ACTA MATHEMATICA VIETNAMICA
Volume 31, Number 1, 2006, pp. 77-98

SOME ALGORITHMS FOR SOLVING MIXED VARIATIONAL INEQUALITIES

TRUONG QUANG BAO AND PHAN QUOC KHANH

Abstract. We propose some algorithms for solving mixed variational inequalities. The global convergence of these algorithms is established under pseudomonotonicity and continuity assumptions without Lipschitz conditions. The implementation is illustrated on a nonsmooth optimization test problem.

1. Introduction

We consider the mixed variational inequality problem, in short (MVI), of finding $x^* \in \mathbb{R}^n$ such that

$$(1) \qquad \langle F(x^*), x - x^* \rangle + \varphi(x) - \varphi(x^*) \geqslant 0, \quad \forall x \in \mathbb{R}^n,$$

where $F : \mathbb{R}^n \to \mathbb{R}^n$ is a single-valued mapping and $\varphi : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is a proper, lower semicontinuous and convex functional. (MVI) is also called a general variational inequality of second type since if φ is the indicator function of a closed convex subset K in $\mathbb{R}^n$, that is,

$$\varphi(x) = I_K(x) = \begin{cases} 0 & \text{if } x \in K, \\ +\infty & \text{otherwise,} \end{cases}$$

(MVI) is equivalent to the usual variational inequality problem (VI): finding $x^* \in K$ such that

$$(2) \qquad \langle F(x^*), x - x^* \rangle \geqslant 0, \quad \forall x \in K,$$

and if F is identically zero, (MVI) becomes a minimization problem.

Observe that the mixed variational inequality problem is equivalent to the problem of finding a zero of the sum of two operators (ZP): finding $x^* \in \mathbb{R}^n$ such that

$$(3) \qquad 0 \in F(x^*) + \partial\varphi(x^*),$$

where $\partial\varphi(.)$ is the subdifferential of the functional φ.

For solving the class of (VI) problem, there are many numerical methods: projection, Weiner-Hoff equation, decomposition and auxiliary principle. Among

Received January 25, 2005; in revised form September 6, 2005.
1991 *Mathematics Subject Classification.* 65K10, 90C25.
Key words and phrases. Mixed variational inequalities, multifunctions, projections, pseudomonotonicity, the bundle method.

these methods, projection algorithms appeared first and are experiencing an explosive development due to their natural arguments, global convergence and simplicity of implementation. The first works were by Goldstein in [9], Levitin et al in [17], and Sibony in [29], where the authors proposed an extension of the projected gradient algorithm for convex minimization problems based on the iteration:

$$x^{k+1} = P_K\big(x^k - \rho F(x^k)\big), \tag{4}$$

where $\rho > 0$ is a parameter and $P_K(.)$ stands for the projection on K. If F is strongly monotone with modulus $\alpha > 0$, i.e.,

$$\langle F(y) - F(x), y - x\rangle \geqslant \alpha\|y - x\|^2 \quad \text{for all } x, y \in \mathbb{R}^n,$$

and Lipschitz continuous with constant L, the classical projection algorithm (4) globally converges to a solution for any $\rho \in \Big(0, \dfrac{2\alpha}{L^2}\Big)$. Moreover, a general scheme for solving (VI) is the auxiliary problem principle which contains the projection method as a special case. In order to guarantee the convergence, this general scheme needs some additional assumptions such as strong monotonicity or cocoercivity, see Cohen [6], Martinet [19], Renaud and Cohen [27] and the recent papers of Anh et al [1], [2] in which the assumptions were weakened and the exact value of the Lipschitz constant is not required to be given.

It is well known that the projection method and its generalizations cannot be used to solve (MVI) due to the presence of the nonlinear term $\varphi(x) - \varphi(x^*)$. But based on the facts that the subdifferential of a proper, lower semicontinuous, convex functional is maximal monotone and that the resolvent of a maximal monotone operator is single-valued, well-defined everywhere and nonexpensive, several methods for (MVI) have been proposed, see e.g. [3], [6], [8], [11], [20]-[22]. Recall that a mapping $T : \mathbb{R}^n \to \mathbb{R}^n$ is said to be monotone if

$$\langle t_x - t_y, x - y\rangle \geqslant 0, \quad \forall x, y \in \mathbb{R}^n, \forall t_x \in T(x), t_y \in T(y),$$

and to be maximal monotone if the graph of any monotone mapping from $\mathbb{R}^n$ into itself cannot properly include the graph of T. The resolvent of operator T, $J_T : \mathbb{R}^n \to \mathbb{R}^n$, is defined by $J_T(x) := (I + \rho T)^{-1}(x)$ for a fixed constant $\rho > 0$ where I is the identity mapping. Resolvents of maximal monotone operators play a crucial role in methods of finding a zero of maximal monotone operator such as the proximal-point algorithms and the operator splitting methods.

The proximal point algorithm for solving $0 \in T(x)$ produces, for any starting point x^0, a sequence $\{x^k\}$ by the iterative scheme

$$x^{k+1} = (I + \rho_k T)^{-1}(x^k).$$

In the splitting methods, the given operator is decomposed into a sum of two operators, whose resolvents are easier to calculate and can be dealt with independently. These methods have been studied by many authors, see e.g. [5]-[8], [11], [15], [16]. In [15], a splitting method was developed by Korpelevich in the spirit of the extragradient method to solve (MVI).

In case of $T = A + B$, where A and B are maximal monotone mappings on $\mathbb{R}^n$ with A being single valued on $\operatorname{dom} A \supset \operatorname{dom} B$, the forward-backward splitting method

$$x^{k+1} = (I + \alpha_k B)^{-1}(I - \alpha_k A)(x^k),$$

where $\alpha_k > 0$, was proposed by Lions and Mercier in [18]. In the case where $B = N_K$, the normal cone of a nonempty closed convex subset K in $\mathbb{R}^n$, this method reduces to a projection method proposed by Sibony in [29] for monotone variational inequalities. In another case, where A is the gradient of a differentiable convex function, it becomes a gradient projection method of Goldstein in [9], and of Levitin and Polyak in [17].

A nice feature of forward-backward method is that the backward (i.e. proximal) step involves B only. However, the forward-backward method has a drawback: it requires A to be Lipschitz continuous on dom$A$ (see [27], [28], [33]). Furthermore, choosing a good step size may be difficult since it entails estimating the Lipschitz constant of A. The Extragradient method of Korperlevich for nonsmooth variational inequalities in [13] modifies the projection method of Sibony by performing an additional forward step at each iteration. By adaptively choosing the stepsize, the method has been shown to converge for monotone continuous mappings in [12], [31].

Choosing α_k requires some care, for it cannot be too large nor can it be too small. If A is Lipschitz continuous, α_k can be chosen to be constant for all iterations. However, it is more practical to choose α_k dynamically using an Armijo-Goldstein stepsize. Specifically, we will choose α_k to be the largest $\alpha \in \{\alpha, \alpha\beta, \alpha\beta^2, \dots\}$ satisfying

$$\alpha \big\| A(J(x^k, \alpha)) - A(x^k) \big\| \leqslant \theta \big\| J(x^k, \alpha) - x^k \big\|,$$

where $\beta \in (0, 1]$ and $\theta \in (0, 1)$ are constants, and $J(x, \alpha) = (I + \alpha B)^{-1}(I - \alpha A)(x)$. Observe that this type of linesearch leads to computing many $J(x^k, \alpha)$ and costs expensively. In [30], [31], [33], [34], another type of linesearch is used: finding a point $y^k \in [x^k, J(x^k, \alpha)]$ such that a Lipschitz condition is fulfilled.

The paper is organized as follows. The remaining part of this Section contains some preliminaries. In Sections 2-4, we present the proposed algorithms and some modified versions, then establish their global convergence. Finally, Section 5 provides some numerical examples.

In the sequel we need the following well known and basic facts (see, e.g., [21]-[23]).

Lemma 1.1. *Let* $\varphi : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}, z \in \mathbb{R}^n$ *and* $\rho > 0$. *Then*

$$x = (I + \rho\partial\varphi)^{-1}(z)$$

if and only if, for any $y \in \mathbb{R}^n$,

$$\langle x - z, y - x \rangle + \rho\varphi(y) - \rho\varphi(x) \geqslant 0. \tag{5}$$

In particular, if φ is the indicator function of a closed convex subset K in $\mathbb{R}^n$, then $(I + \rho\partial\varphi)^{-1} \equiv P_K$, the projection onto K, and Lemma 1.1 reduces to the well known characterization of the projections: $x = P_K(z)$ if and only if $\langle x - z, y - x\rangle \geqslant 0$, $\forall y \in K$.

Lemma 1.2. *$x^* \in \mathbb{R}^n$ is a solution of (MVI) if and only if x^* satisfies the following fixed point formulation*

$$x^* = (I + \rho\partial\varphi)^{-1}(I - \rho F)(x^*). \tag{6}$$

Note that if φ is the indicator function of a closed convex subset K in $\mathbb{R}^n$, Lemma 1.2 reduces to the well known characterization of the solutions to (VI): x^* is a solution to (VI) if and only if $x^* = P_K\big(x^* - \rho F(x^*)\big)$ for any $\rho > 0$.

2. The first algorithm

In the sequel we will use the notations

$$\begin{aligned} \bar{x}(\rho) &:= (I + \rho\partial\varphi)^{-1}(I - \rho F)x := P_\rho(x), \\ r(x,\rho) &:= x - \bar{x}(\rho), \\ \triangle F(x,\rho) &:= F(x) - F(\bar{x}(\rho)). \end{aligned}$$

By Lemma 1.2, x^* is a solution of (MVI) if and only if $r(x^*, \rho) = 0$. This is used as the stopping criteria for our algorithms.

Algorithm 2.1. We require two exogenous parameters $\rho > 0$, $L > 0$ such that $\rho L < 1$.

1. **Initialization**. $x^0 \in \operatorname{dom}\varphi$.

2. **Iteration**. If $r(x_k, \rho_k) = 0$, then stop. Otherwise go to

Linesearch. Choose the smallest nonnegative integer m_k satisfying

$$\big\|\triangle F(x^k, 2^{-m_k}\rho)\big\| \leqslant 2^{m_k} L \big\|r(x^k, 2^{-m_k}\rho)\big\|. \tag{7}$$

Set $\rho_k = 2^{-m_k}\rho$.
Compute

$$\gamma_k := \frac{\big\|r(x^k,\rho_k)\big\|^2 - \rho_k\big\langle\triangle F(x^k,\rho_k), r(x^k,\rho_k)\big\rangle}{\big\|\rho_k\triangle F(x^k,\rho_k) - r(x^k,\rho_k)\big\|^2}.$$

Update

$$x^{k+1} := x^k + \gamma_k[\rho_k\triangle F(x^k,\rho_k) - r(x^k,\rho_k)].$$

Geometric interpretation. Let us compare our algorithm to the well-known *hyperplane projection algorithms* and its modified versions. The main idea of the hyperplane projection algorithm for the variational inequality, i.e., the problem (1) with $\varphi(x) = I_K(x)$, is as follows. Suppose x^k is the current approximation to the solution of (VI). First, we compute the point $P_K[x^k - F(x^k)]$. Next, we search the line segment between x^k and $P_K[x^k - F(x^k)]$ for a point z^k such that the hyperplane $H_k := \{x \in \mathbb{R}^n \mid \langle F(z^k), x - x^k\rangle = 0\}$ strictly separates x^k from any solution x^* of the problem. In Algorithm 2.1 for the mixed variational inequality problem, the projection is replaced by the proximal operator $(I+\rho\partial\varphi)^{-1}$ which becomes the projection in the case $\varphi(x)=I_K(x)$. Moreover, since we do not have nice properties of projections, we use the following set, instead of H_k,

$$\{x\in\mathbb{R}^n \mid \langle\rho\triangle F(x,\rho)-r(x,\rho),x-x^k\rangle=0\}$$

so that the solution set is "sufficient" close to $\bar{x}^k(\rho)=(I+\rho\partial\varphi)^{-1}(I-\rho F)x^k$ (see Proposition 2.2) and the next iteration is "better" than the current one. In our Algorithm, the exact stepsize is chosen as in the proof of Proposition 2.3. Note that x^{k+1} can be infeasible, i.e., Algorithm 2.1 is an infeasible one. So it seems to be incomparable to the Extragradient Method. However, Algorithm 3.1 for (MVI) collapses to Proximal - Extragradient Method for (MVI) or Extragradient Method for (VI) (see Remark 3.4). Also note that the proposed method differs from the one using the auxiliary problem principle, see [1], [2], [6], [27] because we use the projection to a cutting plane separating the solution set and the current iterative as in [30], [31], [34] to obtain an approximate solution instead of using a regularization parameter to get the unique solution of each subproblem.

Remark 2.1. If F is locally Lipschitz with constant l and $L=l$, then $m_k=0$ for all k.

Given $\varphi:\mathbb{R}^n\to\mathbb{R}\cup\{+\infty\}$, recall that a mapping $F:\mathbb{R}^n\to\mathbb{R}^n$ is said to be φ-pseudomonotone at x if, $\forall y\in\mathbb{R}^n$,

$$\big[\langle F(x),y-x\rangle+\varphi(y)-\varphi(x)\geqslant 0\big]\Rightarrow\big[\langle F(y),y-x\rangle+\varphi(y)-\varphi(x)\geqslant 0\big].$$

Hypothesis A. *There exists $l>0$ such that for any convergent sequence x^k and any $\rho_k\to 0$ we have subsequence k_i such that*

$$\big\|F(x^{k_i})-F(P_{\rho_{k_i}}(x^{k_i}))\big\|\leqslant l\big\|x^{k_i}-P_{\rho_{k_i}}(x^{k_i})\big\| \tag{8}$$

for all i large enough.

To establish a global convergence of Algorithm 2.1 we need the following three propositions.

Proposition 2.1. *Assume that F is continuous and (MVI) has a solution $\hat{x}$. Assume further that $x^k\to x$ and $\rho_k\to\rho,\rho_k>0$. Then, $\bar{x}^k(\rho_k)\equiv P_{\rho_k}(x^k)$ contains a subsequence converging to $\bar{x}(\rho)$.*

Proof. We show first that $\{\bar{x}^k(\rho_k)\}$ is bounded. We have

$$\begin{aligned}\big\|\bar{x}^k(\rho_k)-\hat{x}\big\|&=\big\|(I+\rho_k\partial\varphi)^{-1}(I-\rho_kF)x^k-(I+\rho_k\partial\varphi)^{-1}(I-\rho_kF)\hat{x}\big\|\\&\leqslant\big\|(I-\rho_kF)x^k-(I-\rho_kF)\hat{x}\big\|\\&\leqslant\big\|x^k-\hat{x}\big\|+\rho_k\big\|F(x^k)-F(\hat{x})\big\|.\end{aligned}$$

Since F is continuous, $\{\bar{x}^k(\rho_k)\}$ is bounded. Therefore, there is a subsequence, denoted again by $\{\bar{x}^k(\rho_k)\}$, converging to a point y. It remains to show that $y=\bar{x}(\rho)$. We have

$$x^k-\rho F(x^k)\in(I+\rho_k\partial\varphi)\big(\bar{x}^k(\rho_k)\big),$$

or

$$x^k - \bar{x}^k(\rho_k) - \rho_k F(x^k) \in \rho_k \partial\varphi\big(\bar{x}^k(\rho_k)\big).$$

By the definition of a lower limit of a sequence of sets, this implies that

$$x - y - \rho F(x) \in \liminf_{k\to\infty} \rho_k \partial\varphi\big(\bar{x}^k(\rho_k)\big) \subset \rho\overline{\partial\varphi(y)} = \rho\partial\varphi(y).$$

Therefore, $(I - \rho F)x \in (I + \rho\partial\varphi)y$, i.e., $y = \bar{x}(\rho)$. □

Remark 2.2. (i) If F is locally Lipschitz, i.e., $\exists l > 0$, $\forall x \in \mathbb{R}^n$, $\exists U$ (a neighborhood of x), $\forall u, v \in U$,

$$\|F(u) - F(v)\| \leqslant l\|u - v\|,$$

then F satisfies Hypothesis A. Indeed, if $x^k \to x$ and $\rho_k \to 0$, by Proposition 2.1, $\{P_{\rho_k}(x^k)\}$ has a convergent subsequence $P_{\rho_{k_i}}(x^{k_i}) \to P_0(\bar{x}) = \bar{x}$. So for all i large enough we have (8). Moreover, even for the special case where φ is the indicator function of a closed convex subset K, Hypothesis A is weaker than the locally Lipschitz property. Thus, Algorithm 2.1 is a new one for non-Lipschitz variational inequalities.

(ii) (Algorithm 2.1 is well defined) Assume that F satisfies Hypothesis A. Then, for any $x \in \mathbb{R}^n$, there exists a finite positive m such that

$$\big\|\triangle F(x, 2^{-m}\rho)\big\| \leqslant 2^m L\big\|r(x, 2^{-m}\rho)\big\|.$$

Proposition 2.2. *Assume that F is φ-pseudomonotone and $\hat{x}$ is a solution of (MVI). Then, for each $x \in \mathbb{R}^n$ and $\rho > 0$ we have*

$$\big\langle \rho\triangle F(x,\rho) - r(x,\rho), x - \hat{x} - r(x,\rho)\big\rangle \leqslant 0.$$

Proof. By the φ-pseudomonotonicity one has

$$\langle F(\bar{x}(\rho)), \bar{x}(\rho) - \hat{x}\rangle + \varphi(\bar{x}(\rho)) - \varphi(\hat{x}) \geqslant 0. \tag{9}$$

Lemma 1.1 and the definition of $\bar{x}(\rho)$ imply that

$$\langle \bar{x}(\rho) - (x - \rho F(x)), \hat{x} - \bar{x}(\rho)\rangle + \rho\varphi(\hat{x}) - \rho\varphi(\bar{x}(\rho)) \geqslant 0. \tag{10}$$

Multiplying (9) by ρ and then adding it to (10) one gets

$$\big\langle x - \bar{x}(\rho) - \rho\big(F(x) - F(\bar{x}(\rho))\big), \bar{x}(\rho) - \hat{x}\big\rangle \geqslant 0,$$

that is,

$$\langle \rho\triangle F(x,\rho) - r(x,\rho), x - \hat{x} - r(x,\rho)\rangle \leqslant 0.$$

□

Proposition 2.3. *Assume that F is φ-pseudomonotone and $\hat{x}$ is a solution of (MVI). Then, for any sequence $\{x^k\}$ generated by Algorithm 2.1, we have*

$$\big\|x^{k+1} - \hat{x}\big\|^2 \leqslant \big\|x^k - \hat{x}\big\|^2 - \Big(\frac{1-\rho L}{1+\rho L}\Big)^2\big\|r(x^k,\rho_k)\big\|^2. \tag{11}$$

Proof. Observe first, from the definition of m_k, that $\gamma_k \geqslant \dfrac{1-\rho L}{(1+\rho L)^2} > 0$. Next, applying Proposition 2.2 one obtains

$$\begin{aligned}
\|x^{k+1}-\hat{x}\|^2 &= \big\|x^k-\hat{x}+\gamma_k[\rho_k\triangle F(x^k,\rho_k)-r(x^k,\rho_k)]\big\|^2\\
&= \|x^k-\hat{x}\|^2+\gamma_k^2\big\|\rho_k\triangle F(x^k,\rho_k)-r(x^k,\rho_k)\big\|^2\\
&\quad+2\gamma_k\big\langle\rho_k\triangle F(x^k,\rho_k)-r(x^k,\rho_k),x^k-\hat{x}\big\rangle\\
&\leqslant \|x^k-\hat{x}\|^2+\gamma_k^2\big\|\rho_k\triangle F(x^k,\rho_k)-r(x^k,\rho_k)\big\|^2\\
&\quad+2\gamma_k\big\langle\rho_k\triangle F(x^k,\rho_k)-r(x^k,\rho_k),r(x^k,\rho_k)\big\rangle\\
&= \|x^k-\hat{x}\|^2-\frac{\big(\|r(x^k,\rho_k)\|^2-\rho_k\left\langle\triangle F(x^k,\rho_k),r(x^k,\rho_k)\right\rangle\big)^2}{\|\rho_k\triangle F(x^k,\rho_k)-r(x^k,\rho_k)\|^2}\\
&\leqslant \|x^k-\hat{x}\|^2-\Big(\frac{1-\rho L}{1+\rho L}\Big)^2\|r(x^k,\rho_k)\|^2.
\end{aligned}$$

□

Now we are able to prove a global convergence of Algorithm 2.1.

Theorem 2.1. *Assume that F is φ-pseudomonotone, continuous and satisfies Hypothesis A. Assume further that (MVI) has a solution. Then, any sequence $\{x^k\}$ generated by Algorithm 2.1 converges to a solution of (MVI).*

Proof. Adding (11) for $k=0,1,\dots$ we see that $\sum_{k=0}^{\infty}\|r(x^k,\rho_k)\|^2<\infty$ and hence $\lim_{k\to+\infty} r(x^k,\rho_k)=0$. Suppose that there is a subsequence $\rho_{k_i}\to 0^+$, i.e., $m_{k_i}\to+\infty$. By the definition of ρ_k one has

$$\big\|F(x^{k_i})-F(P_{2\rho_{k_i}}(x^{k_i}))\big\|>2^{m_{k_i}-1}L\big\|x^{k_i}-P_{2\rho_{k_i}}(x^{k_i})\big\|. \tag{12}$$

From (11), $\{x^{k_i}\}$ is bounded. Hence, $x^{k_i}\to x^*$ for some x^* (taking subsequence if necessary). Now (12) contradicts Hypothesis A. Therefore, $\{\rho_k\}$ is away from 0. Since $\{\rho_{k_i}\}$ is bounded, there is a convergent subsequence, denoted again by $\{\rho_{k_i}\}, \rho_{k_i}\to\rho_*>0$. By Proposition 2.1,

$$r(x^*,\rho_*)=\lim_{i\to+\infty} r(x^{k_i},\rho_{k_i})=0.$$

Consequenly, x^* is a solution of (MVI) by Lemma 1.2. Since $\|x^{k+1}-x^*\|\leqslant\|x^k-x^*\|$ (by (11)), the whole sequence x^k converges to x^*. □

Example 2.1. To show that Hypothesis A is essential, i.e., the sequence generated by Algorithm 2.1 may disconverge without Hypothesis A, consider the case where $K=\mathbb{R}$, $\varphi=I_K$ and $F(x)=\sqrt{|x|}$. It is obvious that 0 is the unique solution of the problem (VI) and that F does not satisfy Hypothesis A. Let x^k be the sequence produced by Algorithm 2.1 with the linesearch condition

$$\|\triangle F(x^k,\rho_k)\|=\frac{1}{\rho_k}L\|r(x^k,\rho_k)\|$$

instead of choosing the smallest nonnegative integer m_k such that $2^{-m_k}\rho < \rho_k$. It suffices to show that for any $x = a^2$ sufficiently small, $|x_{\cdot+1}|$ is larger than $|x_{\cdot}|$. We have

$$\begin{aligned} x &= a^2, \\ \bar{x} &= a^2 - \rho a, \\ \triangle F(x,\rho) &= \sqrt{a^2} - \sqrt{|a^2 - \rho\sqrt{a^2}|} \\ &= a - \sqrt{\rho a - a^2}, \\ r(x,\rho) &= a^2 - (a^2 - \rho a) = \rho a, \end{aligned}$$

provided that we can choose $\rho > a$, where ρ is the solution of

$$|a - \sqrt{\rho a - a^2}| = \frac{1}{\rho} L\rho a.$$

Furthermore, assume that $\rho > 2a$, i.e., $\sqrt{\rho a - a^2} > a$, we obtain $\rho = a + L + 1$. It is obvious that ρ satisfies all assumptions. Now we have

$$\begin{aligned} \gamma &= \frac{r^2 - \rho\langle \frac{1}{\rho} Lr, r\rangle}{(\rho\frac{1}{\rho}Lr - r)^2} = \frac{1}{1-L}, \\ x_{\cdot+1} &= x + \frac{1}{1-L}\Big(\rho\frac{1}{\rho}Lr - r\Big) \\ &= x - r(x,\rho) \\ &= x - x - \bar{x} \\ &= a^2 - \rho a \\ &= a^2 - (1 + L + a)a \\ &= -(1+L)a. \end{aligned}$$

It is obvious that $|x_{\cdot+1}| > |x_{\cdot}|$. The proof is complete.

The following example shows that Algorithm 2.1 may converge to a solution even in a case where F is non-Lipschitz.

Example 2.2. Consider the problem (1) with $\varphi = I_R$ and F is defined by

$$F(x) = \begin{cases} \dfrac{1}{2k} & \text{if } x = \dfrac{1}{k} \\ \dfrac{1}{2(k+1)} & \text{if } x = \dfrac{1}{k+1} \\ \dfrac{2k+1}{2k(k+1)} & \text{if } x = \dfrac{2k+1}{2k(k+1)} \\ \lambda\dfrac{1}{2k} + (1-\lambda)\dfrac{2k+1}{2k(k+1)} & \text{if } x = \lambda\dfrac{1}{k} + (1-\lambda)\dfrac{2k+1}{2k(k+1)} \\ \lambda\dfrac{1}{2(k+1)} + (1-\lambda)\dfrac{2k+1}{2k(k+1)} & \text{if } x = \lambda\dfrac{1}{k+1} + (1-\lambda)\dfrac{2k+1}{2k(k+1)} \end{cases}$$

for any $k = 1, 2, \cdots$ as in the following figure.

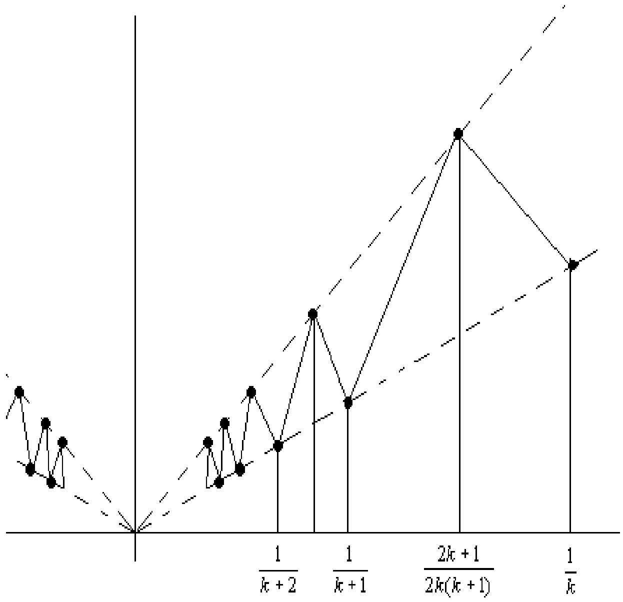

Figure. Graph of $F(s)$

Observe that F is non-Lipschitz (but continuous) everywhere since the linear segments are getting steeper and steeper as they are near zero, which is the solution of the problem. As in the previous example, instead of the checking the linesearch condition (7) we compute the exact ρ_k as the solution of

$$\|\triangle F(x^k, \rho_k)\| = \rho\rho_k^{-1} L \|r(x^k, \rho_k)\|.$$

Hence, $x^{k+1} = \bar{x}^k = x^k - \rho_k F(x^k)$. Since ρ_k is small for k sufficiently large and $0.5x^k < F(x^k) < x^k$, we have $0 < x^{k+1} < x^k$. Therefore the sequence generated by Algorithm 2.1 converges to 0, the unique solution of the problem.

3. The second algorithm

The linesearch in Algorithm 2.1 may lead to performing several proximal points, $J(x_k, \rho)$. To overcome this difficulty, we use the linesearch technique in [33], [34]. This kind of linesearch is to find a point $y^k \in [x^k, J(x^k, \rho)]$ such that some Lipschitz condition is fulfilled. In this section, we write (MVI) in a modified form: finding $x^* \in K$ such that

$$\left\langle F(x^*), x - x^* \right\rangle + \varphi(x) - \varphi(x^*) \geqslant 0, \quad \text{for all } x \text{ in } K.$$

Algorithm 3.1. We require three exogenous positive constants λ, ρ and L such that $\lambda \in (0,1)$ and $\rho L < 1$.

1. **Initialization**. $x^0 \in \operatorname{dom} \varphi$.

2. **Iteration**

STEP 1. Compute

$$\bar{x}^k = (I + \rho\partial\varphi + \rho N_K)^{-1}(I - \rho F)(x^k),$$

or, equivalently, find $\bar{x}^k$ satisfying

$$\langle \bar{x}^k - x^k + \rho F(x^k), z - \bar{x}^k \rangle + \rho\varphi(z) - \rho\varphi(\bar{x}^k) \geqslant 0, \quad \forall z \in K.$$

If $r(x^k) := x^k - \bar{x}^k = 0$, then stop; x^k is a solution of (MVI). Otherwise, go to Step 2.

STEP 2 (Linesearch). Find m_k being the smallest nonnegative integer such that

(13)
$$\langle F(x^k) - F(y^k), r(x^k) \rangle \leqslant L\|x^k - \bar{x}^k\|^2 + \langle s(y^k), x^k - \bar{x}^k \rangle + \varphi(\bar{x}^k) - \varphi(x^k),$$

where $y^k = x^k - \lambda^{m_k} r(x^k)$ and $s(y^k) \in \partial\varphi(y^k)$ is arbitrary.

STEP 3. Compute the projection of x^k onto H_k with

$$H_k := \{ z \in \mathbb{R}^n \ : \langle F(y^k) + s(y^k), z - y^k \rangle = 0 \},$$
$$\tilde{x}^k = P_{H_k}(x^k) = x^k - \gamma_k(F(y^k) + s(y^k)),$$

where

$$\gamma_k = \frac{\langle F(y^k) + s(y^k), x^k - y^k \rangle}{\|F(y^k) + s(y^k)\|^2}.$$

STEP 4. Update the variable

$$x^{k+1} = P_K(\tilde{x}^k) \quad or \quad x^{k+1} = P_{K \cap H_k^-}(\tilde{x}^k),$$

where $H_k^- := \{z \in \mathbb{R}^n \ : \ \langle F(y^k) + s(y^k), z - y^k \rangle \leqslant 0\}$.

Proposition 3.1. *If F is continuous, Algorithm* 3.1 *is well defined.*

Proof. Suppose that for any $y^i = x^k - \lambda^i r(x^k)$, there is $s(y^i) \in \partial\varphi(y^i)$ such that

(14) $$\langle F(x^k) - F(y^i), r(x^k) \rangle > L\|r(x^k)\|^2 + \langle s(y^i), r(x^k) \rangle + \varphi(\bar{x}^k) - \varphi(x^k).$$

Since $y^i \to x^k$, by the maximal monotonicity of $\partial\varphi(.)$, $\{\partial\varphi(y^i)\}$ is bounded. So is $\{s(y^i)\}$. Hence, there is a convergent subsequence $s(y^{i_j}) \to s^* \in \partial\varphi(x^k)$. By the definition of subdifferential, $\varphi(\bar{x}^k) \geqslant \varphi(x^k) + \langle s^*, \bar{x}^k - x^k \rangle$. Now passing (14) to the limit, one obtains

$$0 \geqslant L\|r(x^k)\|^2 + \varphi(\bar{x}^k) - \varphi(x^k) + \langle s^*, x^k - \bar{x}^k \rangle \geqslant L\|r(x^k)\|^2 > 0.$$

This impossibility implies that there exists a finite nonnegative integer satisfying (13) for any $s(y^k) \in \partial\varphi(y^k)$. □

Remark 3.1. In Algorithm 3.1, we need only one $s(y^k) \in \partial\varphi(y^k)$ satisfying (13). So Proposition 3.1 makes the linesearch more easily implementable because we can take arbitrarily $s(y^i) \in \partial\varphi(y^i)$ for checking the condition (13).

Proposition 3.2. *Suppose that $y^k, x^k, \bar{x}^k$ and $s(y^k) \in \partial\varphi(y^k)$ as in Algorithm* 3.1 *satisfy the condition* (13)*. Then, one has (with $t_k = \lambda^{m_k}$)*

$$\langle F(y^k) + s(y^k), x^k - y^k \rangle \geqslant t_k \frac{1 - \rho L}{\rho} \|r(x^k)\|^2.$$

Proof. Since $x^k - y^k = t_k r(x^k)$, by (13) we have the following equivalent inequalities

$$\begin{aligned}
\big\langle F(x^k) - F(y^k), r(x^k)\big\rangle &\leqslant L\|r(x^k)\|^2 + \big\langle s(y^k), r(x^k)\big\rangle \\
&\quad + \varphi(\bar{x}^k) - \varphi(x^k), \\
\Leftrightarrow \rho\big\langle F(y^k), r(x^k)\big\rangle + \rho\big\langle s(y^k), r(x^k)\big\rangle &\geqslant -\rho L\|r(x^k)\|^2 - \rho\varphi(\bar{x}^k) \\
&\quad + \rho\varphi(x^k) + \rho\big\langle F(x^k), r(x^k)\big\rangle, \\
\Leftrightarrow \frac{\rho}{t_k}\big\langle F(y^k) + s(y^k), x^k - y^k\big\rangle &\geqslant -\rho L\|r(x^k)\|^2 - \rho\varphi(\bar{x}^k) \\
&\quad + \rho\varphi(x^k) + \rho\big\langle F(x^k), r(x^k)\big\rangle. \qquad (15)
\end{aligned}$$

By Lemma 1.1, we have

$$\big\langle \bar{x}^k - x^k + \rho F(x^k), x^k - \bar{x}^k\big\rangle + \rho\varphi(x^k) - \rho\varphi(\bar{x}^k) \geqslant 0$$

or

$$\big\langle F(x^k), r(x^k)\big\rangle + \rho\varphi(x^k) - \rho\varphi(\bar{x}^k) \geqslant \|r(x^k)\|^2.$$

Substitute this into (15) we have

$$\big\langle F(y^k) + s(y^k), x^k - y^k\big\rangle \geqslant t_k \frac{1-\rho L}{\rho}\|r(x^k)\|^2.$$

□

Remark 3.2. If F is Lipschitz continuous with the constant l, then (13) holds with $L = l$ and $m_k = 0$, or $y^k = \bar{x}^k$. Indeed,

$$\begin{aligned}
\big\langle F(x^k) - F(\bar{x}^k), r(x^k)\big\rangle &\leqslant L\|r(x^k)\|^2 + \big\langle s(\bar{x}^k), r(x^k)\big\rangle + \varphi(\bar{x}^k) - \varphi(x^k) \\
&\leqslant L\|r(x^k)\|^2,
\end{aligned}$$

the last inequality holds since $s(\bar{x}^k) \in \partial\varphi(\bar{x}^k)$.

Proposition 3.3. *Suppose that x^* is a solution to (MVI) and F is φ-pseudomonotone at that point. The following holds*

$$\big\langle F(y) + s(y), x^* - y\big\rangle \leqslant 0, \quad \forall y \in \mathbb{R}^n,\ \forall s(y) \in \partial\varphi(y).$$

Proof. Since x^* is a solution to (MVI) and F is φ-pseudomonotone at x^*, we have

$$\big\langle F(y), y - x^*\big\rangle + \varphi(y) - \varphi(x^*) \geqslant 0.$$

Since $s(y) \in \partial\varphi(y)$, we have

$$\varphi(x^*) - \varphi(y) + \big\langle s(y), y - x^*\big\rangle \geqslant 0.$$

The required result is obtained by adding the last two inequalities. □

Proposition 3.4. *Assume that $\hat{x}$ is a solution to (MVI) and F is φ-pseudomonotone. Then for any sequence $\{x^k\}$ generated by Algorithm 3.1, one has*

$$\big\|x^{k+1} - \hat{x}\big\|^2 \leqslant \big\|x^k - \hat{x}\big\|^2 - t_k^2 \frac{(1-\rho L)^2}{\rho^2\|F(y^k) + s(y^k)\|^2}\big\|r(x^k)\big\|^4. \qquad (16)$$

Proof. By Proposition 3.2, $\gamma_k > 0$. One has

$$\begin{aligned}
\|\tilde{x}^k - \hat{x}\|^2 &= \|x^k - \hat{x} - \gamma_k(F(y^k) + s(y^k))\|^2 \\
&= \|x^k - \hat{x}\|^2 + \gamma_k^2\|F(y^k) + s(y^k)\|^2 - 2\gamma_k\langle F(y^k) + s(y^k), x^k - \hat{x}\rangle \\
&= \|x^k - \hat{x}\|^2 + \gamma_k^2\|F(y^k) + s(y^k)\|^2 - 2\gamma_k\langle F(y^k) + s(y^k), x^k - y^k\rangle \\
&\quad - 2\gamma_k\langle F(y^k) + s(y^k), y^k - \hat{x}\rangle \\
&\leqslant \|x^k - \hat{x}\|^2 + \gamma_k^2\|F(y^k) + s(y^k)\|^2 - 2\gamma_k\langle F(y^k) + s(y^k), x^k - y^k\rangle \\
&= \|x^k - \hat{x}\|^2 - \frac{\langle F(y^k) + s(y^k), x^k - y^k\rangle^2}{\|F(y^k) + s(y^k)\|^2} \\
&\leqslant \|x^k - \hat{x}\|^2 - t_k^2\frac{(1-\rho L)^2}{\rho^2\|F(y^k) + s(y^k)\|^2}\|r(x^k)\|^4. \qquad (17)
\end{aligned}$$

where the first inequality holds by Proposition 3.3, and the other is due to Proposition 3.2.

Furthermore, $\|x^{k+1} - \hat{x}\| \leqslant \|\tilde{x}^k - \hat{x}\|$ if $x^{k+1} = P_K(\tilde{x}^k)$, since $\hat{x} \in K$. Next, we claim that $\hat{x} \in K \cap H^-$. Indeed, suppose that $\hat{x} \notin K \cap H^-$, i.e.,

$$\langle F(y^k) + s(y^k), \hat{x} - y^k\rangle > 0.$$

Hence,

$$\langle F(y^k), \hat{x} - y^k\rangle + \varphi(\hat{x}) - \varphi(y^k) > 0,$$

and then

$$\langle F(\hat{x}), \hat{x} - y^k\rangle + \varphi(\hat{x}) - \varphi(y^k) > 0,$$

i.e.

$$\langle F(\hat{x}), y^k - \hat{x}\rangle + \varphi(y^k) - \varphi(\hat{x}) < 0.$$

This contradicts the fact that $\hat{x}$ is a solution to (MVI). Hence, we also have $\|x^{k+1} - \hat{x}\| \leqslant \|\tilde{x}^k - \hat{x}\|$ if $x^{k+1} = P_{K\cap H^-}(\tilde{x}^k)$. $\square$

Now we can establish a global convergence of Algorithm 3.1.

Theorem 3.1. *Assume that F is continuous, φ-pseudomonotone and there is a solution $\hat{x}$ to (MVI). Then any sequence $\{x^k\}$ generated by Algorithm* 3.1 *is either finitely terminated or convergent to a solution of (MVI).*

Proof. By Proposition 3.4, the sequence $\{x^k\}$ is bounded, so are $\{\bar{x}^k\}$ and $\{y^k\}$. Hence $\{\|F(y^k) + s(y^k)\|\}$ is bounded above by some M due to the fact that the subdifferential of φ is bounded on any bounded subsets.

Summing (16), we have

$$\sum_{k=0}^{\infty} t_k^2\frac{(1-\rho L)^2}{\rho^2\|F(y^k) + s(y^k)\|^2}\|x^k - \bar{x}^k\|^4 \leqslant \|x^0 - \hat{x}\|^2.$$

Then

$$\sum_{k=0}^{\infty} t_k^2\frac{(1-\rho L)^2}{\rho^2 M}\|x^k - \bar{x}^k\|^4 \leqslant \|x^0 - \hat{x}\|^2,$$

and hence

$$\lim_{k\to\infty} t_k \|x^k - \bar{x}^k\|^2 = 0. \tag{18}$$

Since $\{x^k\}$ is bounded, there is a subsequence $\{x^{k_i}\}$ converging to x^*. Suppose x^* is not a solution of (MVI). Suppose further that $\liminf_{k\to\infty} t_k = 0$. Without loss of generality, we can assume that $\lim_{i\to\infty} t_{k_i} = 0$. Due to the linesearch condition (13), we have

$$\begin{aligned}\Big\langle F(x^{k_i}) - F(x^{k_i} + t_{k_i}\lambda^{-1} r(x^{k_i})), r(x^{k_i})\Big\rangle &> L\|r(x^{k_i})\|^2 \\ + \Big\langle s\big(x^{k_i} + t_{k_i}\lambda^{-1} r(x^{k_i})\big), r(x^{k_i})\Big\rangle &+ \varphi(\bar{x}^{k_i}) - \varphi(x^{k_i}).\end{aligned} \tag{19}$$

Since $x^{k_i} + t_{k_i}\lambda^{-1}x^{k_i} \to x^*$, $\{s(x^{k_i} + t_{k_i}\lambda^{-1}x^{k_i})\}$ has a convergent subsequence, again without loss of generality, we can assume that it converges to $s^* \in \partial\varphi(x^*)$ by the maximal monotonicity of $\partial\varphi$. Passing (19) to the limit, we obtain

$$\begin{aligned}0 &\geqslant L\|r(x^*)\|^2 + \langle s(x^*), x^* - \bar{x}^*\rangle + \varphi(\bar{x}^*) - \varphi(x^*) \\ &\geqslant L\|r(x^*)\|^2.\end{aligned}$$

On the other hand, $0 > \|r(x^*)\|$ since x^* is not a solution. This impossibility shows that $\liminf_{k\to\infty} t_k = t^* > 0$.

Now passing (18) to the limit, we have

$$0 = \lim_{k_i\to+\infty} \|r(x^{k_i})\| = \|r(x^*)\|,$$

i.e., x^* must be a solution to (MVI). Proposition 3.4 with x^* in the place of $\hat{x}$ assures that the whole sequence $\{x^k\}$ converges to x^*. □

Remark 3.3. If $\tilde{x}^k = x^k + \theta_k\gamma_k(F(y^k) + s(y^k))$, where $\theta_k \in (0, 2)$ is called the relaxation parameter, then the estimate in Proposition 3.4 becomes

$$\|x^{k+1} - \hat{x}\|^2 \leqslant \|x^k - \hat{x}\|^2 - \theta_k(2-\theta_k)t_k \frac{1-\rho L}{\rho^2\|F(y^k) + s(y^k)\|^2}\|r(x^k)\|^2.$$

It is clear that Theorem 3.1 still holds whenever the sequence $\{\theta_k\}$ is bounded away from zero, i.e., there is $\bar{\theta} > 0$ such that $\bar{\theta} \leqslant \theta_k$ for all k.

Remark 3.4. If F is Lipschitz continuous, Algorithm 3.1 reduces to the Proximal-Extragradient Method

$$\begin{aligned}y^k &= J_{\partial\varphi}(x^k - \rho F(x^k)), \\ x^{k+1} &= J_{\partial\varphi}(y^k - \rho F(y^k)),\end{aligned}$$

where $J_{\partial\varphi}(\cdot) = (I + \rho\partial\varphi)^{-1}(\cdot)$. If, in addition, $\varphi = I_K$ where K is a closed convex subset, Algorithm 3.1 becomes the Extragradient Method.

4. A modified algorithm

Observe that

$$\left\langle F(\bar{x}^k), x^k - \bar{x}^k \right\rangle + \varphi(\bar{x}^k) - \varphi(x^k) \leqslant 0.$$

Hence, it seems better if we replace the linesearch condition (13) in Algorithm 3.1 by the following condition

$$\left\langle F(x^k) - F(y^k), r(x^k) \right\rangle \leqslant L\|r(x^k)\|^2 + (1-t_k)\left(\left\langle s(y^k), r(x^k) \right\rangle + \varphi(\bar{x}^k) - \varphi(x^k) \right).$$

We modify Algorithm 3.1 as follows.

Algorithm 4.1. We require three exogenous positive constants λ, ρ and L such that $\lambda \in (0,1)$ and $\rho L < 1$.

1. **Initialization**. $x^0 \in \operatorname{dom} \varphi$.

2. **Iteration**

STEP 1. Compute

$$\bar{x}^k = (I + \rho\partial\varphi + \rho N_K)^{-1}(I - \rho F)(x^k).$$

If $r(x^k) = 0$, then stop; x^k is a solution of (MVI). Otherwise, go to Step 2.

STEP 2 (Linesearch). Find m_k being the smallest nonnegative integer such that

(20)

$$\left\langle F(x^k) - F(y^k), r(x^k) \right\rangle \leqslant L\|r(x^k)\|^2 + (1-t_k)\left(\left\langle s(y^k), r(x^k) \right\rangle + \varphi(\bar{x}^k) - \varphi(x^k) \right),$$

where $y^k = x^k - \lambda^{m_k} r(x^k)$, $t_k = \lambda^{m_k}$ and $s(y^k) \in \partial\varphi(y^k)$ arbitrary.

STEP 3. Update the variable

$$\gamma_k = \frac{t_k(1-\rho L)\|r(x^k)\|^2}{\|\rho t_k F(x^k) - \rho F(y^k) - \rho(1-t_k)s(y^k) + t_k r(x^k)\|^2}.$$

$$x^{k+1} = P_K\left[x^k + \gamma_k(\rho t_k F(x^k) - \rho F(y^k) - \rho(1-t_k)s(y^k) + t_k r(x^k))\right].$$

It is obvious that all Propositions 3.1 - 3.4 and Theorem 3.1 still hold. We just show that the estimate in Proposition 3.4 becomes

$$\left\|x^{k+1} - \hat{x}\right\|^2 \leqslant \left\|x^k - \hat{x}\right\|^2 - \frac{t_k^2(1-\rho L)^2\|r(x^k)\|^4}{\|\rho t_k F(x^k) - \rho F(y^k) + \rho(1-t_k)s(y^k) + t_k r(x^k)\|^2}.$$

Proof. Observe that $\gamma_k > 0$. We have

$$\begin{aligned}
&\|x^{k+1} - x^*\|^2 \\
&\leqslant \|x^k - x^* + \gamma_k(\rho t_k F(x^k) - \rho F(y^k) - \rho(1-t_k)s(y^k) + t_k r(x^k))\|^2 \\
&= \|x^k - x^*\|^2 + \gamma_k^2\|\rho t_k F(x^k) - \rho F(y^k) - \rho(1-t_k)s(y^k) + t_k r(x^k)\|^2 \\
&\quad + 2\gamma_k \left\langle \rho t_k F(x^k) - \rho F(y^k) - \rho(1-t_k)s(y^k) + t_k r(x^k), x^k - x^* \right\rangle \\
&= \|x^k - x^*\|^2 + \gamma_k^2\|\rho t_k F(x^k) - \rho F(y^k) - \rho(1-t_k)s(y^k) + t_k r(x^k)\|^2
\end{aligned}$$

$$
\begin{aligned}
&+ 2\gamma_k t_k \left\langle \bar{x}^k - x^k + \rho F(x^k), x^k - x^* \right\rangle \\
&- 2\gamma_k \rho \left\langle F(y^k) + (1-t_k)s(y^k), x^k - x^* \right\rangle \\
= {}& \|x^k - x^*\|^2 + \gamma_k^2 \|\rho t_k F(x^k) - \rho F(y^k) - \rho(1-t_k)s(y^k) + t_k r(x^k)\|^2 \\
&+ 2\gamma_k t_k \left\langle \bar{x}^k - x^k + \rho F(x^k), x^k - \bar{x}^k \right\rangle \\
&- 2\gamma_k \rho \left\langle F(y^k) + (1-t_k)s(y^k), x^k - y^k \right\rangle \\
&+ 2\gamma_k t_k \left\langle \bar{x}^k - x^k + \rho F(x^k), \bar{x}^k - x^* \right\rangle \\
&- 2\gamma_k \rho \left\langle F(y^k) + (1-t_k)s(y^k), y^k - x^* \right\rangle \\
= {}& \|x^k - x^*\|^2 + \gamma_k^2 \|\rho t_k F(x^k) - \rho F(y^k) - \rho(1-t_k)s(y^k) + t_k r(x^k)\|^2 \\
&+ 2\gamma_k t_k \left\langle \bar{x}^k - x^k + \rho F(x^k), x^k - \bar{x}^k \right\rangle \\
&- 2\gamma_k t^k \rho \left\langle F(y^k) + (1-t_k)s(y^k), x^k - \bar{x}^k \right\rangle \\
&+ 2\gamma_k t_k \left\langle \bar{x}^k - x^k + \rho F(x^k), \bar{x}^k - x^* \right\rangle \\
&- 2\gamma_k \rho \left\langle F(y^k), y^k - x^* \right\rangle - 2\gamma_k (1-t_k)\rho \left\langle s(y^k), y^k - x^* \right\rangle \\
:= {}& \|x^k - x^*\|^2 + \gamma_k^2 \|\rho t_k F(x^k) - \rho F(y^k) - \rho(1-t_k)s(y^k) + t_k r(x^k)\|^2 \\
&+ I_1 + I_2 + I_3 + I_4 - 2\gamma_k \rho (1-t_k) \langle s(y^k), y^k - x^* \rangle.
\end{aligned}
$$

We estimate further

$$
\begin{aligned}
&I_1 + I_2 \\
= {}& 2\gamma_k t_k \left\langle \bar{x}^k - x^k + \rho F(x^k), x^k - \bar{x}^k \right\rangle \\
&- 2\gamma_k t^k \rho \left\langle F(y^k) + (1-t_k)s(y^k), x^k - \bar{x}^k \right\rangle \\
\leqslant {}& -2\gamma_k t_k \|r(x^k)\|^2 + 2\gamma_k t_k \rho \left\langle F(x^k) - F(y^k) - (1-t_k)s(y^k), x^k - \bar{x}^k \right\rangle \\
\leqslant {}& -2\gamma_k t_k (1-\rho L)\|r(x^k)\|^2 + 2\gamma_k t_k (\varphi(\bar{x}^k) - \varphi(y^k))
\end{aligned}
$$

The first inequality holds by (20). The second one is satisfied by the convexity of φ: since

$$\varphi(y^k) = \varphi(t_k \bar{x}^k + (1-t_k)x^k) \leqslant t_k \varphi(\bar{x}^k) + (1-t_k)\varphi(x^k),$$

we have

$$(1-t_k)(\varphi(\bar{x}^k) - \varphi(x^k)) \leqslant \varphi(\bar{x}^k) - \varphi(y^k).$$

For I_3 we have

$$
\begin{aligned}
I_3 &= 2\gamma_k t_k \left\langle \bar{x}^k - x^k + \rho F(x^k), x^k - x^* \right\rangle \\
&\leqslant 2\gamma_k t_k \left(\rho\varphi(x^*) - \rho\varphi(\bar{x}^k) \right).
\end{aligned}
$$

This inequality is derived from Lemma 1.1 with $\bar{x}^k = (I + \rho\partial\varphi + \rho N_K)^{-1}(I - \rho F)(x^k)$. Finally, for I_4, since x^* is a solution of (MVI) we have

$$\begin{aligned} I_4 &= -2\gamma_k\rho \left\langle F(y^k), y^k - x^* \right\rangle \\ &\leqslant 2\gamma_k\rho \left(\varphi(y^k) - \varphi(x^*)\right). \end{aligned}$$

Substituting the estimates $I_1 + I_2, I_3$ and I_4 to the estimate of $\|x^{k+1} - x^*\|$ we have

$$\begin{aligned} &\|x^{k+1} - x^*\|^2 \\ &= \|x^k - x^*\|^2 + \gamma_k^2\|\rho t_k F(x^k) - \rho F(y^k) - \rho(1-t_k)s(y^k) + t_k r(x^k)\|^2 \\ &\quad - 2\gamma_k t_k(1-\rho L)\|r(x^k)\|^2 + 2\gamma_k\rho t_k(\varphi(\bar{x}^k) - \varphi(y^k)) \\ &\quad + 2\gamma_k t_k\rho(\varphi(x^*) - \varphi(\bar{x}^k)) + 2\gamma_k\rho(\varphi(y^k) - \varphi(x^*)) \\ &\quad - 2\gamma_k(1-t_k)\rho\left\langle s(y^k), y^k - x^* \right\rangle \\ &= \|x^k - x^*\|^2 + \gamma_k^2\|\rho t_k F(x^k) - \rho F(y^k) - \rho(1-t_k)s(y^k) + t_k r(x^k)\|^2 \\ &\quad - 2\gamma_k t_k(1-\rho L)\|r(x^k)\|^2 + 2\gamma_k(1-t_k)\rho(\varphi(y^k) - \varphi(x^*) \\ &\quad + \left\langle s(y^k), y^k - x^* \right\rangle) \\ &\leqslant \|x^k - x^*\|^2 + \gamma_k^2\|\rho t_k F(x^k) - \rho F(y^k) - \rho(1-t_k)s(y^k) + t_k r(x^k)\|^2 \\ &\quad - 2\gamma_k t_k(1-\rho L)\|r(x^k)\|^2 \\ &= \|x^k - \hat{x}\|^2 - \frac{t_k^2(1-\rho L)^2\|r(x^k)\|^4}{\|\rho t_k F(x^k) - \rho(F(y^k) + (1-t_k)s(y^k)) + t_k r(x^k)\|^2}. \end{aligned}$$

The proof is complete. □

5. Implementing Algorithm 2.1

For implementing Algorithm 2.1, the only step of each iteration to be discussed is, given x, how to define

$$\bar{x}(\rho) := (I + \rho\partial\varphi)^{-1}(I - \rho F)(x). \tag{21}$$

We can apply the bundle method following the technique in [11]. Namely we solve approximately the following quadratic programming problem, for $u^0 = x$,

$$\begin{aligned} &\min_{u\in R^n, v\in\mathbb{R}} \left\{ \frac{1}{2\rho}\|u\|^2 + \langle F(x) - \frac{1}{\rho}x, u - x\rangle + v \right\}, \\ &s.t. \qquad v \geqslant \varphi(u^0) + \langle s(u^0), u - u^0\rangle, \quad s(u^0) \in \partial\varphi(u^0), \end{aligned}$$

to obtain u^1.

If u^{i-1} is not a solution of (21), we continue to solve the similar quadratic programming problem

$$\min_{u\in R^n, v\in\mathbb{R}} \Big\{ \frac{1}{2\rho}\|u\|^2 + \langle F(x) - \frac{1}{\rho}x, u-x\rangle + v \Big\},$$
$$s.t. \qquad v \geqslant \varphi(u^j) + \langle s(u^j), u-u^j\rangle, j=0,1,\ldots,i-1,$$

to obtain u^i.

In [11], it is proved that if $\partial\varphi$ is bounded on bounded subsets of $\mathbb{R}^n$, the sequence $\{u^i\}$ converges to a solution of (21).

Now we consider the same numerical example in [11] but we apply Algorithm 2.1 and use other stopping criteria.

Example 5.1. Consider (MVI), $F(x) = Qx$ for two cases $Q = Q_1$ and $Q = Q_2$, 10×10 nonsymmetric matrices, as below and

$$\begin{aligned}\varphi(x) &= \max_{1\leqslant j\leqslant 5} \varphi_j(x) + I_K(x)\\ &:= \max_{1\leqslant j\leqslant 5}\{x^T C^j x - d^{jT}x\} + I_K(x),\end{aligned}$$

where $I_K(x)$ is the indicator function of the subset K.

The data are given as follows

$$Q_1 = \mathrm{diag}(P^1, P^2, P^3, P^2, P^3),$$
$$Q_2 = \mathrm{diag}(P^4, P^2, P^5, P^3),$$

where

$$P^1 = \begin{pmatrix} 1.6 & -1\\ 1 & 1.6\end{pmatrix}, \quad P^2 = \begin{pmatrix} 1.5 & 1\\ -1 & 1.5\end{pmatrix}, \quad P^3 = \begin{pmatrix} 2 & -1\\ 1 & 2\end{pmatrix},$$

$$P^4 = \begin{pmatrix} 1.5 & 1 & 2 & -1\\ -1 & 1.5 & 1 & 2\\ -2 & 1 & 1.6 & 1\\ -1 & -2 & -1 & 1.6\end{pmatrix}, \quad P^5 = \begin{pmatrix} 2 & 0\\ 0 & 2\end{pmatrix},$$

C^j is 10×10 symmetric matrix defined by

$$C^j_{ik} = \exp\left(\frac{i}{k}\right)\cos(ik)\sin(j), \quad i<k,$$

$$C^j_{ii} = \frac{i}{n}|\sin(j)| + \sum_{k\neq i}|C^j_{ik}|,$$

d^j is the vector in $\mathbb{R}^n$ whose components are $d^j_i = \exp\left(\frac{i}{j}\right)\sin(ij)$, and the subset K is

$$K = \Big\{x\in R^{10} \ : \ \sum_{i=1}^{10} x_i \geqslant 1, \ -5\leqslant x_i\leqslant 5, \ i=1,\ldots,10\Big\}.$$

This example is known as a nonsmooth optimization test problem.

(a) The case of Q_1. We take $L = 2.24$, $\rho = 0.18$, $x^0 = (1, 1, \ldots, 1)$. The stopping criterion for the bundle method to solve (21) for each outer iteration is $\|u^{i+1} - u^i\| < \varepsilon^l$, $l = 1, 2$. Iterations of this bundle method are called inner iterations. The stopping criterion for Algorithm 2.1 is $\|x^k - \bar{x}^k\| < \beta^l$. We use the MATLAB, applying the quadratic - program solver quadprog.m from the MATLAB Optimization Toolbox. For $l = 1$, $\varepsilon^1 = 0.001$, and $\beta^1 = 0.001$, the result is given by the following table.

outer iteration	inner iteration	$r(x^k, \rho_k)$
1	14	1.8633
2	15	0.6985
3	17	0.3021
4	16	0.1051
5	18	0.0295
6	18	0.0092
7	18	0.0044
8	18	0.0027
9	18	0.0018
10	18	0.0012
11	18	0.0008

The obtained approximate solution of (MVI) is

$$x^* = (-0.00, 0.00, 0.09, -0.00, 1.34, 0.00, 0.43, 0.47, 0.46, 0.25).$$

For $l = 2$ and $\varepsilon^2 = \beta^2 = 0.00001$, the result is

outer iteration	inner iteration	$r(x^k, \rho_k)$
1	14	1.863380
2	15	0.698538
3	17	0.302186
4	16	0.105128
5	18	0.029522
6	18	0.009216
7	18	0.004415
8	18	0.002762
9	18	0.001843
10	18	0.001234
11	18	0.000822
12	18	0.000545
13	18	0.000361
14	18	0.000239
15	18	0.000158
16	18	0.000105

outer iteration	inner iteration	$r(x^k, \rho_k)$
17	18	0.000069
18	18	0.000046
19	18	0.000030
20	18	0.000020
21	18	0.000013
22	18	0.000008

The obtained approximate solution of (MVI) is

$$x^* = (0.00, 0.00, 0.09, -0.00, 1.34, 0.00, 0.43, 0.47, 0.46, 0.25).$$

(b) The case of Q_2. We take $L = 3.94$, $\rho = 0.1280$, $x^0 = (1, 1, \dots, 1)$. The stopping criteria are the same as in (a). For $l = 1$, the obtained approximate solution of (MVI) is

$$x^* = (0.00, 0.00, 0.00, -0.00, 1.12, 0.01, 0.40, 0.41, 0.32, 0.17).$$

outer iteration	inner iteration	$r(x^k, \rho_k)$
1	8	1.8797
2	15	0.7664
3	15	0.3309
4	16	0.2190
5	12	0.0909
6	12	0.0430
7	12	0.0252
8	12	0.0173
9	12	0.0141
10	12	0.0118
11	12	0.0095
12	12	0.0077
13	12	0.0062
14	12	0.0049
15	12	0.0039
16	12	0.0029
17	12	0.0021
18	12	0.0015
19	12	0.0011
20	12	0.0008

For $l = 2$ the result is

outer iteration	inner iteration	$r(x^k, \rho_k)$
1	8	1.879768
2	15	0.766487
3	15	0.330983

outer iteration	inner iteration	$r(x^k, \rho_k)$
4	16	0.219021
5	12	0.090969
6	12	0.043078
7	12	0.025216
8	12	0.017342
9	12	0.014188
10	12	0.011811
11	12	0.009585
12	12	0.007717
13	12	0.006248
14	12	0.004995
15	12	0.003933
16	12	0.002941
17	12	0.002137
18	12	0.001556
19	12	0.001152
20	12	0.000862
21	12	0.000639
22	12	0.000462
23	12	0.000326
24	12	0.000362
25	12	0.000226
26	12	0.000115
27	12	0.000086
28	12	0.000065
29	12	0.000048
30	12	0.000035
31	12	0.000024
32	12	0.000017
33	12	0.000012
34	12	0.000009

The obtained approximate solution of (MVI) is

$$x^* = (-0.00, 0.00, 0.00, 0.00, 1.12, 0.01, 0.40, 0.41, 0.32, 0.17).$$

References

[1] P. N. Anh and L. D. Muu, *Coupling the Banach contraction mapping principle and the proximal point algorithm for solving monotone variational inequalities*, Acta Math. Vietnam. **29** (2004), 119-133.

[2] P. N. Anh, L. D. Muu, V. H. Nguyen and J. J. Strodiot, *On the contraction and nonexpansiveness properties of the marginal mappings in generalized variational inequalities involving co-coercive operators*, Generalized Convexity, Generalized Monotonicity and Applications, Springer, New York, 2005, pp. 89-111.

[3] Q. H. Ansari and J. C. Yao, *Iterative schemes for solving mixed variational-like inequalities*, J. Optim. Theory Appl. **108** (2001), 527-541.

[4] T. Q. Bao and P. Q. Khanh, *A projection type algorithm for pseudomonotone nonlipschitz multivalued variational inequalities*, Generalized Convexity, Generalized Monotonicity and Applications, Springer, New York, 2005, pp. 113-129.

[5] H. Brézis, *Operateurs Maximaux Monotones* (1973), North - Holland, Amsterdam.

[6] G. Cohen, *Auxiliary problem principle extended to variational inequalities*, J. Optim. Theory Appl. **59** (1988), 325-333.

[7] J. Douglas and H. H. Rachford, *On the numerical solutions of the heat conduction problem in 2 and 3 space variables*, Trans. Amer. Math. Soc. **82** (1956), 421-439.

[8] F. Facchinei and J. S. Pang, *Finite-dimensional variational inequalities and complementary problems*, Vol. 1 - 2 (2003), Springer, New York.

[9] A. A. Goldstein, *Convex programming in Hilbert space*, Bulletin Amer. Math. Soc. **70** (1964), 709-710.

[10] S. Haubrige, V. H. Nguyen and J. J. Strodiot, *Convergence analysis and application of the Glowinski-Le Tallect splitting method for finding a zero of the sum of two maximal monotone operators*, J. Optim. Theory Appl. **97** (1998), 645-673.

[11] T. T. Hue and J. J. Strodiot, *Convergence analysis of a relaxed extragradient-proximal point algorithm application to variational inequalities*, Optim. **54** (2005), 191-213.

[12] A. N. Iusem, *An iterative algorithm for the variational inequality problem*, Math. Appl. Comput. **13** (1994), 103-114.

[13] A. N. Iusem and L. R. L. Pérez, *An extragradient type algorithm for nonsmooth variational inequalities*, Optimization **48** (2000), 309-332.

[14] A. N. Iusem and B. F. Svaiter, *A variant of Korpelevich's method for variational inequalities with a new search strategy*, Optimization **42** (1997), 309-321.

[15] G. M. Korpelevich, *The extragradient method for finding saddle points and other problems*, Ecomomika i Matematicheskie Metody **12** (1976), 747-756.

[16] C. Lemaréchal and R. Mifflin, *Nonsmooth Optimization*, IIASA Proceedings Series **3** (1978), Pergamon Press, Oxford.

[17] E. S. Levitin and B. T. Polyak, *Constrained minimization problem*, USSR Comput. Math. Math. Phys. **6** (1966), 1-50.

[18] P. L. Lions and B. Mercier, *Splitting algorithms for the sum of two monotone operators*, SIAM J. Num. Anal. **16** (1979), 964-979.

[19] B. Martinet, *Regularization d'inéquations variationelles par approximations successives*, Revue d' Automatique Informatique et Recherche Opérationelle **4** (1970), 154-158.

[20] M. A. Noor, *A new iterative method for monotone mixed variational inequalities*, Math. Comput. Model. **26** (1997), 29-34.

[21] M. A. Noor, *An extraresolvent method for monotone mixed variational inequalities*, Math. Comput. Model. **29** (1999), 95-100.

[22] M. A. Noor, *A modified extragradient method for general monotone variational inequalities*, Comput. Math. Appl. **38** (1999), 19-24.

[23] M. A. Noor, *New extragradient-type methods for general variational inequalities*, J. Math. Anal. Appl. **277** (2003), 379-394.

[24] M. A. Noor, *Pseudomonotone general mixed variational inequalities*, Appl. Math. Appl. **141** (2003), 529-540.

[25] M. A. Noor, *Extragradient methods for pseudomonotone variational inequalities*, J. Optim. Theory Appl. **117** (2003), 475-488.

[26] M. A. Noor, Y. J. Wang and N. H. Xiu, *Projection iterative schemes for general variational inequalities*, J. Ineq. Pure Appl. Math. **3** (2002), 1-8.

[27] A. Renaud and G. Cohen, *An extension of the auxiliary problem principle to nonsymmetric auxiliary operators*, ESAIM: Control Optim. Cal. Vari. **2** (1997), 281-306.

[28] R. T. Rockafellar, *Monotone operators and the proximal point algorithm*, SIAM J. Control Optim. **14** (1976), 877-898.

[29] M. Sibony, *Méthodes itératives pour les équations et inéquations aux dérivées partielles non linéaires de type monotone*, Calcolo **7** (1970), 65-183.
[30] M. V. Solodov and B. F. Svaiter, *A new projection method for variational inequality problems*, SIAM J. Control Optim. **34** (1996), 1814-1830.
[31] M. V. Solodov and P. Tseng, *Modified projection - type methods for monotone variational inequalities*, SIAM J. Control Optim. **34** (1996), 1814-1830.
[32] D. Sun, *A new step size skill for solving a class of nonlinear projection equations*, J. Comput. Math. **13** (1995), 357-368.
[33] P. Tseng, *Alternating projection-proximal methods for convex programming and variational inequalities*, SIAM J. Optim. **7** (1997), 951-965.
[34] P. Tseng, *A modified Forward-Backward splitting method for maximal monotone mapping*, SIAM J. Control Optim. **38** (2000), 431-446.
[35] Y. J. Wang, N. H. Xiu and C. Y. Wang, *Unified framework of extragradient-type methods for pseudomonotone variational inequalities*, J. Optim. Theory Appl. **111** (2001), 641-656.
[36] Y. J. Wang, N. H. Xiu and C. Y. Wang, *A new version of the extragradient method for variational inequality problems*, Comput. Math. Appl. **42** (2001), 969-979.
[37] N. H. Xiu and J. Z. Zhang, *Local convergence analysis of projection-type algorithms: unified approach*, J. Optim. Theory Appl. **115** (2002), 211-230.

Department of Mathematics
Wayne State University, U.S.A.

Department of Mathematics
International University,
Vietnam National University of Hochiminh City.