Available online at www.sciencedirect.com

SCIENCE DIRECT

J. Math. Anal. Appl. 302 (2005) 166–179

Journal of MATHEMATICAL ANALYSIS AND APPLICATIONS

www.elsevier.com/locate/jmaa

On variable-step relaxed projection algorithm for variational inequalities ☆

Qingzhi Yang [a]

[a] *School of Mathematics and LPMC, Nankai University, Tianjin 300071, PR China*

Received 5 May 2004

Available online 2 October 2004

Submitted by M.A. Noor

Abstract

Projection algorithms are practically useful for solving variational inequalities (VI). However some among them require the knowledge related to VI in advance, such as Lipschitz constant. Usually it is impossible in practice. This paper studies the variable-step basic projection algorithm and its relaxed version under weakly co-coercive condition. The algorithms discussed need not know constant/function associated with the co-coercivity or weak co-coercivity and the step-size is varied from one iteration to the next. Under certain conditions the convergence of the variable-step basic projection algorithm is established. For the practical consideration, we also give the relaxed version of this algorithm, in which the projection onto a closed convex set is replaced by another projection at each iteration and latter is easy to calculate. The convergence of relaxed scheme is also obtained under certain assumptions. Finally we apply these two algorithms to the Split Feasibility Problem (SFP).

Keywords: Variational inequality; Weakly co-coercive; Projection; Relaxation; SFP

☆ This work is supported by the National Natural Science Foundation of China, Grant 10171055.

E-mail address: qz-yang@nankai.edu.cn.

0022-247X/$ – see front matter
doi:10.1016/j.jmaa.2004.07.048

1. Introduction

Variational inequality problem is to find an $x \in C$ such that

$$\langle f(x), y - x \rangle \geqslant 0, \quad \text{for all } y \in C, \tag{1.1}$$

where C is a nonempty closed convex subset in R^n and $f(x)$ is a mapping from R^n to itself and $\langle \, , \rangle$ denotes the inner product. Due to their wide applications in various fields, variational inequalities (VI) have received great attention since 1970s, and achieved the fruitful results in the theory as well as applications. The interested reader may consult a two-volume monograph by F. Facchinei and J.S. Pang, which presents a comprehensive, state-of-the art treatment of the finite dimensional variational inequality and complementarity problems [1].

Numerous algorithms for VI have been proposed. Among them, projection-type methods are simple in form and useful in practice provided the projection onto C is easy to calculate. Various projection algorithms, such as basic projection algorithm, extragradient projection algorithm and hyperplane projection algorithm, have been designed to solve the different class of VIs (see, e.g., [1,2,4,6,12,14–21] and references therein). Generally speaking, each projection algorithm is confined in certain class of VIs so that the convergence of the algorithm can be guaranteed. So one usually hopes that an effective algorithm may be used in a broader scope if it possible. For example, consider the following basic projection algorithm with a constant step

$$x^{k+1} = P_C\left[x^k - \gamma f\left(x^k\right)\right]. \tag{1.2}$$

In the early stages of studying projection methods, $f(x)$ was required to be strongly monotone and Lipschitz continuous with small γ for the convergence of the algorithm. Later, this condition is weaken to only require the co-coercivity of $f(x)$ (see [1, p. 1111], [9,13]) while γ is chosen in an interval related to the co-coercive constant. However, in practice we usually cannot get the knowledge of that constant in advance.

Therefore some algorithms are specially proposed so that they may be performed without the requirement of prior knowledge related to $f(x)$. Generally the step-size in this class of algorithms is varied from one iteration to the next in order to guarantee the convergence of an algorithm. In this paper we study the variable-step basic projection algorithm solving VI, which has the following form:

$$x^{k+1} = P_C\left[x^k - \gamma_k f\left(x^k\right)\right]. \tag{1.3}$$

Though in ([1, Algorithm 12.1.4], [20,21]) a variable-step basic projection algorithm is given, actually there variable steps depend on the co-coercive constant. So in this paper we focus on a variable-step basic projection algorithm, where the variable steps are independent of the co-coercive constant. The first algorithm discussed below was actually proposed by Auslender [4] in 1970s. Later Fukushima [6] gave its relaxed version for practical purpose. Both Auslender and Fukushima assumed the strong monotonicity of $f(x)$ for the convergence of their algorithms. This is a strict limit. We establish the convergence of Auslender's algorithm as well as Fukushima's algorithm under weaker conditions. Mainly the strong monotonicity is replaced by the weak co-coercivity, whose a special case is the co-coercivity, in the proofs of convergence. Therefore we extend the Auslender's algorithm as well as Fukushima's algorithm to a broader class of VIs. Really Algorithm 1 below may also be viewed as an improvement of ([1, Algorithm 12.1.4], [20,21]) since the weak co-coercivity is a weaker condition than co-coercivity. Finally we consider an application of algorithms discussed above to the Split Feasibility Problem (SFP), which arises in image reconstruction. An algorithm proposed by Byrne for solving the SFP is improved by transforming the SFP equivalently into a special case of VI.

This paper is organized as follows. In Section 2 we establish the convergence of Auslender's algorithm under certain conditions, and an example is given to show that all conditions supposed in this section are satisfied but $f(x)$ is neither strongly monotone nor co-coercive. In Section 3, based on the similar way used in [6], the convergence of the relaxed projection algorithm is obtained under certain assumptions. In Section 4 we apply the algorithms discussed in above two sections to the SFP.

2. A basic projection algorithm with variable steps

First we make some assumptions as following.

(A_1) $f(x)$ is a continuous mapping from $C \to R^n$.

(B_1) There is a positive continuous function $g(x, y)$ on C such that

$$\langle f(x) - f(y), x - y\rangle \geqslant g(x, y)\|f(x) - f(y)\|_2^2, \quad \forall x, y \in C. \tag{2.1}$$

(C_1) For some $z \in C$, there exists a $\beta > 0$ and a bounded open set $D \subset R^n$ such that

$$\langle f(x), x - z\rangle \geqslant \beta \cdot \|f(x)\|_2, \quad \forall x \in C/D.$$

(D_1) If there is a solution x^* of (1.1) with $f(x^*) \neq 0$, then $f(x) = f(x^*)$ implies $x = x^*$ for $x \in C$.

Remark. If (B_1) holds, we call $f(x)$ weakly co-coercive. If $g(x, y)$ is a constant $\eta > 0$ or $g(x, y)$ has an infimum $\eta > 0$, then $f(x)$ is called co-coercive [1] or inverse strongly monotone(ism) [9]. However if C is unbounded and $g(x, y)$ tends to zero as $\|x\|_2$ or $\|y\|_2$ approaches infinity, $f(x)$ is not co-coercive. So the weak co-coercivity is a weaker condition than co-coercivity. In Example 1 below we are to discuss a variational inequality problem with weakly co-coercive but no co-coercive $f(x)$.

Let $\{\rho_k\}$ be a sequence of positive parameters satisfying

$$\sum_{k=1}^{\infty} \rho_k = \infty, \qquad \sum_{k=1}^{\infty} \rho_k^2 < +\infty. \tag{2.2}$$

We may state following algorithm without necessity of knowing $g(x, y)$ in advance.

Algorithm 1. For an arbitrary x^1, set $k = 1$. If $f(x^k) = 0$, stop, x^k is a solution of (1.1); Otherwise calculate the projection of $x^k - \frac{\rho_k}{\|f(x^k)\|_2} \cdot f(x^k)$ onto C, let it be x^{k+1}:

$$x^{k+1} = P_C\left[x^k - \frac{\rho_k}{\|f(x^k)\|_2} \cdot f\left(x^k\right)\right]$$

then $k = k + 1$. If $x^{k+1} = x^k$, terminate, x^k is a solution of (1.1); otherwise repeat.

Remark. Algorithm 1 was first proposed in [4] in 1970s, Auslender proved its convergence by assuming the strong monotonicity and boundedness of $f(x)$. Later based on it Fukushima [6] established the convergence of its relaxed version under the strong monotonicity and other conditions.

It is clear that x^k is a solution of (1.1) if the algorithm terminates at x^k. So in the remainder of this section we focus on the convergence of Algorithm 1 for the case of $\{x^k\}$ being an infinite sequence.

Lemma 2.1. *x^* is a solution of* (1.1) *if and only if $x^* = P_C[x^* - \alpha \cdot f(x^*)]$ for any given $\alpha > 0$.*

Lemma 2.2. *For any $x, y \in R^n$*

$$\left\|P_C(x) - P_C(y)\right\|_2^2 \leqslant \|x - y\|_2^2 - \left\|\left(P_C(x) - x\right) - \left(P_C(y) - y\right)\right\|_2^2.$$

These two lemmas are well known for VI.

Lemma 2.3. *Assume that (A_1), (B_1), (C_1) hold and $\{x^k\}$ is the sequence generated by Algorithm* 1, *then $\{x^k\}$ is bounded.*

This lemma can be easily proved with the same way as the proof of Lemma 3 in [6].

Theorem 2.1. *Suppose that (A_1), (B_1), (D_1) hold and* (1.1) *has at least a solution. Assume that $\{x^k\}$ is an infinite sequence produced by Algorithm* 1 *and is bounded. Then any accumulation point of $\{x^k\}$ is a solution of* (1.1). *Furthermore, if* (1.1) *has no solution at which $f(x)$ vanishes, then the solution of* (1.1) *is unique and $\{x^k\}$ converges to this solution.*

Proof. Let x^* be a solution of (1.1). By Lemma 2.1 we have $x^* = P_C[x^* - \frac{\rho_k}{\|f(x^k)\|_2} \cdot f(x^*)]$ for each k. Then

$$\begin{aligned}
\left\|x^{k+1} - x^*\right\|_2^2 &= \left\|P_C\left[x^k - \frac{\rho_k}{\|f(x^k)\|_2} \cdot f\left(x^k\right)\right] - P_C\left[x^* - \frac{\rho_k}{\|f(x^*)\|_2} \cdot f(x^*)\right]\right\|_2^2 \\
&\leqslant \left\|\left(x^k - \frac{\rho_k}{\|f(x^k)\|_2} \cdot f\left(x^k\right)\right) - \left(x^* - \frac{\rho_k}{\|f(x^k)\|_2} \cdot f(x^*)\right)\right\|_2^2 \\
&\quad - \left\|\left(x^{k+1} - \left(x^k - \frac{\rho_k}{\|f(x^k)\|_2} \cdot f\left(x^k\right)\right)\right)\right.
\end{aligned}$$

$$
\begin{aligned}
&\quad - \left(x^* - \left(x^* - \frac{\rho_k}{\|f(x^k)\|_2} \cdot f(x^*) \right) \right) \Big\|_2^2 \\
&= \|x^k - x^*\|_2^2 - 2\frac{\rho_k}{\|f(x^k)\|_2} \langle x^k - x^*, f(x^k) - f(x^*) \rangle \\
&\quad + \left(\frac{\rho_k}{\|f(x^k)\|_2} \right)^2 \|f(x^k) - f(x^*)\|_2^2 \\
&\quad - \left\| (x^{k+1} - x^k) + \frac{\rho_k}{\|f(x^k)\|_2} (f(x^k) - f(x^*)) \right\|_2^2 \\
&= \|x^k - x^*\|_2^2 - 2\frac{\rho_k}{\|f(x^k)\|_2} \langle x^k - x^*, f(x^k) - f(x^*) \rangle \\
&\quad - \|x^{k+1} - x^k\|_2^2 - 2\frac{\rho_k}{\|f(x^k)\|_2} \langle x^{k+1} - x^k, f(x^k) - f(x^*) \rangle \\
&\leqslant \|x^k - x^*\|_2^2 - 2\frac{\rho_k}{\|f(x^k)\|_2} g(x^k, x^*) \|f(x^k) - f(x^*)\|_2^2 \\
&\quad - \|x^{k+1} - x^k\|_2^2 - 2\frac{\rho_k}{\|f(x^k)\|_2} \langle x^{k+1} - x^k, f(x^k) - f(x^*) \rangle.
\end{aligned} \tag{2.3}
$$

The first inequality follows from Lemma 2.2 and the last inequality from condition (B_1).

Because

$$
\begin{aligned}
&-2\frac{\rho_k}{\|f(x^k)\|_2} \langle x^{k+1} - x^k, f(x^k) - f(x^*) \rangle \\
&\quad \leqslant 2\frac{\rho_k}{\|f(x^k)\|_2} g(x^k, x^*) \|f(x^k) - f(x^*)\|_2^2 \\
&\qquad + \frac{1}{2g(x^k, x^*)} \frac{\rho_k}{\|f(x^k)\|_2} \|x^{k+1} - x^k\|_2^2.
\end{aligned}
$$

Then we deduce from (2.3)

$$
\|x^{k+1} - x^*\|_2^2 \leqslant \|x^k - x^*\|_2^2 - \left(1 - \frac{1}{2g(x^k, x^*)} \frac{\rho_k}{\|f(x^k)\|_2} \right) \|x^{k+1} - x^k\|_2^2. \tag{2.4}
$$

Since we assume that $\{x^k\}$ is bounded, then there exists a closed ball $B_M = \{x \in R^n \mid \|x\|_2 \leqslant M\}$ containing $\{x^k\}$ and x^*. Since by assumption (B_1), $g(x, y)$ is continuous on C, then $g(x, y)$ is uniformly continuous on $C \cap B_M$, which means that there exists a $\delta > 0$ such that

$$
g(x, x^*) \geqslant \delta, \quad \forall x \in C \cap B_M. \tag{2.5}
$$

As a result we conclude that $g(x^k, x^*) \geqslant \delta$, $\forall k$. Therefore from (2.4) it follows

$$
\|x^{k+1} - x^*\|_2^2 \leqslant \|x^k - x^*\|_2^2 - \left(1 - \frac{1}{2\delta} \frac{\rho_k}{\|f(x^k)\|_2} \right) \|x^{k+1} - x^k\|_2^2. \tag{2.6}
$$

Next we discuss two cases.

Case 1. $\inf\{\|f(x^k)\|_2\} > 0$. In this case there exists a $\xi > 0$ such that $\|f(x^k)\|_2 \geqslant \xi$ for all k. Then we have from (2.6)

$$
\|x^{k+1} - x^*\|_2^2 \leqslant \|x^k - x^*\|_2^2 - \left(1 - \frac{1}{2\delta\xi} \rho_k \right) \|x^{k+1} - x^k\|_2^2.
$$

Since ρ_k tends to zero, there exists $\bar{k} > 0$ such that

$$1 - \frac{1}{2\delta\xi}\rho_k \geqslant 1/2, \quad \forall k \geqslant \bar{k}.$$

Therefore $\{\|x^k - x^*\|_2\}$ is monotonically decreasing for $k \geqslant \bar{k}$, and $\|x^{k+1} - x^k\|_2 \to 0$ as $k \to +\infty$. Hence we conclude that $x^k \to \bar{x}$ as $k \to +\infty$ and $\bar{x} \in C \cap B_M$.

Notice

$$\begin{aligned} &-2\frac{\rho_k}{\|f(x^k)\|_2}\langle x^{k+1} - x^k, f(x^k) - f(x^*)\rangle \\ &\quad \leqslant \|x^{k+1} - x^k\|_2^2 + \left(\frac{\rho_k}{\|f(x^k)\|_2}\right)^2 \|f(x^k) - f(x^*)\|_2^2. \end{aligned}$$

From (2.3) we derive

$$\begin{aligned} \|x^{k+1} - x^*\|_2^2 \leqslant{}& \|x^k - x^*\|_2^2 - 2\delta\frac{\rho_k}{\|f(x^k)\|_2}\|f(x^k) - f(x^*)\|_2^2 \\ &+ \left(\frac{\rho_k}{\|f(x^k)\|_2}\right)^2 \|f(x^k) - f(x^*)\|_2^2. \end{aligned} \tag{2.7}$$

Denote $\tau = \max_{x \in C \cap B_M}\{\|f(x)\|_2\}$, then $\xi \leqslant \|f(x^k)\|_2 \leqslant \tau$, $\forall k$. From (2.7) we get

$$\|x^{k+1} - x^*\|_2^2 \leqslant \|x^k - x^*\|_2^2 - \left(2\frac{\delta\rho_k}{\tau} - \left(\frac{\rho_k}{\xi}\right)^2\right)\|f(x^k) - f(x^*)\|_2^2. \tag{2.8}$$

If $\inf\{\|f(x^k) - f(x^*)\|_2\} = 0$, then $f(\bar{x}) = f(x^*)$ which implies $\bar{x}$ is a solution of (1.1) from (D_1). Otherwise we have $\|f(x^k) - f(x^*)\|_2 \geqslant \sigma$ for all k, where $\sigma > 0$ is a constant. Thus we get from (2.8)

$$\|x^{k+1} - x^*\|_2^2 \leqslant \|x^k - x^*\|_2^2 - \left(2\frac{\delta\rho_k}{\tau} - \left(\frac{\rho_k}{\xi}\right)^2\right)\sigma^2 \tag{2.9}$$

for $k \geqslant \bar{\bar{k}}$, where $\bar{\bar{k}}$ is a sufficiently large positive scalar.

Because $\sum_{k=1}^{\infty}\rho_k = \infty$, $\sum_{k=1}^{\infty}\rho_k^2 < +\infty$, from (2.9) we deduce

$$\|x^{\bar{\bar{k}}+l} - x^*\|_2^2 \leqslant \|x^{\bar{\bar{k}}} - x^k\|_2^2 - \left(\frac{2\delta}{\tau}\sum_{i=\bar{\bar{k}}}^{\bar{\bar{k}}+l-1}\rho_i - \frac{1}{\xi^2}\sum_{i=\bar{\bar{k}}}^{\bar{\bar{k}}+l-1}\rho_i^2\right)\sigma^2$$

which leads to

$$\|x^{\bar{\bar{k}}+l} - x^*\|_2^2 < 0$$

for sufficiently large l. This contradiction shows that the case of $\|f(x^k) - f(x^*)\|_2 \geqslant \sigma$ for all k would not occur.

Case 2. $\inf\{\|f(x^k)\|_2\} = 0$. Since $\{x^k\}$ is bounded, there exists at least an accumulation point of $\{x^k\}$. Consequently for every accumulation point $\bar{\bar{x}}$ of $\{x^k\}$, $f(\bar{\bar{x}}) = 0$ which means $\bar{\bar{x}}$ is a solution of (1.1). If this is not true, we assume $x^{k_i} \to \bar{\bar{x}} \in C$ as $k_i \to +\infty$ with $f(\bar{\bar{x}}) \neq 0$. Since

$$x^{k_i+1} = P_C\left[x^{k_i} - \frac{\rho_{k_i}}{\|f(x^{k_i})\|_2} \cdot f(x^{k_i})\right]$$

by the algorithm, we have $x^{k_i+1} \to P_C[\bar{\bar{x}}] = \bar{\bar{x}}$ as $k_i \to +\infty$. Repeat this step, we obtain $x^k \to \bar{\bar{x}}$, which contradicts $\inf\{\|f(x^k)\|_2\} = 0$.

Combining these two Cases we obtain the desirable conclusions.

Furthermore, if (1.1) has no solution at which $f(x)$ vanishes, then we deduce from (B_1) and (D_1) that the solution of (1.1) is unique. From above argument we conclude that $\{x^k\}$ converges to this solution. □

Theorem 2.2. *Let the set C be bounded and (A_1), (B_1) hold. Then* (1.1) *exists a solution and $\{x^k\}$ is bounded. In addition, if (D_1) holds, then the conclusions of Theorem* 2.1 *remain valid.*

Proof. Since $f(x)$ is monotone and continuous on C, we know that (1.1) has a solution as long as C is bounded [1, Proposition 2.2.3], [7]. And obviously $\{x^k\}$ is bounded. So this corollary directly follows from Theorem 2.1. □

Theorem 2.3. *If the set C is unbounded. Assume (A_1), (B_1), (C_1) hold. Then* (1.1) *exists a solution and $\{x^k\}$ is bounded. In addition, if (D_1) holds, then the conclusions of Theorem* 2.1 *remain valid.*

Proof. If $f(x) = 0$ for some $x \in C$, then clearly x is a solution of (1.1). Otherwise, since $f(x) \neq 0$ for any $x \in C/D$, it follows from (C_1)

$$\big\langle f(x), x - z\big\rangle > 0, \quad \forall x \in C/D.$$

Hence (1.1) has a solution [1, Proposition 2.2.3], [7]. Additionally we see that $\{x^k\}$ is bounded by Lemma 2.3. Thus we conclude that the conclusions of Theorem 2.1 remain valid. □

Now we give an illustrative example, in which (A_1), (B_1), (C_1), (D_1) are satisfied, but $f(x)$ is neither strongly monotone nor co-coercive.

Example 1 (see [6, p. 60]). Consider variational inequality problem (1.1) with $f(x) = 1 - e^{-x}$, $C = (-\infty, +\infty)$.

We verify conditions (A_1), (B_1), (C_1) and (D_1).

(1) Obviously $f(x)$ is continuous.
(2) For any $x, y \in C$,

$$\big\langle f(x) - f(y), x - y\big\rangle = (e^{-y} - e^{-x})(x - y) \geqslant \min(e^x, e^y)\big|f(x) - f(y)\big|^2.$$

Denote $g(x, y) = \min(e^x, e^y)$, then (B_1) holds.
(3) Take $z = 1 \in C$ and $D = (-1, 2)$, then

$$\big\langle f(x), x - z\big\rangle = f(x)(x - 1) \geqslant \big|f(x)\big| \quad \text{for any } x \in C/D.$$

So (C_1) is satisfied. Moreover $f(x) \neq 0$ for any $x \in C/D$.
(4) For $f(x) = 1 - e^{-x}$, obviously $f(x) = f(y)$ implies $x = y$.

If $\langle f(x) - f(y), x - y\rangle > 0$ $\forall x, y \in C$ and $x \neq y$, then $f(x)$ is called strictly monotone. It is known that the solution of VI is unique under strict monotonicity provided VI has a solution. In above example, it is easy to see that $f(x)$ is strictly monotone, so this problem has a unique solution from Theorem 2.3. It is clear that $x^* = 0$ is the unique solution of the problem. Since $f(x^*) = 0$, so really it need not to check (D_1). Then from Theorem 2.1 and Theorem 2.3 we know that the sequence $\{x^k\}$ produced by Algorithm 1 converges to the unique solution x^*.

Remark. Note that in (2) $g(x, y)$ tends to zero as x or y approaches $-\infty$. Actually it is easy to verify that $f(x)$ in example is neither strongly monotone nor co-coercive. Therefore Algorithm 1 is applicable to a broader class of VIs.

3. A relaxed basic projection algorithm with variable steps

For projection-type algorithms, whether or not the projection can be solved efficiently is a crucial problem. Except for some particular situations, such as C is the nonnegative orthant R^n_+, generally it is not a trivial work to compute a projection onto C. Even it is impossible in some cases to get an exact projection onto C. If so, the overall efficiency of a projection method will be seriously affected. To overcome this difficulty, some inexact projection algorithms were proposed (see, e.g, [3,5,6,14]). Among them, the relaxed projection algorithm for solving (1.1) proposed by Fukushima [6] is quite attractive. The essential idea of that method is to utilize outer approximations to the closed convex set C. In detail, at kth iteration, the projection P_C onto C is replaced by P_{C^k} while latter may be solved easily. Fukushima established the convergence of his relaxed projection algorithm with strong monotonicity and other assumptions. In [1, p. 1223] authors comment that "The computational advantages of the approach are evident but should be weighted against the rather strong assumptions needed for convergence." Therefore it is significant to weaken assumption conditions if it is permitted. Like in previous section, we here extend the Fukushima's algorithm to a broader scope.

First we recall some notations.

Define the distance from a point $x \in R^n$ to C by

$$\mathrm{dist}[x, C] = \min\bigl\{\|x - z\|_2 \mid z \in C\bigr\}$$

and denote for each $\delta > 0$

$$C_\delta = \bigl\{x \in R^n \mid \mathrm{dist}[x, C] < \delta\bigr\}.$$

Assume that the following conditions are satisfied.

(A_2) $f(x)$ is a mapping from R^n to itself and is continuous on C_δ for some $\delta > 0$.

(B_2) There is a positive continuous function $g(x, y)$ on C_δ such that

$$\bigl\langle f(x) - f(y), x - y\bigr\rangle \geqslant g(x, y)\bigl\|f(x) - f(y)\bigr\|_2^2, \quad \forall x, y \in C_\delta.$$

(C_2) For some $z \in C$, there exists a $\beta > 0$ and a bounded open set $D \subset R^n$ such that

$$\bigl\langle f(x), x - z\bigr\rangle \geqslant \beta \cdot \bigl\|f(x)\bigr\|_2, \quad \text{for all } x \notin D.$$

(D_2) The set C is given by

$$C = \{x \in R^n \mid c(x) \leqslant 0\}$$

where $c: R^n \to R$ is a convex function and there exists a point x^0 such that $c(x^0) < 0$.

(E_2) For any $x \in R^n$, at least one subgradient $g \in \partial c(x)$ can be calculated, where

$$\partial c(x) = \{g \in R^n \mid c(y) \geqslant c(x) + \langle g, y - x\rangle,\ \forall y \in R^n\}.$$

(F_2) If there is a solution x^* of (1.1) with $f(x^*) \neq 0$, then $f(x) = f(x^*)$ implies $x = x^*$ for $x \in C$.

For convenience, we restate the algorithm proposed by Fukushima [6]. Then we prove the convergence of the algorithm under above conditions.

Let $\{\rho_k\}$ be that sequence given in previous section.

Algorithm 2.

Step 0: Select a starting point x^1 and set $k = 1$;

Step 1: Choose $g^k \in \partial c(x^k)$ and let

$$C^k = \{x \in R^n \mid c(x^k) + \langle g^k, x - x^k\rangle \leqslant 0\};$$

Step 2: Obtain the projection of $x^k - \frac{\rho_k}{\|f(x^k)\|_2} f(x^k)$ onto the halfspace C^k and let it be x^{k+1}, i.e.,

$$x^{k+1} = P_{C^k}\left[x^k - \frac{\rho_k}{\|f(x^k)\|_2} f(x^k)\right];$$

Step 3: If $x^{k+1} = x^k$, then terminate. Otherwise, set $k = k + 1$ and return to step 1.

It is obvious that $\{x^k\}$ is a solution of (1.1) provided $x^{k+1} = x^k$ for some k. So in the sequel we assume that $\{x^k\}$ is an infinite sequence.

We carefully check the propositions and their proofs in [6] and find that under above conditions all lemmas and their proofs remain valid except the proof of Theorem 2, in which $f(x)$ was assumed to be strongly monotone and the generated sequence $\{x^k\}$ was proved to converge to the unique solution of (1.1). In this section we will show that the assertion of Theorem 2 in [6] is still right under our assumptions. To this end we need recall some lemmas in [6]. In the rest part of this section we assume that all above conditions hold.

Lemma 3.1 (Lemma 3 in [6]). *The sequence $\{x^k\}$ generated by Algorithm* 2 *is bounded.*

Lemma 3.2 (Lemma 4 in [6]).

$$\lim_{k\to\infty} \operatorname{dist}[x^k, C] = 0.$$

Lemma 3.3 (Lemma 5 in [6]).

$$\lim_{k\to\infty} \|x^{k+1} - x^k\|_2 = 0.$$

Based above lemmas we may establish the convergence of Algorithm 2 under our assumptions.

Theorem 3.1. *Assume that* (A_2)–(F_2) *are satisfied. Then the sequence* $\{x^k\}$ *generated by Algorithm* 2 *converges to a solution of* (1.1). *Furthermore, if* (1.1) *has no solution at which* $f(x)$ *vanishes, then the solution of* (1.1) *is unique and* $\{x^k\}$ *converges to this solution.*

Proof. From Theorem 2.3 we see that (1.1) has a solution. Let x^* be a solution of (VI). By Lemma 3.2, we have $x^k \in C_\delta$ for all k sufficiently large, where C_δ is the set given in conditions (A_2) and (B_2). From (B_2) we have

$$\begin{aligned}\left\langle f\left(x^k\right), x^{k+1} - x^*\right\rangle \geqslant{}& g\left(x^k, x^*\right)\left\|\left(f\left(x^k\right) - f(x^*)\right)\right\|^2 + \left\langle f(x^*), x^k - x^*\right\rangle \\ &+ \left\langle f\left(x^k\right), x^{k+1} - x^k\right\rangle \end{aligned} \tag{3.1}$$

for sufficiently large k. Since $g(x, y)$ is a positive continuous function on C_δ and $\{x^k\}$ is bounded, then there exists an $\eta > 0$ such that $g(x^k, x^*) \geqslant \eta$ for sufficiently large k.

Let ε be an arbitrary positive number. Since x^* is a solution of (1.1). From Lemmas 3.1 and 3.2 it follows

$$\left\langle f(x^*), x^k - x^*\right\rangle \geqslant -\varepsilon \tag{3.2}$$

for all sufficiently large k.

Moreover since $f(x)$ is continuous on C_δ, from Lemmas 3.1, 3.2, and 3.3 we obtain

$$\left\langle f\left(x^k\right), x^{k+1} - x^k\right\rangle \geqslant -\left\| f\left(x^k\right)\right\|_2 \left\|x^{k+1} - x^k\right\|_2 \geqslant -\varepsilon \tag{3.3}$$

for all k large enough. Consequently we derive following inequality from (3.1), (3.2) and (3.3)

$$\left\langle f\left(x^k\right), x^{k+1} - x^*\right\rangle \geqslant \eta \left\| f\left(x^k\right) - f(x^*)\right\|_2^2 - 2\varepsilon \tag{3.4}$$

for all k large enough.

If $\inf\{\|f(x^k) - f(x^*)\|_2\} > 0$, then there exists a $\gamma > 0$, such that

$$\left\| f\left(x^k\right) - f(x^*)\right\| \geqslant \gamma \quad \text{for all } k,$$

then it follows from (3.4)

$$\left\langle f\left(x^k\right), x^{k+1} - x^*\right\rangle \geqslant \eta\gamma - 2\varepsilon \tag{3.5}$$

for all sufficiently large k.

Since $C \subseteq C^k$ for all k and $x^* \in C$, then $x^* \in P_{C^k}(x^*)$.

Since $\{f(x^k)\}$ is bounded, let $\|f(x^k)\|_2 \leqslant M$ for all k. Then we obtain from Lemma 2.2

$$\begin{aligned}\left\|x^{k+1} - x^*\right\|_2^2 &= \left\| P_{C^k}\left[x^k - \frac{\rho_k}{\|f(x^k)\|_2} f\left(x^k\right)\right] - P_{C^k}(x^*)\right\|_2^2 \\ &\leqslant \left\| x^k - \frac{\rho_k}{\|f(x^k)\|_2} f\left(x^k\right) - x^*\right\|_2^2 \\ &\quad - \left\| x^{k+1} - \left(x^k - \frac{\rho_k}{\|f(x^k)\|_2} f\left(x^k\right)\right)\right\|_2^2\end{aligned}$$

$$\leqslant \|x^k - x^*\|_2^2 - 2\frac{\rho_k}{\|f(x^k)\|_2}\langle f(x^k), x^{k+1} - x^*\rangle$$
$$\leqslant \|x^k - x^*\|_2^2 - 2\rho_k\frac{(\eta\gamma - 2\varepsilon)}{M} = \|x^k - x^*\|_2^2 - \frac{\eta\gamma}{M}\rho_k \tag{3.6}$$

for all sufficiently large k, and last equality follows with taking $\varepsilon = \eta\gamma/4$.

Let k_0 be a large positive constant and (3.5) holds for $k \geqslant k_0$. By adding the inequalities (3.6) from k_0 to $k_0 + l$ we get

$$\|x^{k_0+l+1} - x^*\|_2^2 \leqslant \|x^{k_0} - x^*\|_2^2 - \frac{\eta\gamma}{M}\sum_{i=k_0}^{k_0+l}\rho_i$$

for any $l > 0$. However it is impossible since $\sum_{k=k_0}^{\infty}\rho_k = +\infty$. So we conclude that $\inf\{\|f(x^k) - f(x^*)\|_2\} = 0$. Since $\{x^k\}$ is bounded and $\|x^{k+1} - x^k\| \to 0$ as $k \to +\infty$ by Lemma 3.3. Then $x^k \to \bar{x}$ as $k \to +\infty$. Therefore $f(\bar{x}) = f(x^*)$. This implies that $\bar{x}$ is a solution of (1.1) from condition (F_2).

Furthermore, if (1.1) has no solution at which $f(x)$ vanishes, then from (B_2) and (D_2) we conclude that the solution of (1.1) is unique and so $\{x^k\}$ converges to this solution.

This completes the proof. □

Theorem 3.2. *Let the set C be bounded. Then* (1.1) *has a solution and the conclusions of Theorem* 3.1 *remain valid.*

Theorem 3.3. *If the set C is unbounded. Then* (1.1) *exists a solution and the conclusions of Theorem* 3.1 *remain valid.*

Remark. It is easily verified that for Example 1 all conditions in this section are satisfied. Then the sequence $\{x^k\}$ by Algorithm 2 is convergent. However $f(x) = 1 - e^{-x}$ is not strongly monotone on C_δ. Therefore our result shows that Fukushima's algorithm applies in a broader scope.

4. An application

In this section we apply Algorithms 1 and 2 to the Split Feasibility Problem (SFP).

Let C and Q be the closed convex subsets in R^n and R^m, respectively. The SFP is to find $x \in C$ with $Ax \in Q$ if such x exist, where A is a real m by n matrix [8–11]. A number of image reconstruction problems can be formulated as SFP. In [11], Censor and Elfving used their multidistance method to obtain iterative algorithms for solving the SFP. Their algorithms, as well as others obtained later (see, e.g., [10]) involve matrix inverses at each step. In [8,9], Byrne proposed an iterative method, called the CQ algorithm, to solve the SFP. The feature of the CQ algorithm is that matrix inverses are not involved. The CQ algorithm solving the SFP may be restated as follows [8,9]):

Let x^0 be arbitrary. For $k = 0, 1, \ldots,$ let

$$x^{k+1} = P_C[x^k - \gamma A^T(I - P_Q)Ax^k], \tag{4.1}$$

where $\gamma \in (0, 2/\rho(A^T A)$ and $\rho(A^T A)$ denotes the spectral radius of $A^T A$.

It is known that x is a solution of following variational inequality problem: to find $x \in C$ such that

$$\left\langle A^T(I - P_Q)Ax, y - x\right\rangle \geqslant 0 \quad \text{for all } y \in C, \tag{4.2}$$

if and only if x is a solution of the SFP whenever the SFP has a solution (see [9]).

By Lemma 2.1, (4.2) may be equivalently written as

$$x = P_C\left[x - \gamma A^T(I - P_Q)Ax\right]. \tag{4.3}$$

Furthermore, it has been known from [9, Lemma 8.1] that $f(x) = A^T(I - P_Q)Ax$ is co-coercive in R^n with constant $1/\rho(A^T A)$, i.e.,

$$\left\langle f(x) - f(y), x - y\right\rangle \geqslant \frac{1}{\rho(A^T A)}\left\| f(x) - f(y)\right\|_2^2.$$

Therefore the CQ algorithm is really a particular case of (2.2). Hence the convergence of the CQ algorithm immediately follows from [1, Theorem 12.1.8]. In [9] Byrne proposed a different proof based on Dolidze's Theorem [9,13] for his CQ algorithm. However it is not trivial work to gain the $\rho(A^T A)$. For practical reasons a quick method to estimate $\rho(A^T A)$ was given by Byrne [8,9]. However Byrne's method is only available for A being ε-sparse. For general A, it may cost a large of amount of computation to obtain $\rho(A^T A)$. Therefore it is meaning to present an improvement of the CQ algorithm without the requirement of knowledge of $\rho(A^T A)$.

Since (4.1) is a special case of (2.2), it is natural to apply Algorithm 1 or Algorithm 2 to solve the SFP. Of course, to guarantee the convergence, it is essential to check whether the conditions in Section 2 or Section 3 are satisfied. Obviously $f(x) = A^T(I - P_Q)Ax$ is continuous in R^n and $f(x) = 0$ whenever x is a solution of the SFP. Hence it is only necessary to verify if (C_1) or (C_2) holds for the convergence of Algorithm 1 or Algorithm 2 with $f(x) = A^T(I - P_Q)Ax$.

Lemma 4.1. *Assume that Q is bounded and A is an m by n matrix with full column rank. Then for any given $z \in C$, there exists a bounded set $D \subseteq R^n$ such that*

$$\left\langle A^T(I - P_Q)Ax, x - z\right\rangle \geqslant \left\|A^T(I - P_Q)Ax\right\|_2 \quad \textit{for all } x \notin D.$$

Proof. First we have

$$\begin{aligned}
&\left\langle A^T(I - P_Q)Ax, x - z\right\rangle \\
&\quad = \left\langle (I - P_Q)Ax, Ax - Az\right\rangle \\
&\quad = x^T A^T Ax - \langle Ax, Az\rangle - \left\langle P_Q(Ax), Ax\right\rangle + \left\langle P_Q(Ax), Az\right\rangle.
\end{aligned}$$

Since A has full column rank, then $A^T A$ is positive definite. Note that Q is bounded, hence there is a big closed ball $D = \{x \in R^n \mid \|x\|_2 \leqslant d\}$ such that

$$\begin{aligned}
&x^T A^T Ax - \langle Ax, Az\rangle - \left\langle P_Q(Ax), Ax\right\rangle + \left\langle P_Q(Ax), Az\right\rangle \\
&\quad \geqslant \left\|A^T Ax\right\|_2 + \left\|A^T P_Q(Ax)\right\|_2 \\
&\quad \geqslant \left\|A^T(I - P_Q)Ax\right\|_2 \quad \text{for all } x \notin D
\end{aligned}$$

where d is sufficiently large positive number.

That is

$$\langle A^T(I-P_Q)Ax, x-z\rangle \geqslant \|A^T(I-P_Q)Ax\|_2 \quad \text{for all } x \notin D. \qquad \square$$

Therefore we immediately get following theorem from Lemma 4.1 and Theorem 2.1.

Theorem 4.1. *Assume Q is bounded and A is an m by n matrix with full column rank. Then the sequence $\{x^k\}$ generated by Algorithm* 1 *with $f(x) = A^T(I-P_Q)Ax$ converges to a solution of the SFP whenever the SFP has a solution.*

When apply Algorithm 2 to solve the SFP, one needs note that mapping $f(x)$ itself contains an exact orthogonal projection P_Q. If we suppose that P_C can be solved without great expense, then we have following theorem from Lemma 4.1 and Theorem 3.1.

Theorem 4.2. *Assume Q is bounded and A is an m by n matrix with full column rank, conditions (E_2), (F_2) hold. Then the sequence $\{x^k\}$ generated by Algorithm* 2 *converges to a solution of the SFP provided the SFP has a solution.*

In [22], the relaxed version of the CQ algorithm is given, where the projections P_C and P_Q are replaced by P_{C^k} and P_{Q^k} respectively at kth iteration, C^k and Q^k are the halfspaces associated with C and Q respectively. The convergence of the relaxed CQ algorithm is established under mild assumptions. When applying Algorithm 2 to solve the SFP with using the Q^k in place of Q in $f(x) = A^T(I-P_Q)Ax$, whether or not the convergence still holds is a subject deserving research.

References

[1] F. Facchinei, J.S. Pang, Finite-Dimensional Variational Inequality and Complementarity Problems, vols. 1 and 2, Springer-Verlag, New York, 2003.

[2] P. Harker, J.S. Pang, Finite-dimensional variational inequality and nonlinear complementarity problems: a survey of theory, algorithms and applications, Math. Program. 48 (1990) 161–220.

[3] J.S. Pang, Inexact Newton methods for the nonlinear complementarity problems, Math. Program. 36 (1986) 54–71.

[4] A. Auslender, Optimisation: Méthodes Numérique, Masson, Paris, 1976.

[5] M. Fukushima, An outer approximation algorithm for solving general convex programs, Oper. Res. 31 (1983) 101–113.

[6] M. Fukushima, A relaxed projection method for variational inequalities, Math. Program. 35 (1986) 58–70.

[7] D. Kinderlehrer, G. Stampacchia, An Introduction to Variational Inequalities and Their Applications, Academic Press, New York, 1980.

[8] C. Byrne, Iterative oblique projection onto convex sets and the split feasibilities problem, Inverse Problems 18 (2002) 441–453.

[9] C. Byrne, A unified treatment of some iterative algorithms in signal processing and image reconstruction, Inverse Problems 20 (2004) 103–120.

[10] C. Byrne, Bregman–Legendre multidistances projection algorithm for convex feasibility and optimization, in: D. Butnariu, Y. Censor, S. Reich (Eds.), Inherently Parallel Algorithms in Feasility and Optimization and their Applications, Elsevier, Amsterdam, 2001, pp. 87–100.

[11] Y. Censor, T. Elfving, A multiprojection algorithm using Bregman projections in a production space, Numer. Algorithms 8 (1994) 221–239.

[12] Y. Censor, A. Iusem, S. Zenios, An interior point method with Bregman functions for the variational inequality problem with paramonotone operators, Math. Program. 81 (1998) 370–400.

[13] Z. Dolidze, Solution of variational inequalities associated with a class of monotone maps, Ekonomik i Matem. Metody 18 (1982) 925–927.

[14] B.S. He, Inexact implicit methods for monotone general variational inequalities, Math. Progam. Ser. A 86 (1999) 199–217.

[15] M. Solodov, P. Tseng, Modified projection-type methods for monotone variational inequalities, SIAM J. Control Optim. 34 (1996) 1814–1830.

[16] D. Sun, A class of iterative methods for solving nonlinear projection equations, J. Optim. Theory Appl. (1996) 123–140.

[17] P. Tseng, Alternating projection-proximal methods for variational inequalities, SIAM J. Optim. 7 (1997) 951–965.

[18] N. Xiu, J. Zhang, Some recent advances in projection-type methods for variational inequalities, J. Comput. Appl. Math. 152 (2003) 559–587.

[19] M. Noor, Y. Wang, N. Xiu, Some projection methods for variational inequalities, Appl. Math. Comput. 137 (2003) 423–435.

[20] P. Marcotte, J. Wu, On the convergence of projection methods. Applications to the decomposition of affine variational inequalities, J. Optim. Theory Appl. 85 (1995) 347–362.

[21] E. Golshtein, N. Tretyakov, Modified Lagrangians and Monotone Maps in Optimization, Wiley, New York, 1996.

[22] Q. Yang, The relaxed CQ algorithm solving the Split Feasibility Problem, Inverse Problems 20 (2004) 1261–1266.