Subgradient-like algorithms for variable inequalities[*]

J.Y. Bello Cruz[1]   G. Bouza Allende[2]   L. R. Lucambio Pérez[3]

[1]Universidade Federal de Goiás, IME, CEP 74001-970 GO, Goiânia, Brazil, e-mail: yunier@impa.br
[2] Departamento de Matemática Aplicada, Facultad de Matemática y Computación, Universidad de La Habana, Habana, Cuba.
email: gema@matcom.uh
[3]Instituto de Matemática e Estatística, Universidade Federal de Goiás, Campus Samambaia, CEP 74001-970 GO, Goiânia, Brazil
email: lrlp@mat.ufg.br

Abstract

Consider the variable inequality problem, that is to find a solution of the inclusion given by the sum of a function and a point-to-cone application. This problem can be seen as a generalization of the classical system inequality problem taking a variable order structure. Exploiting this special structure, we propose variants of the subgradient algorithm for solving a system of variable inequalities. Their convergence is analyzed under convex-like conditions.

Keywords: Convexity; Projection methods; Strong convergence; Subgradient algorithm

Mathematical Subject Classification(2010): 90C25 47N10

1 Introduction

Given $F : \mathbb{R}^n \to \mathbb{R}^m$ and C a non-empty subset of $\mathbb{R}^n$, consider the problem of finding $x \in C$ such that the following inequalities are satisfied

$$F(x) \leq 0,$$

or equivalently that

$$0 \in F(x) + \mathbb{R}^m_+.$$

This problem can be generalized as follows

$$\text{find } x \in C \text{ such that } 0 \in F(x) + K, \tag{1}$$

where $K \subset \mathbb{R}^m$ is a cone.

Suppose that K is a closed, convex and pointed cone in $\mathbb{R}^m$. A partial order $\preceq$, induced in $\mathbb{R}^m$ by K, is defined as $\bar{y} \preceq y$ if and only if $y - \bar{y} \in K$. Then Problem (1) can be formulated as the K-inequalities problem

$$\text{find } x \in C \text{ such that } F(x) \preceq_K 0. \tag{2}$$

If the function F is convex with respect to the partial order given by K, i.e. for all $\bar{x}, x \in \mathbb{R}^n$ and $\alpha \in [0, 1]$, $F(\alpha \bar{x} + (1 - \alpha)x) \preceq_K \alpha F(\bar{x}) + (1 - \alpha)F(x)$, then Problem (2) is called K-convex inequalities problem; see [19, 34, 21].

The variable case considers the set valued application $K : \mathbb{R}^m \rightrightarrows \mathbb{R}^m$, where $K(y)$ is a pointed, convex and closed cone, for all $y \in \mathbb{R}^m$. Then the variable version of (1) can be reformulated as finding a point $x \in C$ such that that

$$0 \in F(x) + K(F(x)). \tag{3}$$

This problem is an inclusion model, which has been studied in many papers; see for instance [41, 12]. In this case, the set valued application $K(F(x))$ has a special structure, which will be strongly exploited. Based on the variable order structures given by $z \preceq_{K(z)} y$ if and only if $y - z \in K(z)$, Problem (3) is equivalent to the K-inequality problem

$$\text{find } x \in C \text{ such that } F(x) \preceq_{K(F(x))} 0. \tag{4}$$

The solution set of this problem is denoted by S^*.

In this paper we propose a subgradient approach for solving Problem (4), which combines a subgradient iteration with a simple projection step, onto the intersection of C with suitable halfspaces containing S^*. For the proposed conceptual algorithm we present two variants called Algorithm R, based on Robinson's subgradient algorithm given in [37], and Algorithm S, which corresponds to a special modification of the subgradient algorithm for scalar problems exposed in [4]. Their main difference lies in how the projection is done. For the convergence of the algorithm, we assume that $S^* \neq \emptyset$ and that function F is convex with respect to the variable order structure defined by $K(F(x))$. That is, F is K-convex if for all $\bar{x}, x \in \mathbb{R}^n$ and $\alpha \in [0, 1]$,

$$F(\alpha \bar{x} + (1 - \alpha)x) \preceq_{K(F(\alpha \bar{x} + (1 - \alpha)x))} \alpha F(\bar{x}) + (1 - \alpha)F(\bar{x}).$$

In this case we say that Problem (4) is a K-convex inequality problem, where K is understood as a set valued function.

Note that if K is a constant application, Problem (4) corresponds with (1), This model has been already studied in [37]. Moreover, if K is the Pareto cone, i.e., $K = \mathbb{R}^m_+$, it is equivalent to the convex feasibility problem, which has been well-studied and has many applications in optimization theory, approximation theory, image reconstruction and so on; see [36, 40, 11].

The paper is organized as follows. First we outline the main definitions and preliminary results. In Section 3 some analytical results for K-convex functions are shown. Section 4 is devoted to the presentation of the algorithms and finally in section 5 their convergence is shown.

2   Preliminaries

In this section, we present some definitions and results, which are needed in the convergence analysis. Next we deal with the so called Fejér convergence and its properties. We begin with some classical notations.

The inner product in $\mathbb{R}^n$ is denoted by $\langle \cdot, \cdot \rangle$, the norm, determined by this inner product, by $\| \cdot \|$ and $B[x, \rho]$ is the closed ball centered at $x \in \mathbb{R}^n$ with radio ρ, i.e., $B[x, \rho] = \{y \in \mathbb{R}^n : \|y - x\| \leq \rho\}$. A set valued application $K : \mathbb{R}^m \rightrightarrows \mathbb{R}^m$ is closed iff $gr(K) = \{(x, y) : F(x) = y\}$ is a closed set.

The set C will be a closed and convex subset of $\mathbb{R}^n$. For an element $x \in \mathbb{R}^n$, we define the orthogonal projection of x onto C, $P_C(x)$, as the unique point in C, such that $\|P_C(x) - y\| \leq \|x - y\|$ for all $y \in C$. Next proposition shows two well known facts of the orthogonal projections, which will be used throughout the paper.

Proposition 1. *For all $x, y \in \mathbb{R}^n$ and all $z \in C$,*

i) $\langle x - P_C(x), z - P_C(x) \rangle \leq 0.$

ii) $\|P_C(x) - P_C(y)\|^2 \leq \|x - y\|^2 - \|(P_C(x) - x) - (P_C(y) - y)\|^2.$

Proof. See Proposition 1.1.9 of [7]. □

Now we present the Fejér convergence.

Definition 1. *Let S be a nonempty subset of $\mathbb{R}^n$. A sequence $\{x^k\}$ is said to be Fejér convergent to S, if and only if for all $x \in S$, there exists $\bar{k}$ such that $\|x^{k+1} - x\| \leq \|x^k - x\|$ for all $k \geq \bar{k}$.*

This definition was introduced in [10] and has been further elaborated in [30]. A useful result on Fejér sequences is the following.

Theorem 1. *If $\{x^k\}$ is Fejér convergent to S then,*

  i) The sequence $\{x^k\}$ is bounded,

  ii) if a cluster point of the sequence $\{x^k\}$ belongs to S, then the whole sequence $\{x^k\}$ converges.

Proof. See Theorem 2.16 of [2]. □

3 On convexity

Convexity is a very helpful concept in optimization. Convex functions satisfy nice properties such as existence of directional derivative and subgradient. In this section we will study the fulfillment of these properties in the variable order case.

We begin with the analysis of the epigraph. In the variable order case the epigraph of F is defined as

$$epi(F) = \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m : F(x) \in y - K(F(x))\}.$$

In non-variable orders, the convexity of $epi(F)$ is equivalent to the convexity of F; see [33]. However, as it is shown in the next proposition, in the variable order setting, this important characterization does not hold.

Proposition 2. *Suppose that F is a K-convex function. Then, $epi(F)$ is convex if and only if $K(F(x)) = K$, for all $x \in \mathbb{R}^n$.*

Proof. Suppose that for some $x, \hat{x} \in \mathbb{R}^n$ such that $F(x) \neq F(\hat{x})$, there exists $z \in K(F(x)) \setminus K(F(\hat{x}))$. Take the points $(x, F(x) + 2\alpha z)$ and $(2\hat{x} - x, F(2\hat{x} - x))$, with $\alpha > 0$. They belong to $epi(F)$.
Consider the following convex combination:

$$\frac{(x, F(x) + 2\alpha z)}{2} + \frac{(2\hat{x} - x, F(2\hat{x} - x))}{2} = \left(\hat{x}, \frac{F(x) + F(2\hat{x} - x)}{2} + \alpha z \right).$$

This point belongs to $epi(F)$ if and only if

$$F(\hat{x}) = \frac{F(x) + F(2\hat{x} - x)}{2} + \alpha z - k(\alpha),$$

where $k(\alpha) \in K(F(\hat{x}))$. By the K-convexity of F,

$$F(\hat{x}) = \frac{F(x) + F(2\hat{x} - x)}{2} - k_1,$$

where $k_1 \in K(F(\hat{x}))$. So,

$$\alpha z + k_1 = k(\alpha). \tag{5}$$

Since $K(F(\hat{x}))$ is closed and convex, and $z \notin K(F(\hat{x}))$, $\{z\}$ and $K(F(\hat{x}))$ may be strictly separated in $\mathbb{R}^m$ by a hyperplane, *i.e.* there exists some $p \in \mathbb{R}^m \setminus \{0\}$ such that

$$p^T k \geq 0 > p^T z, \tag{6}$$

for all $k \in K(F(\hat{x}))$. Therefore, after multiplying (5) by p^T and using (6) with

$$k = k(\alpha) \in K(F(\hat{x})),$$

we obtain that

$$\alpha p^T z + p^T k_1 = p^T k(\alpha) \geq 0.$$

Taking limits as α goes to ∞, the contradiction is established, since

$$0 \leq \alpha p^T z + p^T k_1 \to -\infty.$$

Hence, $K(F(x)) \equiv K$ for all $x \in \mathbb{R}^n$. □

In the following we present some analytical properties of K-convex functions. For the non-differentiable model, we generalize the classical assumptions given in the case of constant cones; see [13, 33]. Let us first present the definition of Daniell cone, for more details; see [35].

Definition 2. *Given a cone $\mathcal{K}$ we say that $\{y^k\}$ is a $\mathcal{K}$-non increasing sequence if $y^k - y^{k+1} \in \mathcal{K}$ for $k = 1, 2, \ldots$.*
We say that $\inf_k\{y^k\} = y^*$, if $y^k - y^* \in \mathcal{K}$, for all $k \in \mathbb{N}$ and there is not $y \in y^* + \mathcal{K} \setminus \{0\}$ such that $y^k - y \in \mathcal{K}$.
A cone $\mathcal{K}$ is Daniell, if for all non-increasing sequence $\{y^k\}$ such that for some $\hat{y}$, $\{y^k - \hat{y}\} \subset \mathcal{K}$, then $\lim_{k \to \infty} y^k = \inf_k\{y^k\}$.

In our framework, it is well known every pointed, closed and convex cone is a Daniell cone; see [31].

Lemma 1. *Suppose that there exists $\mathcal{K}$ a Daniell cone such that $K(F(x)) \subset \mathcal{K}$ for all x in a neighborhood of x^*. If F is a K-convex function, then F is locally Lipschitz around x^*.*

Proof. If F is K convex, then F is $\mathcal{K}$-convex in the non-variable sense. By Theorem 3.1 of [34], F is locally Lipschitz. □

Proposition 3. *Suppose that for each $\bar{x}$ there exists $\varepsilon > 0$ such that $\cup_{x \in B[\bar{x}, \varepsilon]} K(F(x)) \subset \mathcal{K}$, where $\mathcal{K}$ is a Daniell cone. Then, the directional derivative of F at $\bar{x}$ exists along $d = (x - \bar{x})$, i.e.*

$$F'(\bar{x}; x - \bar{x}) = \lim_{t \to 0^+} \frac{F(\bar{x} + td) - F(\bar{x})}{t}.$$

Proof. By the convexity of F,

$$F(\bar{x} + t_1 d) - \frac{t_1}{t_2} F(\bar{x} + t_2 d) - \left(\frac{t_2 - t_1}{t_2}\right) F(\bar{x}) \in -K(F(\bar{x} + t_1 d)),$$

for all $0 < t_1 < t_2 < \varepsilon$. Dividing by t_1, we have that

$$\frac{F(\bar{x} + t_1 d) - F(\bar{x})}{t_1} - \frac{F(\bar{x} + t_2 d) - F(\bar{x})}{t_2} \in -K(F(\bar{x} + t_1 d)) \subset -\mathcal{K}.$$

Hence, $\dfrac{F(\bar{x} + t_1 d) - F(\bar{x})}{t_1}$ is a non-increasing function. Similarly, as

$$F(\bar{x}) - \frac{t_1}{t_1 + 1} F(\bar{x} - d) - \frac{1}{t_1 + 1} F(\bar{x} + t_1 d) \in -K(F(\bar{x})),$$

it holds that

$$\frac{F(\bar{x} + t_1 d) - F(\bar{x})}{t_1} - F(\bar{x} - d) - F(\bar{x}) \in K(F(\bar{x})) \subset \mathcal{K}.$$

4

Since $\mathcal{K}$ is a Daniell cone, $\dfrac{F(\bar{x} + t_1 d) - F(\bar{x})}{t_1}$ has a limit as t_1 goes to 0. Hence, the directional derivative exists. $\qquad\square$

Let us present the definition of subgradient.

Definition 3. *We say that $\epsilon_{\bar{x}}$ is a subgradient of F at $\bar{x}$ if for all $x \in \mathbb{R}^n$,*

$$F(x) - F(\bar{x}) \in \epsilon_{\bar{x}}(x - \bar{x}) + K(F(\bar{x})).$$

The set of all subgradients of F at $\bar{x}$ is denoted as $\partial F(\bar{x})$.

Proposition 4. *If for all $x \in \mathbb{R}^n$, $\partial F(x) \neq \emptyset$, then F is K-convex.*

Proof. As $\partial F(x) \neq \emptyset$, there exists $\epsilon_{\alpha x_1 + (1-\alpha)x_2}$ and $k_1, k_2 \in K(F(\alpha x_1 + (1 - \alpha)x_2))$, such that

$$F(x_2) - F(\alpha x_1 + (1 - \alpha)x_2) = \alpha \epsilon_{\alpha x_1 + (1-\alpha)x_2}(x_2 - x_1) + k_1,$$

and

$$F(x_1) - F(\alpha x_1 + (1 - \alpha)x_2) = (\alpha - 1)\epsilon_{\alpha x_1 + (1-\alpha)x_2}(x_2 - x_1) + k_2.$$

Multiplying the previous equalities by $(1 - \alpha)$ and α respectively, their addition leads to

$$\alpha F(x_1) + (1 - \alpha)F(x_2) - F(\alpha x_1 + (1 - \alpha)x_2) = \alpha k_2 + (1 - \alpha)k_1.$$

Since $K(F(\alpha x_1 + (1 - \alpha)x_2))$ is convex, the result follows. $\qquad\square$

Proposition 5. *If K is a closed application, then ∂F is closed.*

Proof. Assume that $\{x^k\}$ and $\{A^k\} \subset \partial F(x^k)$ are sequences such that $\lim_{k \to \infty} x^k = x^*$ and $\lim_{k \to \infty} A^k = A$. For every x, one has
$$F(x) - F(x^k) - A^k(y - x^k) \in K(F(x^k)).$$

Taking k going to ∞, as $\lim_{k \to \infty} F(x^k) = F(x^*)$ and K is a closed mapping, we get that

$$F(x) - F(x^*) - A(x - x^*) \in K(F(x^*)).$$

Hence, $A \in \partial F(x^*)$, establishing that $\partial F(x^*)$ is closed. $\qquad\square$

Proposition 6. *Let F be a K-convex function. If $gr(K)$ is closed, then for all $\bar{x} \in \mathbb{R}^n$, where F is differentiable, $\nabla F(\bar{x}) = \partial F(\bar{x})$.*

Proof. First we show that $\nabla F(\bar{x})$ belongs to $\partial F(\bar{x})$. Since F is a differentiable function, fixed $\bar{x}$, we get

$$F(\alpha x + (1 - \alpha)\bar{x}) = F(\bar{x}) + \alpha \nabla F(\bar{x})(x - \bar{x}) + o\left(\alpha\|x - \bar{x}\|\right).$$

By K-convexity

$$F(\bar{x}) + \alpha \nabla F(\bar{x})(x - \bar{x}) + o\left(\alpha\|x - \bar{x}\|\right) \in \alpha F(x) + (1 - \alpha)F(\bar{x}) - K(F(\alpha x + (1 - \alpha)\bar{x})).$$

So,

$$\alpha\left(F(x) - F(\bar{x}) - \nabla F(\bar{x})(x - \bar{x}) + \frac{o(\alpha)}{\alpha}\right) \in K(F(\alpha x + (1 - \alpha)\bar{x})).$$

As K is a cone, it follows that

$$F(x) - F(\bar{x}) - \nabla F(\bar{x})(x - \bar{x}) + \frac{o(\alpha)}{\alpha} \in K(F(\alpha x + (1 - \alpha)\bar{x})).$$

By taking limits as α goes to 0 and recalling that F is a continuous function and K is a closed application, by Lemma 1 it holds that

$$F(x) - F(\bar{x}) - \nabla F(\bar{x})(x - \bar{x}) \in K(F(\bar{x})),$$

and hence, $\nabla F(\bar{x}) \in \partial F(\bar{x})$.

Suppose that $\varepsilon \in \partial F(\bar{x})$. Fixed $d \in \mathbb{R}^n$, we get that, for all $\alpha > 0$,

$$F(\bar{x} + \alpha d) - F(\bar{x}) = \alpha \nabla F(\bar{x})d + o(\alpha) \in \alpha \varepsilon d + k(\alpha),$$

where $k(\alpha) \in K(F(\bar{x}))$. Dividing by $\alpha > 0$, and taking limits as α approaches 0, it follows that

$$[\nabla F(\bar{x}) - \varepsilon]d \in K(F(\bar{x})),$$

recall that $K(F(\bar{x}))$ is a closed set. Repeating the same analysis for $-d$, we obtain that

$$-[\nabla F(\bar{x}) - \varepsilon]d \in K(F(\bar{x})).$$

Taking into account that $K(F(\bar{x}))$ is a pointed cone, $[\nabla F(\bar{x}) - \varepsilon]d = 0$. As the previous equality is valid for all $d \in \mathbb{R}^n$,

$$\nabla F(\bar{x}) = \varepsilon,$$

establishing the desire equality. $\qquad\square$

Theorem 2. *Suppose that there exists $\mathcal{K}$ a Daniell cone such that $K(F(x)) \subset \mathcal{K}$ for all x in a neighborhood of x^*. If F is K-convex and K is a closed application, then $\partial F(x^*) \neq \emptyset$.*

Proof. By Lemma 1, F is a locally Lipschitz continuous function. By Rademacher's Theorem, for all x^*, F is differentiable almost everywhere on some neighborhood of x^*. Moreover, due to the boundedness of ∇F whenever exists, there exists a sequence x^k convergent to x^* such that $A = \lim_{k \to \infty} \nabla F(x^k)$. By Proposition 6, it holds that $\nabla F(x^k) = \partial F(x^k)$. By Proposition 5, $A \in \partial F(x^*)$, hence $\partial F(x^*) \neq \emptyset$. $\qquad\square$

Remark 1. *Given x^* and V a bounded neighborhood of x^*, under the assumptions of the previous Lemma, the set $\partial F(x)$ is uniformly bounded in V. Indeed as F is K-convex, locally around x^*, F will be also $\mathcal{K}$-convex. Now, as the domain of F is a finite dimensional space, the fact follows directly after [34, Theorem 4.12(ii)].*

4 The algorithms

In this part we will consider two variants of subgradient method for solving Problem (4). The algorithms generate a sequence of projections on special sets. From now on we assume that the following assumptions hold.

Assumptions

(A1) The subgradients of F are locally bounded.

(A2) F is K-convex.

(A3) K is a closed application.

(A4) For all $x^* \in S^*$ and $x \in C$.

$$K(F(x^*)) \subseteq K(F(x)),$$

We emphasize that Hypothesis (A1) is a typical assumption for proving the convergence of the subgradient-scalar methods in infinite dimension setting; see [4, 1, 36, 29, 3]. As stated in [34], for the scalar and vector framework, this assumption holds trivially in finite-dimensional spaces. A sufficient condition can be found in Remark 1.

The existence of subgradient is guaranteed under (A2)-(A3).

Assumption (A4) implies that there exists a cone $\mathbb{K}$ such that $K(F(x)) \equiv \mathbb{K}$ for all $x \in S^*$. In this case Problem (4) is equivalent to the non-variable inequality problem

$$\text{find } x \in C \text{ such that } F(x) \preceq_{\mathbb{K}} 0.$$

However, as $\mathbb{K}$ is not known, this equivalence is not useful from a practical viewpoint. Next example shows a function and an order structure fulfilling (4).

Remark 2. *Given Problem* (4) *with* $C = \mathbb{R}$, $K(y) = \{(r, \theta), r \geq 0, \theta \in [0, \theta(y)]\}$ *and* $F : \mathbb{R} \to \mathbb{R}^2$, $F(x) = (x^2, x)$. *Here*

$$\theta(y) = \begin{cases} \dfrac{\pi}{2}, & \text{if } y_1 = 0, \\[2ex] \dfrac{3\pi}{4} - \dfrac{\arctan(y_2^2/y_1^2)}{2}, & \text{otherwise.} \end{cases}$$

Evidently

$$\mathbb{R}_+ \times \{0\} \subset K(y) \subset \mathbb{R}_+ \times \mathbb{R}.$$

Moreover, $F(x) \in -K(F(x))$ *if and only if* $x = 0$. *Therefore,* $S^* = \{0\}$ *and due to*

$$\theta(y) \geq \frac{\pi}{2} = K(0),$$

Asumption (A4) *holds.*
As x^2 *is convex*

$$F(\alpha\bar{x} + (1 - \alpha)x) - \alpha F(\bar{x}) - (1 - \alpha)F(x) \in -\mathbb{R}_+ \times \{0\} \subset -K(F(x)).$$

Hence, F *is* K-*convex. Moreover, the continuity of* θ *implies that* K *is a closed application.* □
Now we will present the conceptual algorithm.

Conceptual Algorithm

Initialization step. Take $x^0 \in C$, and set $k = 0$.

Iterative step. Given x^k, $U^k \in \partial F(x^k)$. Compute

$$x^{k+1} := \mathcal{F}(x^k, U^k). \tag{7}$$

If $x^{k+1} = x^k$ then stop.

We consider two variants of it. As they are based on the algorithms proposed in [37, 5], the extensions are called Algorithms R and S respectively. The main difference is given by the definition of the function $\mathcal{F}$, which is defined as follows

$$\mathcal{F}_R(x^k, U^k) = P_{C \cap H(x^k, U^k)}(x^k)); \tag{8}$$

$$\mathcal{F}_S(x^k, U^k) = P_{C \cap W(x^k) \cap H(x^k, U^k)}(x^0); \tag{9}$$

where
$$H(x, U) = \{z \in \mathbb{R}^n : F(x) + \langle U, z - x \rangle \in -K(F(x))\} \tag{10}$$

and
$$W(x) = \left\{z \in \mathbb{R}^n \ : \ \langle z - x, x^0 - x \rangle \leq 0\right\}. \tag{11}$$

Before we start with the formal analysis of the convergence properties of the algorithm, we make the following remark on the complexity of the projection steps.

Remark 3. *As $H(x^k, U^k)$ and $W(x^k)$ are halfspaces, the projections defined in (8) and (9) does not entail any significant additional computational cost over the computation of the projection onto C itself. Actually, if C is described by nonlinear constrains, the projections onto these smaller sets may be easier than onto the feasible set; see [3].*

5   Convergence Analysis

In this part we will prove the convergence of the algorithms. We will divide this section in three subsections. First we study the properties of the solution set S^*. The convergence of the variants will be discussed separately in the last two parts.

5.1   Properties of the solution set

Proposition 7. *The set S^* is closed and convex.*

Proof. Take $\hat{x}, \bar{x} \in S^*$. Then, it holds that

$$F(\alpha \hat{x} + (1 - \alpha)\bar{x}) \in \alpha F(\hat{x}) + (1 - \alpha)F(\bar{x}) - K(F(\alpha \hat{x} + (1 - \alpha)\bar{x})),$$

for all $\alpha \in [0, 1]$. As $F(\hat{x}) \preceq_{K(F(\hat{x}))} 0$ and $F(\bar{x}) \preceq_{K(F(\bar{x}))} 0$, it follows from (A4) that

$$K(F(\hat{x})) = K(F(\bar{x})) \subset K(F(\alpha \hat{x} + (1 - \alpha)\bar{x})).$$

Hence,
$$F(\alpha \hat{x} + (1 - \alpha)\bar{x}) \in -K(F(\alpha \hat{x} + (1 - \alpha)\bar{x})),$$

and therefore $\alpha \hat{x} + (1 - \alpha)\bar{x} \in S^*$.

For the closeness, consider any sequence $\{\bar{x}^k\} \subset S^*$ convergent to $\hat{x}$. As F is a continuous function; see Lemma 1, $\lim_{k \to \infty} F(\bar{x}^k) = F(\hat{x})$ and taking into account that $F(\bar{x}^k) \in -K(F(\bar{x}^k))$ and the closeness of K leads to $F(\hat{x}) \in -K(F(\hat{x}))$. So, $\hat{x} \in S^*$.   □

We assume that S^* is a nonempty set.

Lemma 2. *For all $x \notin S^*$ and $U \in \partial F(x)$, it holds that $S^* \subset H(x, U)$.*

Proof. Take $\hat{x} \in S^*$. Then, $F(\hat{x}) \in -K(F(z))$ and by the K convexity of F,

$$F(x) + \langle U, \hat{x} - x \rangle - F(\hat{x}) \in -K(F(x)),$$

for all $x \in \mathbb{R}^n$ and all $U \in \partial F(x)$. Hence, using the above inclusion and (4), we get that

$$F(x) + \langle U, \hat{x} - x \rangle \in -K(F(x)) - K(F(\hat{x})) \subseteq -K(F(x)),$$

for all $x \notin S^*$. So, $\hat{x} \in H(x, U)$.   □

Lemma 3. *If $x \in H(x, U) \cap C$ for some $U \in \partial F(x)$, then $x \in S^*$.*

Proof. Suppose that $x \in H(x, U) \cap C$ for some $U \in \partial F(x)$, then $x \in C$ and

$$F(x) \in -K(F(x)),$$

i.e. $x \in S^*$. □

The above lemma shows that the stop criterion of the conceptual algorithm is well defined. The following proposition gives the validity of the stopping criterion.

Proposition 8. *If Algorithm R, respectively Algorithm S, stops at iteration k, then, $x^k \in S^*$.*

Proof. For Algorithm S if $x^{k+1} = x^k$, it follows from (8) that $x^k \in H^k(x^k, U^k) \cap C$. In the case of Algorithm R, $x^{k+1} = x^k$ and (9), lead to $x^k \in W^k(x^k) \cap H^k(x^k, U^k) \cap C \subset H^k(x^k, U^k) \cap C$. So, in both cases, by Lemma 3, $x^k \in S^*$. □

5.2 Convergence of Algorithm R

Proposition 9. *The sequence generated by Algorithm R is Féjer convergent to S^*. Moreover, it is bounded and*

$$\lim_{k \to \infty} \|x^{k+1} - x^k\| = 0. \tag{12}$$

Proof. Take $\hat{x} \in S^*$. By Lemma 2, $\hat{x} \in H(x^k, U^k)$, for all $k \in \mathbb{N}$. Then

$$\|x^{k+1} - \hat{x}\|^2 - \|x^k - \hat{x}\|^2 + \|x^{k+1} - x^k\|^2 = 2\langle \hat{x} - x^{k+1}, x^k - x^{k+1} \rangle \leq 0,$$

using Proposition 1(i) and (8) in the last inequality. So,

$$\|x^{k+1} - \hat{x}\|^2 \leq \|x^k - \hat{x}\|^2 - \|x^{k+1} - x^k\|^2. \tag{13}$$

The above inequality implies that $\{x^k\}$ is Féjer convergent to S^* and hence $\{x^k\}$ is bounded. Also, we get

$$0 \leq \|x^{k+1} - \hat{x}\|^2 \leq \|x^k - \hat{x}\|^2.$$

So, $\{\|x^k - \hat{x}\|^2\}$ is a convergent sequence. Therefore, using (13), we get that

$$\lim_{k \to \infty} \|x^{k+1} - x^k\| = 0.$$

□

Theorem 3. *The sequence generated by Algorithm R converges to some point in S^*.*

Proof. By Proposition 9, $\{x^k\}$ is bounded. So, using (A1), $\{U^k\}$ is bounded, *i.e.* there exists $L \geq 0$ such that

$$\|U^k\| \leq L, \tag{14}$$

for all k.

Fix $k \in \mathbb{N}$. As $K(F(x^k))$ is a closed convex cone, it is clear that $y \in \mathbb{R}^m$ can be uniquely written as

$$y = y_+ + y_-,$$

with $y_+ \in K^*(F(x^k))$, $y_- \in -K(F(x))$ and $\langle y_+, y_- \rangle = 0$. For $y = F(x^k)$, consider $F(x^k)_+$ and $F(x^k)_-$. Now

$$\|F(x^k)_+\|^2 = \langle F(x^k)_+, F(x^k)_+ + F(x^k)_- \rangle = \langle F(x^k)_+, F(x^k) \rangle \tag{15}$$
$$= \langle F(x^k)_+, F(x^k) + U^k(x^{k+1} - x^k) \rangle - \langle F(x^k)_+, U^k(x^{k+1} - x^k) \rangle. \tag{16}$$

But $F(x^k)_+ \in K^*(F(x^k))$, so $\langle F(x^k)_+, F(x^k) + U^k(x^{k+1} - x^k)\rangle \leq 0$, and therefore

$$\|F(x^k)_+\|^2 \leq -\langle F(x^k)_+, U^k(x^{k+1} - x^k)\rangle.$$

Applying the Cauchy Schwartz inequality and recalling (14), it follows that

$$\|F(x^k)_+\|^2 \leq L\|F(x^k)_+\|\|x^{k+1} - x^k\|.$$

As $x^k \notin S^*$, $F(x^k)_+ \neq 0$. So, dividing by $\|F(x^k)_+\|$, we obtain

$$\|F(x^k)_+\| \leq L\|x^{k+1} - x^k\|. \tag{17}$$

Recalling (12), it follows that

$$\lim_{k \to \infty} \|F(x^k)_+\| = 0. \tag{18}$$

Now consider a convergent subsequence $\{x^{j_k}\}$ of $\{x^k\}$. Denote $\hat{x}$ as its limit. It follows from (18) that $F(\hat{x})_+ = 0$. Henceforth, $F(\hat{x}) = F(\hat{x})_-$. Moreover as

$$\lim_{k \to \infty} F(x^{j_k})_- = \lim_{k \to \infty} F(x^{j_k}) - \lim_{k \to \infty} F(x^{j_k})_+,$$

we get that

$$\lim_{k \to \infty} F(x^{j_k})_- = F(x^*).$$

As $F(x^{j_k})_- \in -K(F(x^{j_k}))$ and (A3) is fulfilled,

$$F(x^*) \in -K(F(x^*)),$$

i.e. $x^* \in S^*$. Therefore the accumulation points of $\{x^k\}$ belong to S^*. Finally, by the Féjer convergence, the sequence converge to a point in S^*. □

5.3 Convergence of Algorithm S

Observe that, in virtue of their definitions, $W_k := W(x^k)$ and $H_k := H(x^k, U^k)$ for some $U^k \in \partial F(x^k)$ are convex and closed sets, for each k. Therefore $C \cap H_k \cap W_k$ is a convex and closed set, for each $k = 0, 1, \ldots,$. So, if $C \cap H_k \cap W_k$ is nonempty then, the next iterate, x^{k+1}, is well-defined. Next lemma guarantees this fact.

Lemma 4. *For all $k = 0, 1, \ldots,$, it holds that $S^* \subset C \cap H_k \cap W_k$.*

Proof. We proceed by induction. By definition, $S^* \subset C$. By Lemma 2, $S^* \subset C \cap H_k$, for all k. For $k = 0$, as $W_0 = \mathbb{R}^m$, $S^* \subset C \cap H_0 \cap W_0$.

Assume that $S^* \subset C \cap H_\ell \cap W_\ell$, for $\ell \leq k$. Henceforth, $x^{k+1} = P_{C \cap H_k \cap W_k}(x^0)$ is well defined. Then, by Lemma 2, for all $x^* \in S^*$, we get that

$$\langle x^* - x^{k+1}, \, x^0 - x^{k+1}\rangle = \langle x^* - P_{C \cap H_k \cap W_k}(x^0), \, x^0 - P_{C \cap H_k \cap W_k}(x^0)\rangle \leq 0, \tag{19}$$

using the induction hypothesis. The inequality (19) implies that $x^* \in W_{k+1}$ and hence, S^* is a subset of $C \cap H_{k+1} \cap W_{k+1}$. □

Corollary 1. *Algorithm S is well-defined.*

Proof. By the previous lemma, $S^* \subset C \cap H_k \cap W_k$, for $k = 0, 1, \ldots,$. As $S^* \neq \emptyset$, then, given x^0, the sequence $\{x^k\}$ is computable. □

Before proving the convergence of the sequence, we will study its boundness. Next Lemma shows that the sequence remains in a ball determined by the initial point.

Lemma 5. *The sequence $\{x^k\}$ is bounded. Furthermore,*

$$\{x^k\} \subset B\left[\frac{1}{2}(x^0 + x^*), \frac{1}{2}\rho\right], \quad k = 0, 1, \ldots, \tag{20}$$

where $x^* = P_{S^*}(x^0)$ and $\rho = \text{dist}(x^0, S^*)$.

Proof. Lemma 4 says that $S^* \subset C \cap W_k \cap H_k$ for $k = 0, 1, \ldots$, and by definition of x^{k+1}; see (7), it is true that

$$\|x^{k+1} - x^0\| \leq \|z - x^0\|, \tag{21}$$

for $k = 0, 1, \ldots$, and all $z \in S^*$. Henceforth, taking in (21) $z = x^*$,

$$\|x^{k+1} - x^0\| \leq \|x^* - x^0\| = \rho, \tag{22}$$

for all k. Hence, $\{x^k\}$ is bounded. Without loss of generality, take $z^k = x^k - \frac{1}{2}(x^0 + x^*)$ and $z^* = x^* - \frac{1}{2}(x^0 + x^*)$. It follows from the fact $x^* \in W_{k+1}$ that

$$
\begin{aligned}
0 &\geq 2\langle x^* - x^{k+1}, x^0 - x^{k+1}\rangle \\
&= 2\left\langle z^* + \frac{1}{2}(x^0 + x^*) - z^{k+1} - \frac{1}{2}(x^0 + x^*), z^0 + \frac{1}{2}(x^0 + x^*) - z^{k+1} - \frac{1}{2}(x^0 + x^*)\right\rangle \\
&= 2\langle z^* - z^{k+1}, z^0 - z^{k+1}\rangle = \langle z^* - z^{k+1}, -z^* - z^{k+1}\rangle = \|z^{k+1}\|^2 - \|z^*\|^2,
\end{aligned}
$$

using in the third equality that $z^* = -z^0$. So,

$$\left\|x^{k+1} - \frac{x^0 + x^*}{2}\right\| \leq \left\|x^* - \frac{x^0 + x^*}{2}\right\| = \frac{\rho}{2}. \tag{23}$$

$\square$

Now we will focus on the properties of the accumulation points.

Lemma 6. *All accumulation points of $\{x^k\}$ are elements of S^*.*

Proof. Since $x^{k+1} \in W_k$,

$$0 \geq 2\langle x^{k+1} - x^k, x^0 - x^k\rangle = \|x^{k+1} - x^k\|^2 - \|x^{k+1} - x^0\|^2 + \|x^k - x^0\|^2.$$

Equivalently

$$0 \leq \|x^{k+1} - x^k\|^2 \leq \|x^{k+1} - x^0\|^2 - \|x^k - x^0\|^2,$$

establishing that $\{\|x^k - x^0\|\}$ is a monotone nondecreasing sequence. It follows from Lemma 5 that $\{\|x^k - x^0\|\}$ is bounded and thus, it is a convergent sequence. Therefore,

$$\lim_{k \to \infty} \|x^{k+1} - x^k\| = 0. \tag{24}$$

Let $\bar{x}$ be an accumulation point of $\{x^k\}$ and $\{x^{j_k}\}$ be a convergent subsequence to $\bar{x}$. Since x^{k+1} belongs to H_k, we have that

$$F(x^{j_k}) + U^{j_k}\left(x^{j_k+1} - x^{j_k}\right) \preceq_{K(F(x^k))} 0. \tag{25}$$

By Remark 1, $\{U^{j_k}\}$ is bounded. So, $\{U^{j_k}(x^{j_k+1} - x^{j_k})\}$ converges to zero. By taking limits in (25) and recalling that K is closed application, we obtain that

$$\lim_{k \to \infty} F(x^{j_k}) + U^{j_k}\left(x^{j_k+1} - x^{j_k}\right) = F(\bar{x}) \in -K(F(\bar{x})). \tag{26}$$

$\square$

Finally, we are ready to prove the convergence of the sequence $\{x^k\}$ generated by Algorithm S to the solution which lies closest to x^0.

11

Theorem 4. *Define $x^* = P_{S^*}(x^0)$. Then $\{x^k\}$ converges to x^*.*

Proof. By Lemma 5, $\{x^k\} \subset B\left[\frac{1}{2}(x^0 + x^*), \frac{1}{2}\rho\right]$ is bounded. Let $\{x^{j_k}\}$ be a convergent subsequence of $\{x^k\}$, and let $\hat{x}$ be its limit. Evidently $\hat{x} \in B\left[\frac{1}{2}(x^0 + x^*), \frac{1}{2}\rho\right]$. Furthermore, by Lemma 6, $\hat{x} \in S^*$. As

$$S^* \cap B\left[\frac{1}{2}(x^0 + x^*), \frac{1}{2}\rho\right] = \{x^*\},$$

recall that S^* is a convex and closed set, we conclude that x^* is the unique limit point of $\{x^k\}$. Hence $\{x^{j_k}\}$ converges to $x^* \in S^*$. □

Final remarks

In this paper we have presented two algorithms for finding a solution of the K-convex variable inequalities problem. Using the same hypotheses, their convergence is shown. At Algorithm S the projection step involves more specifications than Algorithm R. However, the sequence generated by the first one, has better properties. In fact it converges to a solution of the problem, which lies closest to the starting point. We emphasize, that this last special feature is interesting and it is useful in specific applications, such as image reconstruction [20, 27, 38]. The main drawback of extending these algorithms to the infinite dimensional spaces is that the existence of the subgradient has not been shown in the variable order case.

As studied in [15], variable order can be considered in two ways,

$$y \preceq_K^1 \bar{y} \text{ if and only if } \bar{y} - y \in K(y)$$

or

$$y \preceq_K^2 \bar{y} \text{ if and only if } \bar{y} - y \in K(\bar{y}).$$

Problem (4) corresponds with the inequality defined by $\preceq_K^1$. If the order is given by $\preceq_K^2$, the inequality problem becomes

$$\text{find } x \in C \text{ such that } F(x) \preceq_{K(0)} 0.$$

As the cone $K(0)$ is fixed, the previous model is a non-variable K-inequality problem and it can be solved by the solution algorithm proposed in [37].

References

[1] Alber, Ya.I., Iusem, A.N., Solodov, M.V.: On the projected subgradient method for nonsmooth convex optimization in a Hilbert space. Math. Program. **81**, 23-37 (1998)

[2] Bauschke, H.H., Borwein, J. M.: On projection algorithms for solving convex feasibility problems. SIAM Rev. **38**, 367-426 (1996)

[3] Bello Cruz, J.Y., Iusem, A.N.: A strongly convergent direct method for monotone variational inequalities in Hilbert spaces. *Numerical Functional Analysis and Optimization* **30** 23-36 (2009)

[4] Bello Cruz, J.Y., Iusem, A.N.: A strongly convergent method for nonsmooth convex minimization in Hilbert spaces. Numer. Funct. Anal. Optim. **32**, 1009 -1018 (2011)

[5] Bello Cruz, J.Y., Lucambio Pérez, L.R.: Convergence of a projected gradient method variant for quasiconvex objectives. Nonlinear Anal. **9**, 2917-2922 (2010)

[6] Bello Cruz, J.Y., Lucambio Pérez, L.R., Melo, J.G. Convergence of the projected gradient method for quasiconvex multiobjective optimization. *Nonlinear Analysis* **74** 5268-5273 (2011)

[7] Bertsekas, D.: Convex optimization theory. Athena Scientific, Nashua (2009)

[8] Bolintinéanu, S.: Approximate efficiency and scalar stationarity in unbounded nonsmooth convex vector optimization problems. J. Optim. Theory Appl. **106**, 265-296 (2000)

[9] Bonnel, H., Iusem, A.N., Svaiter, B.F.: Proximal methods in vector optimization. SIAM J. Optim. **15**, 953-970 (2005)

[10] Browder, F.E.: Convergence theorems for sequences of nonlinear operators in Banach spaces. Math. Z. **100**, 201-225 (1967)

[11] Censor, Y, Herman, G.T.: Block-iterative algorithms with underrelaxed Bregman projections. SIAM J. Optim. **13** 283-297 (2002)

[12] Combettes, P. L., Pesquet, J-C.: Primal-Dual Splitting Algorithm for Solving Inclusions with Mixtures of Composite, Lipschitzian, and Parallel-Sum Type Monotone Operators. Set-Valued Var. Anal. **20**, 307-330 (2012).

[13] Chen, G.Y., Johannes Jahn.: Optimality conditions for set-valued optimization problems. Math. Meth. Oper. Res. **48**, 187-200 (1998)

[14] Eichfelder, G., Duc Ha, T. X.: Optimality conditions for vector optimization problems with variable ordering structures. Optimization, DOI: 10.1080/02331934.2011.575939 (2011)

[15] Eichfelder, G.: Optimal elements in vector optimization with variable ordering structure. J. Optim. Theory Appl. **151**, 217-240 (2011)

[16] Engau, A.: Variable preference modeling with ideal-symmetric convex cones, J. Global Optim. **42**, 295-311 (2008)

[17] Fliege, J., Graña Drummond, L.M., Svaiter, B.F.: Newton's method for multiobjective optimization. SIAM J. Optim. **20**, 602-626 (2009) Math. Program. **111**, 201-216 (2008)

[18] Fliege, J., Svaiter, B.F.: Steepest descent methods for multicriteria optimization. Math. Methods Oper. Res. **51**, 479-494 (2000)

[19] Fukuda, E.H., Graña Drummond, L.M.: On the convergence of the projected gradient method for vector optimization. Optimization **60**, 1009-1021 (2011)

[20] Gordon, R., Herman, G.T.: Reconstruction of pictures from their projections. Comm. ACM **14**, 759-768 (1971)

[21] Graña Drummond, L.M., Iusem, A.N.: A projected gradient method for vector optimization problems. Comput. Optim. Appl. **28**, 5-30 (2004)

[22] Graña Drummond, L.M., Maculan, N., Svaiter, B.F.: On the choice of parameters for the weighting method in vector optimization. Math. Program. **111**, 201-216 (2008)

[23] Graña Drummond, L.M., Svaiter, B.F.: A steepest descent method for vector optimization. J. Comput. Appl. Math. **175**, 395-414 (2005)

[24] Jahn, J.: Vector optimization - Theory, applications and extensions. Springer, Berlin (2004)

[25] Jahn, J.: Mathematical vector optimization in Partially Ordered Linear Spaces. Verlag Peter D. Lang, Frankfurt (1986)

[26] Jahn, J.: Scalarization in vector optimization. Math. Program. **29**, 203-218 (1984)

[27] Hudson, H.M., Larkin, R.S.: Accelerated image reconstruction using ordered subsets of projection data. IEEE Trans. Med. Imag. **13**, 601-609 (1994)

[28] Isac, G., Tammer, C.: Application of a vector-valued Ekeland-type variational principle for deriving optimality conditions. In: Nonlinear Analysis and Variational Problems: In Honor of George Isac, Springer Optimization and Appications **35** 343-365 (2010)

[29] Iusem, A.N., Sosa, W.: Iterative algorithms for equilibrium problems. Optimization **52**, 301-316 (2003)

[30] Iusem, A.N., Svaiter, B.F., Teboulle, M.: Entropy-like proximal methods in convex programming. Math. Oper. Res. **19**, 790-814 (1994)

[31] Luc, D. T.: Pareto Optimality, Game Theory And Equilibria. In Pareto Optimality, Springer Optimization and its Applications **17**, 481-515 (2008)

[32] Luc, D.T.: Scalarization of vector optimization problems. J. Optim. Theory Appl. **55**, 85-102 (1987)

[33] Luc, D.T.: Theory of vector optimization. Lecture Notes in Economics and Mathematical Systems **319**, Springer, Berlin (1989)

[34] Luc, D.T., Tan, N.X., Tinh, P.N.: Convex vector functions and their subdifferential. Acta Math. Vietnam. **23**, 107-127 (1998)

[35] Peressini, A.L.: Ordered Topological Vector Space. Harper and Row, (1967)

[36] Polyak, B.T.: Minimization of unsmooth functionals. U.S.S.R. Comput. Math. and Math. Phys. **9**, 14-29 (1969)

[37] Robinson, S. M.: A subgradient algorithm for solving K-convex inequalities. Optimization and operations research. Lecture Notes in Economics and Mathematical Systems , Springer, Berlin **117**, 237-245 (1976)

[38] Rockmore, A.J., Macovski, A.: A maximum likelihood approach to transmission image reconstruction from projections. IEEE Trans. Nucl. Sci. **24**, 1929-1935 (1977)

[39] Solodov, M.V., Svaiter, B.F.: Forcing strong convergence of proximal point iterations in a Hilbert space. *Mathematical Programming* **87** 189-202 (2000)

[40] Von Neumann, J.: Functional Operators, vol. 2: The Geometry of Orthogonal Spaces. Princeton University Press, Princeton (1950)

[41] Verma, Ram U.: Hybrid inexact proximal point algorithms based on RMM frameworks with applications to variational inclusion problems. J. Appl. Math. Comput. **39**, 345-365 (2012)