex function. The proximal point method generates a sequence $\{x^k\} \subset \mathbb{R}^n$ corresponding to the recursion

$$g^{k+1} + \beta_k \nabla_1 d(x^{k+1}, x^k) = 0, \tag{1.2}$$

where $g^{k+1} \in \partial_{\varepsilon_k} f(x^{k+1})$, $\{\beta_k\}$ is a bounded exogenous sequence of positive real numbers (called regularization parameters), $\nabla_1 d(\cdot, y)$ denotes the gradient map of function $d(\cdot, y)$ with respect to the first variable, d is some proximity measure, and x^k the current iterate.

With the choice $d(x,y) = 2^{-1}\|x - y\|^2$ and $\varepsilon_k = 0$, for all $k \in \mathbb{N}$, we may recover the classic proximal algorithm, whose origins can be traced back to the 1960s, see, e.g., (Moreau, 1965; Martinet, 1970, 1972; Rockafellar, 1976). In this case, the sequence $\{x^k\}$ produced by the above algorithm does not necessarily belong to $\text{int}(\Omega)$. Thus the proximal term $d(x,y)$ will play the role of a distance-like function, satisfying certain properties, see Section 2, which will force the iterates of the produced sequence to stay in $\text{int}(\Omega)$ and thus automatically eliminate the constraints.

It has been proved in Auslender and Teboulle (2006) that the sequence $\{x^k\}$ generated by the proximal point method (1.2) belongs to $\text{int}(\Omega)$ and converges to some solution of problem (1.1), under some conditions on d.

* Corresponding author. Tel.: +55 21 2562 8672/3; fax: +55 21 2562 8676.
E-mail addresses: kelydvv@cos.ufrj.br (K.D.V. Villacorta), poliveir@cos.ufrj.br (P.R. Oliveira).

[1] Supported by CAPES/CNPq, IEL Nacional, Brazil.
[2] Partial supported by CNPq, Brazil.

0377-2217/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.ejor.2011.05.006

In the so-called inexact versions of the method, x^{k+1} need not to be an exact solution of the subproblem (1.2), but only an approximate solution of it. Clearly, the inexact version is essential if one wants to get the convergence results for actual implementations of this kind of method. Inexact versions were proposed as early as 1976, see, e.g., (Rockafellar, 1976), in which the kth subproblem was allowed to be solved within a prescribed tolerance ε_k, and it was demanded that $\sum_{k=0}^{\infty}\varepsilon_k < \infty$. Similar error criteria, requiring summability of the tolerances, appeared in several papers later on.

Kaplan and Tichatschke (2007a,b, 2008) use Bregman functions with a modified "convergence sensing condition" that enables them to construct a generalized proximal method for solving (1.1) on sets that are not necessarily polyhedral. For this, they admit a successive approximation of the operator ∂f and an inexact calculation of the proximal iterate. This method generates sequences $\{x^k\} \subset \mathbb{R}^n$ and $\{\theta^k\} \subset \mathbb{R}^n$ corresponding to the recursion

$$\theta^{k+1} \in Q^k(x^{k+1}) + \beta_k \nabla_1 D_{\hat{h}}(x^{k+1}, x^k) \tag{1.3}$$

where $\{\beta_k\}$ is a bounded exogenous sequence of positive real numbers, $D_{\hat{h}}$ is the Bregman distance induced by $\hat{h} = h + \eta$ and $\partial f \subset Q^k \subset \partial_{\varepsilon_k} f$.

It has been proved in Kaplan and Tichatschke (2007a,b) that if h, η, θ^k and ε_k satisfy certain properties, the sequence $\{x^k\}$ generated, by the proximal point method (1.3), belongs to int(Ω) and converges to some solution of problem (1.1).

The above discussion refers, of course, to the proximal method for scalar-valued convex optimization. This paper consists of the extension of both the exact proximal method (1.2), $\varepsilon_k = 0$, and inexact counterpart (1.3) to the vector-valued optimization problem introduced at the beginning of this section. Basically, in the exact case the kth subproblem will consist of finding weakly efficient solutions of

$$F(x) + \beta_k d(x, x^k) e_k \tag{1.4}$$

restricted to the set $\Omega_k \subset \Omega$ defined as $\Omega_k = \{x \in \Omega : F(x) \preceq F(x^k)\}$, where d is a proximal distance on int(Ω) and e_k is an exogenously selected vector belonging to int(C), with $\|e_k\| = 1$.

For our inexact version, we consider the positive polar cone $C^* \subset \mathbb{R}^m$, given by $C^* = \{z \in \mathbb{R}^m : \langle y, z\rangle \geqslant 0 \text{ for all } y \in C\}$, and the indicator function I_{Ω_k}, of set Ω_k, defined as above. We take an exogenous sequence $\{z_k\} \subset C^*$, with $\|z_k\| = 1$ for all $k \in \mathbb{N}$, and define, at iteration k, the function $f_k : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ as

$$f_k(x) = \langle F(x), z_k\rangle + I_{\Omega_k}. \tag{1.5}$$

Then we take as x^{k+1} any vector $x \in \Omega$ such that there exists $\theta^{k+1} \in \mathbb{R}^n$, $\varepsilon_k \in \mathbb{R}_+$ satisfying

$$\theta^{k+1} \in Q^k(x) + \beta_k \langle e_k, z_k\rangle \nabla_1 \tilde{d}(x, x^k), \tag{1.6}$$

where $\partial f_k \subset Q^k \subset \partial_{\varepsilon_k} f_k$ and $\tilde{d}$ is a convenient proximal distance.

We will establish that any sequence generated by either our exact or inexact version converges to a weakly efficient solution of F on Ω under the following hypothesis:

- (**A1**) F is C-convex on Ω, i.e., $F(\lambda x + (1-\lambda)x') \preceq \lambda F(x) + (1-\lambda)F(x')$ for all $x, x' \in \Omega$ and all $\lambda \in [0,1]$.
- (**A2**) F is positively lower semicontinuous, i.e., $\langle F(\cdot), z\rangle$ is lower semicontinuous for every $z \in C^*$.
- (**A3**) The set $(F(x^0) - C) \cap F(\Omega)$ is C-complete; i.e., for every sequence $\{a_k\} \subset \Omega$, with $a_0 = x^0$, such that $F(a_{n+1}) \preceq F(a_k)$ for all $k \in \mathbb{N}$, there exists $a \in \Omega$ such that $F(a) \preceq F(a_k)$ for all $k \in \mathbb{N}$.

In the absence of assumption A3, we establish convergence results, namely, that the generated sequence is a minimizing one for our problem, meaning that $\{F(x_k)\}$ approaches the set of infimal values of F, that will be showed in Propositions 4.2 and 4.3 of Section 4.

The paper is organized as follows: Section 2 recalls and introduces some required preliminary material. Section 3 states formally the problem. The exact version of the method is analyzed in Section 4. Finally, Section 5 develops the inexact version.

2. Preliminaries

We adopt the following convex analysis notation (Rockafellar, 1970). For a proper convex and lsc function $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, its effective domain is set by dom$f = \{x : f(x) < +\infty\}$, and for all $\epsilon \geqslant 0$ its $\epsilon-$subdifferential at x is defined by $\partial_\epsilon f(x) = \{g \in \mathbb{R}^n : \forall\, z \in \mathbb{R}^n,\ \ f(z) + \epsilon \geqslant f(x) + \langle g, z - x\rangle\}$, which coincides with the usual subdifferential $\partial f = \partial_0 f$ whenever $\epsilon = 0$. We set dom$\partial f = \{x \in \mathbb{R}^n : \partial f(x) \neq \emptyset\}$. For any convex set $S \subset \mathbb{R}^n$, I_S denotes the indicator function of S, ri(S) its relative interior, int(S) its interior, $\overline{S}$ its closedness and $N_S(x) = \partial I_S(x) = \{v \in \mathbb{R}^n : \langle v, z - x\rangle \leqslant 0\ \forall\, z \in S\}$ the normal cone to S at $x \in S$.

Now, we recall some useful properties of convex analysis and nonnegative sequences.

Lemma 2.1 Rockafellar (1970, Corollary 6.5.2). *Let S_1 be a convex set. Let S_2 be a convex set contained in $\overline{S_1}$ but not entirely contained in the relative boundary of S_1. Then* ri(S_2) $\subset$ ri(S_1).

Lemma 2.2 Rockafellar (1970, Theorem 27.4). *Let f be a proper convex function, and let S be a nonempty convex set. In order that x^* be a point where the infimum of f, relative to S, is attained, it is sufficient that there exists a vector $y^* \in \partial f(x^*)$ such that $-y^*$ is normal to S at x^*. This condition is necessary, as well as sufficient, if* ri(dom f) *intersects* ri(S), *or if S is polyhedral and* ri(dom f) *merely intersects S.*

Lemma 2.3 Polyak (1987, Lemma 2.2.2). *Let $\{\xi_k\}$, $\{\upsilon_k\}$ and $\{\zeta_k\}$ be nonnegative sequences of real numbers satisfying $\xi_{k+1} \leqslant (1 + \upsilon_k)\xi_k + \zeta_k$ and such that $\sum_{k=1}^{\infty}\zeta_k < \infty$, $\sum_{k=1}^{\infty}\upsilon_k < \infty$. Then, the sequence $\{\xi_k\}$ converges.*

2.1. Proximal distances

In this part we remember definitions of proximal distance d and induced proximal distance H, presented in Auslender and Teboulle (2006). Furthermore, we will introduce a subclass of the induced proximal distance H, which is slightly modified of that was given by them.

Definition 2.1. A function $d : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_+ \cup \{+\infty\}$ is called a *proximal distance* with respect to an open nonempty convex set $S \subset \mathbb{R}^n$ if for each $y \in S$ it satisfies the following properties:

- (P_1) $d(\cdot, y)$ is proper, lsc, convex, and continuously differentiable on S;
- (P_2) dom $d(\cdot, y) \subset \overline{S}$ and dom$\nabla_1 d(\cdot, y) = S$, where $\nabla_1 d(\cdot, y)$ denotes the gradient map of function $d(\cdot, y)$ with respect to the first variable;
- (P_3) $d(\cdot, y)$ is level bounded on $\mathbb{R}^n$, i.e., $\lim_{\|x\| \to \infty} d(x, y) = +\infty$;
- (P_4) $d(y, y) = 0$.

Just as in Auslender and Teboulle (2006), we also denote by $\mathcal{D}(S)$ the family of functions d satisfying Definition 2.1.

The next definition associates for each given $d \in \mathcal{D}(S)$ another function satisfying some convenient properties.