**Nonlinear Funct. Anal. & Appl., Vol. 11, No. 3 (2006), pp. 495–506**

# AN APPROXIMATE NONLINEAR PROXIMAL PREDICTION-CORRECTION ALGORITHM FOR MAXIMAL MONOTONE OPERATORS

Min Li

Abstract. This paper introduces an approximate nonlinear proximal prediction-correction algorithm for finding the zero points of $\hat{T}(\cdot)$, where $\hat{T}$ is a maximal monotone operator. In the prediction step, the presented algorithm allows for constant relative error tolerance which should be easier to verify and enforce in practice than those given in earlier analysis of approximate generalized proximal point algorithms. In the correction step, to make more progress, a general decent direction and a suitable step length are used. And the global convergence is easily established under weaker conditions on the algorithm parameters. As for applications, we give two methods, one is for solving monotone variational inequalities and the other is for the choice of the decent direction in the correction step.

## 1. Introduction

$\hat{T}$ is called a monotone operator on $R^n$, i.e., it has the property

$$(x, y), \quad (x', y') \in \hat{T} \ \Rightarrow \ (x - x')^T (y - y') \geq 0.$$

$\hat{T}$ is maximal if it is not properly contained in any other set with the above property. We define $\hat{T}(x) = \{y \mid (x, y) \in \hat{T}\}$, since it is natural to view $\hat{T}$ as the graph of point-to-set mapping. We also define the sum $T + Q$ of two operators $T$ and $Q$ on $R^n$ by

$$T + Q = \{(x, y + z) \mid (x, y) \in T, \ (x, z) \in Q\}.$$

Received October 28, 2005. Revised May 17, 2006.

2000 Mathematics Subject Classification: 47J20, 65K10, 90C25.

Key words and phrases: Maximal monotone operator, prediction-correction method, monotone variational inequalities.

In this paper we consider $\hat{T}$, where $\hat{T} = T + N_\Omega$ (See [7], p117), $T$ is a maximal monotone operator on $R^n$, and $N_\Omega$ is the normal cone operator with respect to $\Omega$, that is

$$N_\Omega = \{(x, y) \mid x \in \Omega,\ y^T(z - x) \le 0 \quad \forall z \in \Omega\}, \tag{1.1}$$

where $\Omega$ is closed and convex. Note that $N_\Omega(x) = \{y \mid (x, y) \in N_\Omega\}$ is a cone and hence $\beta N_\Omega(x) = N_\Omega(x)$ for all $x \in \Omega$ and $\beta > 0$. Obviously $N_\Omega$ is a maximal monotone operator. The main problem of this paper is to find one of the zero points of $\hat{T}(\cdot)$, defined to be points $x^* \in \Omega$ such that $0 \in \hat{T}(x^*)$. This kind of issue often appears in a wide variety of equilibrium problems such as convex programming and monotone variational inequalities. First, we make the following assumptions:

**Assumption 1.1.** *$(int\ \Omega) \cap (relint\ dom\ T) \neq \emptyset$. Here, "int $X$" denotes the interior of the set $X$, "relint $X$" the relative interior of the set $X$, and "dom $T$" the domain of the operator $T$, where dom $T = \{x \mid (x, y) \in T\} = \{x \mid T(x) \neq \emptyset\}$.*

**Remark 1.1.** We can easily prove that $\hat{T}$ ($\hat{T} = T + N_\Omega$) is maximal and monotone since $T$ and $N_\Omega$ are both maximal monotone and satisfy Assumption 1.1.

For a Bregman function $h$, the approximate Bregman-function-based proximal algorithm [7] generates a pair of sequences $\{x^k\}$, $\{e^k\}$ conforming to the recursion:

$$\nabla h(x^k) + e^k \in \nabla h(x^{k+1}) + \beta_k \hat{T}(x^{k+1}), \tag{1.2}$$

where $\{\beta_k\}$ is a sequence of positive scalars and $\{e^k\}$ satisfies the conditions:

$$\sum_{k=1}^{\infty} \|e^k\| < \infty, \tag{1.3}$$

$$\sum_{k=1}^{\infty} \langle e^k, x^k \rangle \text{ exists and is finite.} \tag{1.4}$$

The first aim of this paper is to extend the above method. We use (1.2) to develop the predictor. On the other hand, from a practical point of view, it is interesting to improve the error sequence considered in [7]. Such summability conditions on the errors are rather restrictive and somewhat undesirable because they impose increasing precision along the iterative process. In this paper, a fixed relative error tolerance is permitted which is particular appropriate for computer implementation and does not affect the global convergence of the algorithm.

The second goal of this paper is to accelerate the method. We make a correction, in which a general decent direction and a suitable step length are used. This strategy may help to make more progress in each iteration.

The framework of the proposed method is that for given $x^k$ and $\beta_k$ ($\inf \beta_k = \beta > 0$), let $\tilde{x}^k$ and $e^k$ satisfy the following set-valued equation:

$$\nabla h(x^k) + e^k \in \nabla h(\tilde{x}^k) + \beta_k \hat{T}(\tilde{x}^k), \tag{1.5}$$

where $\{e^k\}$ is regarded as an error sequence and usually obeys

$$\|e^k\| \leq \mu\eta_k \|x^k - \tilde{x}^k\|, \qquad \sup_k \eta_k = \eta < 1. \tag{1.6}$$

From the definition of $\hat{T}$ ($\hat{T} = T + N_\Omega$), we can find a $y^k \in T(\tilde{x}^k)$, such that

$$\nabla h(\tilde{x}^k) + \beta_k y^k - \nabla h(x^k) - e^k \in -N_\Omega(\tilde{x}^k). \tag{1.7}$$

We use the notation

$$d^k = \nabla h(x^k) - \nabla h(\tilde{x}^k) + e^k, \tag{1.8}$$

and then a new iteration $x^{k+1}$ satisfies

$$\nabla h(x^{k+1}) - \nabla h(x^k) + \alpha_k \beta_k y^k \in -N_\Omega(x^{k+1}), \tag{1.9}$$

where

$$\alpha_k = \gamma_k \alpha_k^*, \quad \gamma_k \in [\gamma_{_L}, \gamma_{_U}] \subset [1, 2) \quad \text{and} \quad \alpha_k^* = \frac{(x^k - \tilde{x}^k)^T d^k}{\|d^k\|^2}. \tag{1.10}$$

For convenience, we make some additional assumptions to guarantee that the problem under consideration is solvable and the algorithm is well defined.

**Assumption 1.2.** *Suppose $h$ satisfies the following conditions:*

(i) *$h$ is a twice differentiable real-valued function on $\Omega_0$, where $\Omega_0$ is an open set and $\Omega_0 \supseteq \Omega$.*

(ii) $\nabla h$ *is uniformly strong monotone on* $\Omega$*, i.e., there exists a constant* $\mu > 0$*, such that*

$$(\nabla h(x) - \nabla h(y))^T (x - y) \geq \mu \|x - y\|^2, \qquad \forall\ x, y \in \Omega. \tag{1.11}$$

*And we also assume that we have already known the value of* $\mu$*.*

(iii) $\nabla h$ *is Lipschitz continuous on* $\Omega$*, i.e., there exists a constant* $L > 0$*, such that*

$$\|\nabla h(x) - \nabla h(y)\| \leq L \|x - y\|, \qquad \forall\ x, y \in \Omega. \tag{1.12}$$

(iv) $h$ *is strongly convex with positive constant* $\nu$*, i.e.,*

$$\nabla^2 h(x) \succeq \nu I, \qquad \forall\ x \in \Omega.$$

*In the proposed method, we usually assume* $\nu = 1$ *for reasons of simplicity of the presentation.*

(v) *If* $\{y^k\} \in \Omega$ *is a convergent sequence with limit* $y^*$*, then* $D_h(y^*, y^k) \to 0$*. (*$D_h(x, y)$ *will be defined in* (1.13)*.)*

**Remark 1.2.** (ii) can be satisfied from (iv). The conditions on $h$ are similar to those of Bregman functions [5], but are stronger.

In the following, we give three examples to explain that we can easily find such $h$ to satisfy Assumption 1.2.

**Example 1.** *Let* $h(x) = \frac{1}{2}\|x\|_Q^2$*, where* $Q$ *is a symmetry positive definite matrix and* $\Omega$ *is a closed convex set.*

**Example 2.** *Let* $\Omega$ *be a bounded and closed convex subset of* $R^n_{++}$*, and set*

$$h(x) = \frac{1}{q} \sum_{j=1}^{n} x_j^q,$$

*where* $q > 1$*. Note that* $R^n_{++}$ *is a strict positive orthants, that is,*

$$R^n_{++} = \{x \in R^n \mid x > 0\}.$$

**Example 3.** *Let* $\Omega$ *be defined as in Example* 2 *and*

$$h(x) = \sum_{j=1}^{n} x_j \log x_j - x_j.$$

**Remark 1.3.** If $h$ satisfies Assumption 1.2, so does $h(x) + a^T x + b$, for any $a \in R^n$ and $b \in R$.

**Assumption 1.3.** *The sequences* $\{\tilde{x}^k\} \subset \Omega$ *and* $\{e^k\} \subset R^n$ *conforming to the recursion* (1.5) *exist.*

**Remark 1.4.** The above assumption can be satisfied by some simple conditions on $\nabla h$ such as in [3-5, 7]. If $h(x) = \frac{1}{2}\|x\|^2$, the existence of $\{\tilde{x}^k\}$ and $\{e^k\}$ is assured in Theorem 2 of [6]. Also for monotone variational inequalities, i.e., $T = F$ where $F$ is a continuous monotone mapping from $R^n$ into itself, it is easy to prove that Assumption 1.3 is satisfied without any extra conditions.

**Assumption 1.4.** *The zero points set of* $\hat{T}(\cdot)$, *denoted by* $\Omega^*$, *is nonempty.*

## 2. Preliminaries

For convenience, we consider the problems only under the Euclidean norm. We use the notation $P_\Omega : R^n \to R^n$ to denote any function for which $P_\Omega(y)$ is a projection of $y$ on $\Omega$, i.e., for all $y \in R^n$,

$$P_\Omega(y) = \operatorname{argmin}\{\|z - y\| \mid z \in \Omega\},$$

where $P_\Omega$ is a projection on $\Omega$.

**Lemma 2.1.** ([1], p267) *For any* $\beta > 0$, $y \in -N_\Omega(x)$ *if and only if*

$$x = P_\Omega[x - \beta y].$$

From $h$, one may obtain a kind of distance measure or "D-functions" $D_h$ by the construction [2]

$$D_h(x, y) = h(x) - h(y) - \nabla h(y)^T (x - y). \tag{2.1}$$

$D_h(x, y)$ may be interpreted as the difference between $h(x)$ and the value at $x$ of a linearized approximation of $h$ around $y$. By the strict convexity of $h$, $D_h$ is nonnegative and $D_h(x, y) = 0$ if and only if $x = y$. For example, if $h(x) = \frac{1}{2}\|x\|^2$, then $D_h(x, y) = \frac{1}{2}\|x - y\|^2$. From Assumption 1.2 (iv), we obtain

$$\begin{aligned} D_h(x, y) &= h(x) - h(y) - \nabla h(y)^T (x - y) \\ &= \frac{1}{2}(x - y)^T \nabla^2 h(z)(x - y) \\ &\geq \frac{1}{2}\|x - y\|^2, \end{aligned} \tag{2.2}$$

for some $z$ on the line segment $[x, y]$.

It follows from the notation (1.8), Assumption 1.2 and (1.6) that

$$\begin{aligned}(x^k - \tilde{x}^k)^T d^k &= (x^k - \tilde{x}^k)^T[\nabla h(x^k) - \nabla h(\tilde{x}^k)] + (x^k - \tilde{x}^k)^T e^k \\ &\geq \mu\|x^k - \tilde{x}^k\|^2 - \|x^k - \tilde{x}^k\|\|e^k\| \\ &\geq \mu(1-\eta)\|x^k - \tilde{x}^k\|^2,\end{aligned} \tag{2.3}$$

and

$$\begin{aligned}\|d^k\| &\leq \|\nabla h(x^k) - \nabla h(\tilde{x}^k)\| + \|e^k\| \\ &\leq (L + \mu\eta)\|x^k - \tilde{x}^k\|.\end{aligned} \tag{2.4}$$

To choose a better step-size $\alpha$, it is natural to let the new iterate $x^{k+1}$ satisfy:

$$\nabla h(x_\alpha^{k+1}) - \nabla h(x^k) + \alpha\beta_k y^k \in -N_\Omega(x_\alpha^{k+1}). \tag{2.5}$$

From Lemma 2.1, we know that

$$x_\alpha^{k+1} = P_\Omega\{x_\alpha^{k+1} - [\nabla h(x_\alpha^{k+1}) - \nabla h(x^k) + \alpha\beta_k y^k]\}, \tag{2.6}$$

where $y^k \in T(\tilde{x}^k)$.

## 3. Convergence

Now let us observe the difference between $D_h(x^*, x^k)$ and $D_h(x^*, x_\alpha^{k+1})$. Denote

$$\Theta_k(\alpha) := D_h(x^*, x^k) - D_h(x^*, x_\alpha^{k+1}). \tag{3.1}$$

We call $\Theta_k(\alpha)$ a profit function because it measures the progress obtained in the k+1-th iteration.

To achieve rapid convergence, some suitable step length should be carefully chosen. The following theorem tells us how to choose $\alpha$.

**Theorem 3.1.** *For given $x^k \in R^n$ and $\beta_k > 0$, let $\tilde{x}^k$ and $e^k$ conform to set-valued equation* (1.5) *and condition* (1.6)*. Then the new iterate $x_\alpha^{k+1}$ given by* (2.6) *for any $\alpha > 0$ satisfies*

$$\Theta_k(\alpha) \geq \alpha(x^k - \tilde{x}^k)^T d^k - \frac{1}{2}\alpha^2\|d^k\|^2, \tag{3.2}$$

*where $d^k$ is defined in* (1.8)*.*

*Proof.* By direct algebra,

$$\begin{aligned} & D_h(x^*, x^k) - D_h(x^*, x_\alpha^{k+1}) \\ &= D_h(x_\alpha^{k+1}, x^k) + (x^* - x_\alpha^{k+1})^T[\nabla h(x_\alpha^{k+1}) - \nabla h(x^k)]. \end{aligned} \tag{3.3}$$

Substituting this and (2.2) into (3.1), we obtain

$$\Theta_k(\alpha) \geq \frac{1}{2}\|x_\alpha^{k+1} - x^k\|^2 + (x^* - x_\alpha^{k+1})^T[\nabla h(x_\alpha^{k+1}) - \nabla h(x^k)]. \tag{3.4}$$

From the definition of $N_\Omega$, (2.5) is equivalent to

$$x_\alpha^{k+1} \in \Omega \qquad (x' - x_\alpha^{k+1})^T[\nabla h(x_\alpha^{k+1}) - \nabla h(x^k) + \alpha\beta_k y^k] \geq 0, \qquad \forall\ x' \in \Omega.$$

Let $x' = x^*$. Then we have

$$(x^* - x_\alpha^{k+1})^T[\nabla h(x_\alpha^{k+1}) - \nabla h(x^k)] \geq \alpha\beta_k(x_\alpha^{k+1} - x^*)^T y^k.$$

Substituting this into (3.4), we get

$$\Theta_k(\alpha) \geq \frac{1}{2}\|x_\alpha^{k+1} - x^k\|^2 + \alpha\beta_k(x_\alpha^{k+1} - x^*)^T y^k. \tag{3.5}$$

Since $x^*$ is a zero point of $\hat{T}$, $y^k \in \hat{T}\ (\tilde{x}^k)$ $(y^k \in T(\tilde{x}^k), 0 \in N_\Omega(\tilde{x}^k),$ $\hat{T} = T + N_\Omega)$ and $\hat{T}$ is monotone, we have

$$(\tilde{x}^k - x^*)^T(y^k - 0) \geq 0,$$

that is,

$$(x_\alpha^{k+1} - x^*)^T y^k \geq (x_\alpha^{k+1} - \tilde{x}^k)^T y^k. \tag{3.6}$$

Then from (3.5), (3.6) and the notation (1.8), we obtain

$$\begin{aligned} \Theta_k(\alpha) \geq & -\frac{1}{2}\alpha^2\|d^k\|^2 + \frac{1}{2}\|(x_\alpha^{k+1} - x^k) + \alpha d^k\|^2 \\ & + \alpha\beta_k(x_\alpha^{k+1} - \tilde{x}^k)^T y^k + \alpha(x^k - x_\alpha^{k+1})^T d^k. \end{aligned} \tag{3.7}$$

Now we consider the last term in the right-hand-side of (3.7). (1.1) and (1.9) yield

$$\tilde{x}^k \in \Omega, \qquad (x' - \tilde{x}^k)^T[\nabla h(\tilde{x}^k) + \beta_k y^k - \nabla h(x^k) - e^k] \geq 0, \qquad \forall\ x' \in \Omega.$$

By using $x' = x_\alpha^{k+1}$ and the notation (1.8) in the above inequality, we get

$$(x_\alpha^{k+1} - \tilde{x}^k)^T (d^k - \beta_k y^k) \le 0,$$

that is

$$(x^k - x_\alpha^{k+1})^T (d^k - \beta_k y^k) \ge (x^k - \tilde{x}^k)^T (d^k - \beta_k y^k),$$

and thus

$$(x^k - x_\alpha^{k+1})^T d^k \ge (x^k - \tilde{x}^k)^T d^k - \beta_k (x_\alpha^{k+1} - \tilde{x}^k)^T y^k.$$

Substituting this into (3.7), we obtain

$$\Theta_k(\alpha) \ge -\frac{1}{2}\alpha^2 \|d^k\|^2 + \alpha (x^k - \tilde{x}^k)^T d^k.$$

The assertion is proved. □

According to Theorem 3.1, it is natural to choose step-size as

$$\alpha_k^* = \frac{(x^k - \tilde{x}^k)^T d^k}{\|d^k\|^2}.$$

Using (2.3) and (2.4), we have

$$\begin{aligned} \alpha_k^* &\ge \frac{\mu(1-\eta)\|x^k - \tilde{x}^k\|^2}{(L+\mu\eta)^2 \|x^k - \tilde{x}^k\|^2} \\ &= \frac{\mu(1-\eta)}{(L+\mu\eta)^2}. \end{aligned} \tag{3.8}$$

For fast convergence, we get the step-size $\alpha_k$ by multiplying $\alpha_k^*$ by a relaxation factor $\gamma_k$ as in the SOR methods, i.e.,

$$\alpha_k = \gamma_k \alpha_k^*, \qquad \gamma_k \in [\gamma_L, \gamma_U] \subset [1, 2).$$

**Theorem 3.2.** *Let $\{x^k\}$, $\{\tilde{x}^k\}$, $\{e^k\}$ and $\{\eta_k\}$ be the sequences conforming to the set-valued equation* (1.5) *and condition* (1.6). *Then the sequence $\{x^k\}$ generated by* (1.9) *satisfies*

$$D_h(x^*, x^{k+1}) \le D_h(x^*, x^k) - \frac{\mu^2(1-\eta)^2 \gamma_L (2-\gamma_U)}{2(L+\mu\eta)^2} \|x^k - \tilde{x}^k\|^2. \tag{3.9}$$

*Proof.* From (3.2) and (1.10), we obtain

$$\begin{aligned}\Theta_k(\gamma_k\alpha_k^*) &\geq \gamma_k\alpha_k^*(x^k-\tilde{x}^k)^T d^k - \frac{1}{2}(\gamma_k\alpha_k^*)^2\|d^k\|^2\\ &= \gamma_k\alpha_k^*(x^k-\tilde{x}^k)^T d^k - \frac{1}{2}(\gamma_k^2\alpha_k^*)(x^k-\tilde{x}^k)^T d^k\\ &\geq \frac{1}{2}\alpha_k^*\gamma_L(2-\gamma_U)(x^k-\tilde{x}^k)^T d^k.\end{aligned}$$

And then it follows from (3.1) and (2.3) that

$$\begin{aligned}D_h(x^*,x^{k+1}) &\leq D_h(x^*,x^k) - \frac{1}{2}\alpha_k^*\gamma_L(2-\gamma_U)(x^k-\tilde{x}^k)^T d^k\\ &\leq D_h(x^*,x^k) - \frac{1}{2}\alpha_k^*\gamma_L(2-\gamma_U)\mu(1-\eta)\|x^k-\tilde{x}^k\|^2\end{aligned}$$

The assertion follows from (3.8) immediately. □

From Inequality (2.2) and (3.9), we can see that

$$\frac{1}{2}\|x^*-x^k\|^2 \leq D_h(x^*,x^k) \leq D_h(x^*,x^0) \text{ and } \lim_{k\to\infty}\|x^k-\tilde{x}^k\| = 0. \quad (3.10)$$

Consequently, $\{x^k\}$ and $\{\tilde{x}^k\}$ are both bounded. In the following, we will prove the convergence of the approximate nonlinear proximal prediction-correction algorithm.

**Theorem 3.3.** *Let $\{x^k\}$, $\{\tilde{x}^k\}$, $\{e^k\}$ and $\{\eta_k\}$ be the sequences conforming to the set-valued equation* (1.5) *and condition* (1.6)*. Then the sequence $\{x^k\}$ generated by* (1.9) *converges to one zero point of $\hat{T}(\cdot)$.*

*Proof.* From Theorem 3.2, we know that $\{\tilde{x}^k\}$ is bounded. Therefore it has one cluster point at least. Let $x^\infty$ be a cluster point of $\{\tilde{x}^k\}$ and the subsequences $\{\tilde{x}^{k_j}\}$ converges to $x^\infty$. Define

$$z^k = \frac{1}{\beta_k}[\nabla h(x^k) - \nabla h(\tilde{x}^k) + e^k],$$

then $z^{k_j} \in \hat{T}(\tilde{x}^{k_j})$(according to (1.5)). Using $\lim_{k\to\infty}\|x^k-\tilde{x}^k\| = 0$, $e^k \to 0$ (see (1.6)), $\inf_k \beta_k = \beta > 0$ and $\nabla h(\cdot)$ is Lipschitz continuous, we obtain

$$\lim_{j\to\infty} z^{k_j} = \lim_{j\to\infty}\frac{1}{\beta_{k_j}}[\nabla h(x^{k_j}) - \nabla h(\tilde{x}^{k_j}) + e^{k_j}] = 0.$$

Because $\hat{T}$ is maximal, it is a closed set in $R^n \times R^n$, then

$$\lim_{j\to\infty} (\tilde{x}^{k_j}, z^{k_j}) = (x^\infty, 0) \in \hat{T},$$

and $x^\infty$ is a zero point of $\hat{T}(\cdot)$. Note that Inequality (3.9) is true for all zero points of $\hat{T}(\cdot)$, hence we have

$$D_h(x^\infty, x^{k+1}) \le D_h(x^\infty, x^k), \qquad \forall\ k \ge 0. \tag{3.11}$$

Since $\{\tilde{x}^{k_j}\} \to x^\infty$ and $x^k - \tilde{x}^k \to 0$, $\{x^{k_j}\} \to x^\infty$. From Assumption 1.2 (v), we know that $D_h(x^\infty, x^{k_j}) \to 0$, i.e., for any given $\epsilon > 0$, there is an $l > 0$, such that

$$D_h(x^\infty, x^{k_l}) < \epsilon. \tag{3.12}$$

Therefore, for any $k \ge k_l$, it follows from (2.2), (3.11) and (3.12) that

$$\frac{1}{2}\|x^\infty - x^k\|^2 \le D_h(x^\infty, x^k) \le D_h(x^\infty, x^{k_l}) < \epsilon,$$

and thus the sequence $\{x^k\}$ converges to $x^\infty$. □

## 4. Two methods for application

**Method 1.** (For monotone variational inequalities) *Let $\Omega$ be a nonempty closed convex subset of $R^n$ and $F$ be a continuous monotone mapping from $R^n$ into itself. A variational inequality problem, denoted by VI($\Omega, F$), is to determine a vector $x^* \in \Omega$ such that*

$$(x - x^*)^T F(x^*) \ge 0, \qquad \forall\ x \in \Omega.$$

*Another popular reformulation for VI($\Omega, F$) is the multi-valued equation*

$$0 \in \hat{T}(x) := T(x) + N_\Omega(x),$$

*where $T(x) = F(x)$. Obviously, in the proposed method $y^k$ is defined as $F(\tilde{x}^k)$ and from* (1.7) *and Lemma* 2.1*, we get*

$$\tilde{x}^k = P_\Omega\{\tilde{x}^k - [\nabla h(\tilde{x}^k) - \nabla h(x^k) - e^k + \beta_k F(\tilde{x}^k)]\} \tag{4.1}$$

*with $\|e^k\| \le \mu\eta_k\|x^k - \tilde{x}^k\|$. We can find an $\tilde{x}^k$ satisfying* (4.1) *through the following method. First, similar to proximal point algorithms, for any given $h(x)$ which satisfies Assumption* 1.2*, there exists an $x \in \Omega$, such that*

$$(x' - x)^T[\nabla h(x) - \nabla h(x^k) + \beta_k F(x)] \ge 0, \qquad \forall\ x' \in \Omega. \tag{4.2}$$

*Due to Lemma* 2.1*, this is equivalent to*

$$x = P_\Omega\{x - [\nabla h(x) - \nabla h(x^k) + \beta_k F(x)]\}.$$

*The approximate solution* $\bar{x}$ *of* (4.2) *can be found through some iteration methods, i.e.,*

$$\bar{x} \approx P_\Omega\{\bar{x} - [\nabla h(\bar{x}) - \nabla h(x^k) + \beta_k F(\bar{x})]\}. \tag{4.3}$$

*Define* $\tilde{x}^k$ *as follows:*

$$\begin{aligned}\tilde{x}^k &= P_\Omega\{\bar{x} - [\nabla h(\bar{x}) - \nabla h(x^k) + \beta_k F(\bar{x})]\} \\ &= P_\Omega\{\tilde{x}^k - [\nabla h(\tilde{x}^k) - \nabla h(x^k) - e^k + \beta_k F(\tilde{x}^k)]\},\end{aligned} \tag{4.4}$$

*where*

$$e^k = [\nabla h(\tilde{x}^k) - \nabla h(\bar{x})] - [\tilde{x}^k - \bar{x}] + \beta_k[F(\tilde{x}^k) - F(\bar{x})]$$

*and*

$$\|e^k\| \le \mu\eta_k\|x^k - \tilde{x}^k\|. \tag{4.5}$$

**Remark 4.1.** From (4.3) and (4.4), we obtain that $\bar{x} \approx \tilde{x}^k$ and thus the condition (4.5) on $e^k$ can be met from the continuities of $\nabla h(x)$ and $F(x)$, when $\|\bar{x} - \tilde{x}^k\|$ is small enough. Method 1 is the extension of Section 4 in [8].

**Method 2.** (Example for choice of $y^k$)

1. *Given* $x^k \in \Omega$*, find an* $\tilde{x}^k \in \Omega$ *such that*

$$\nabla h(x^k) + e^k \in \nabla h(\tilde{x}^k) + \beta_k T(\tilde{x}^k) \tag{4.6}$$

*for some* $e^k$ *satisfying*

$$\|e^k\| \le \mu\eta_k\|x^k - \tilde{x}^k\| \quad \textit{and} \quad \sup_k \eta_k < 1.$$

2. *Set*

$$y^k = \frac{1}{\beta_k}[\nabla h(x^k) + e^k - \nabla h(\tilde{x}^k)], \tag{4.7}$$

$$\nabla h(x^{k+1}) - \nabla h(x^k) + \alpha_k\beta_k y^k \in -N_\Omega(x^{k+1}),$$

*where* $\alpha_k$ *is defined in* (1.10)*.*

3. *Terminate if a stopping criterion is met, otherwise set* $k = k+1$ *and goto* 1*.*

**Remark 4.2.** From (4.6) and (4.7), we know that $y^k \in T(\tilde{x}^k)$. Since $0 \in N_\Omega(\tilde{x}^k)$ and $\hat{T} = T + N_\Omega$,

$$\nabla h(x^k) + e^k \in \nabla h(\tilde{x}^k) + \beta_k \hat{T}(\tilde{x}^k).$$

This method gives us an example for the choice of $y^k$ in application.

## 5. Conclusion remarks

In this paper, we suggest an approximate nonlinear proximal prediction-correction algorithm for finding the zero points of maximal monotone operators. It is shown that the method has global convergence under proper assumptions. Our main work is using $\nabla h(u)$ to substitute $u$ in the proximal algorithms and we only assume $h(u)$ satisfies Assumption 1.2. Moreover, in the prediction step, the presented algorithm allows for constant relative error tolerance, and in the correction step, a general decent direction and a suitable step length are used. It is our belief that the research on the choice of $h(u)$ is important in application and we hope this paper may stimulate further investigation in this direction.

## References


1. P. Bertsekas and J.N. Tsitsiklis, *Parallel and Distributed Computation, Numerical Methods*, Prentice-Hall, Englewood Cliffs, NJ, 1989.
2. L. M. Bregman, *The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming*, USSR Comput. Math. Math. Phys. **7** (1967), 200-217.
3. R. S. Burachik and A.N. Iusem, *A generalized proximal point algorithm for the variational inequality problem in a Hilbert space*, SIAM J. Optim. **8** (1998), 197-216.
4. Y. Censor and S.A. Zenios, *The proximal minimization algorithm with D-functions*, J. Optim. Theo. Appl. **73** (1992), 451-464.
5. J. Eckstein, *Nonlinear proximal point algorithms using Bregman functions, with applications to convex programming*, Math. Oper. Res. **18** (1993), 202-226.
6. J. Eckstein and D.P. Bertsekas, *On the Douglas-Rachford splitting method and the proximal points algorithm for maximal monotone operators*, Math. Program. **55** (1992), 293-318.
7. J. Eckstein, *Approximate iterations in Bregman-function-based proximal algorithms*, Math. Program. **83** (1998), 113-123.
8. B. S. He, *A new approximate proximal point algorithm for maximal monotone operator*, Science in China, Series A **46**(2) (2003), 200-206.



Min Li
Department of Mathematics
Nanjing University
22 Hankou Road
Nanjing, Jiangsu 210093
P. R. China
*E-mail address*: zrlmnju@yahoo.com.cn