GENERAL FRAMEWORK FOR PROXIMAL POINT ALGORITHMS ON \((A, \eta)\)-MAXIMAL MONOTONICITY FOR NONLINEAR VARIATIONAL INCLUSIONS

Ram U. Verma

Abstract. General framework for proximal point algorithms based on the notion of \((A, \eta)\)-maximal monotonicity (also referred to as \((A, \eta)\)-monotonicity in literature) is developed. Linear convergence analysis for this class of algorithms to the context of solving a general class of nonlinear variational inclusion problems is successfully achieved along with some results on the generalized resolvent corresponding to \((A, \eta)\)-monotonicity. The obtained results generalize and unify a wide range of investigations readily available in literature.

1. Introduction

Let \(X\) be a real Hilbert space with the norm \(\| \cdot \|\) and the inner product \(\langle \cdot, \cdot \rangle\). We consider a class inclusion problems of the form: find a solution to

\[
0 \in M(x),
\]

where \(M : X \rightarrow 2^X\) is a set-valued mapping on \(X\).

Motivated by the algorithmic advances [1, 2, 3, 7, 10], we develop a hybrid relaxed proximal point algorithm considered by Eckstein and Bertsekas [2] based on the notions of \(A\)-maximal monotonicity [10] and \((A, \eta)\)-maximal monotonicities [11] for solving general inclusion problems. These new notions generalize the general class of maximal monotone set-valued mappings, including the notion of \(H\)-maximal monotonicity introduced by Fang and Huang [3] in a Hilbert space setting. This greatly impacts on a general class of problems of variational character, including minimization or maximization (whether constraint or not) of functions, variational problems, and minimax problems, that can be unified into the form (1). Recently, the author [11] introduced and studied the notion of \((A, \eta)\)-maximal monotonicity to the context of approximating the solution of an inclusion problem based on the generalized resolvent operator technique. The generalized resolvent operator technique can also be applied...
to other problems, such as equilibria problems in economics, management sciences, optimization and control theory, operations research, and mathematical programming. For more details on the resolvent operator technique and its applications, we refer the reader [1-17].

2. \((A, \eta)\)-maximal monotonicity

In this section we discuss some results based on basic properties of \((A, \eta)\)-maximal monotonicity and its variant forms. Let \(M : X \to 2^X\) be a multivalued mapping on \(X\). We shall denote both the map \(M\) and its graph by \(M\), that is, the set \(\{(x, y) : y \in M(x)\}\). This is equivalent to stating that a mapping is any subset \(M\) of \(X \times X\), and \(M(x) = \{y : (x, y) \in M\}\). If \(M\) is single-valued, we shall still use \(M(x)\) to represent the unique \(y\) such that \((x, y) \in M\) rather than the singleton set \(\{y\}\). This interpretation shall much depend on the context.

The domain of a map \(M\) is defined (as its projection onto the first argument) by

\[
D(M) = \{x \in X : \exists y \in X : (x, y) \in M\} = \{x \in X : M(x) \neq \emptyset\}.
\]

\(D(M) = X\) shall denote the full domain of \(M\), and the range of \(M\) is defined by

\[
R(M) = \{y \in X : \exists x \in X : (x, y) \in M\}.
\]

The inverse \(M^{-1}\) of \(M\) is \(\{(y, x) : (x, y) \in M\}\). For a real number \(\rho\) and a mapping \(M\), let \(\rho M = \{x, \rho y) : (x, y) \in M\}\). If \(L\) and \(M\) are any mappings, we define

\[
L + M = \{(x, y + z) : (x, y) \in L, (x, z) \in M\}.
\]

**Definition 2.1.** Let \(M : X \to 2^X\) be a multivalued mapping on \(X\). The map \(M\) is said to be:

(i) \((r)\)-strongly monotone if there exists a positive constant \(r\) such that

\[
\langle u^* - v^*, u - v \rangle \geq r\|u - v\|^2 \forall (u, u^*), (v, v^*) \in M.
\]

(ii) \((m)\)-relaxed monotone if there exists a positive constant \(m\) such that

\[
\langle u^* - v^*, u - v \rangle \geq (m\|u - v\|^2 \forall (u, u^*), (v, v^*) \in M.
\]

**Definition 2.2.** Let \(M : X \to 2^X\) be a multivalued mapping on \(X\), and let \(\eta : X \times X \to X\) be another mapping. The map \(M\) is said to be:

(i) \((r, \eta)\)-strongly monotone if there exists a positive constant \(r\) such that

\[
\langle u^* - v^*, \eta(u, v) \rangle \geq r\|u - v\|^2 \forall (u, u^*), (v, v^*) \in M.
\]

(ii) \((1, \eta)\)-strongly monotone if

\[
\langle u^* - v^*, \eta(u, v) \rangle \geq \|u - v\|^2 \forall (u, u^*), (v, v^*) \in M.
\]

(iii) \(\eta\) is said to be \((\tau)\)-Lipschitz continuous if there is a positive constant \(\tau\) such that

\[
\|\eta(u, v)\| \leq \tau\|u - v\|.
\]
**Definition 2.3** ([10]). Let $A : X \to X$ be $(r)$-strongly monotone. The map $M : X \to 2^X$ is said to be $A$-maximal monotone if
1. $M$ is $(m)$-relaxed monotone,
2. $R(A + \rho M) = X$ for $\rho > 0$.

**Definition 2.4.** Let $A : X \to X$ be $(r, \eta)$-strongly monotone. The map $M : X \to 2^X$ is said to be $(A, \eta)$-maximal monotone if
1. $M$ is $(m, \eta)$-relaxed monotone,
2. $R(A + \rho M) = X$ for $\rho > 0$.

**Definition 2.5.** Let $A : X \to X$ be an $(r, \eta)$-strongly monotone mapping and let $M : X \to 2^X$ be an $(A, \eta)$-maximal monotone mapping. Then the generalized resolvent operator $J_{\rho, A}^{M, \eta} : X \to X$ is defined by
$$J_{\rho, A}^{M, \eta}(u) = (A + \rho M)^{-1}(u) \text{ for } r - \rho m > 0.$$

**Proposition 2.1.** Let $A : X \to X$ be an $(r, \eta)$-strongly monotone mapping and let $M : X \to 2^X$ be an $(A, \eta)$-maximal monotone mapping. Then the operator $(A + \rho M)^{-1}$ is single-valued for $r - \rho m > 0$.

**Proof.** Suppose that for some $z \in X$, there are $x, y \in (A + \rho M)^{-1}(z)$. Then we have
$$-A(x) + z \in \rho M(x) \text{ and } A(y) + z \in \rho M(y).$$
Since $M$ is $(A, \eta)$-maximal and $A$ is $(r, \eta)$-strongly monotone, it follows that
$$-\langle A(x) - A(y), \eta(x, y) \rangle \geq -\rho m \|x - y\|^2$$
$$\Rightarrow -\rho m \|x - y\|^2 \leq -\langle A(x) - A(y), \eta(x, y) \rangle \leq -r \|x - y\|^2$$
$$\Rightarrow (r - \rho m) \|x - y\|^2 \leq 0$$
$$\Rightarrow x = y \text{ for } (r - \rho m) > 0.$$\hfill \Box

### 3. Hybrid proximal point algorithms

This section deals with a hybrid proximal point algorithm to the relaxed version of the proximal point algorithm [2] and its application to approximation solvability of the inclusion problem (1) based on the $(A, \eta)$-maximal monotonicity.

**Lemma 3.1** ([11]). Let $X$ be a real Hilbert space, let $A : X \to X$ be $(r, \eta)$-strongly monotone, let $M : X \to 2^X$ be $(A, \eta)$-maximal monotone, and let $\eta : X \times X \to X$ be $(r)$-Lipschitz continuous. Then the generalized resolvent operator associated with $M$ and defined by
$$J_{\rho, A}^{M, \eta}(u) = (A + \rho M)^{-1}(u) \forall u \in X,$$
is $(\frac{r}{r - \rho m})$-Lipschitz continuous.

Furthermore, we have
$$\langle J_{\rho, A}^{M, \eta}(u) - J_{\rho, A}^{M, \eta}(v), u - v \rangle \leq \frac{r}{r - \rho m} \|u - v\|^2$$
for \( r - \rho m > 0 \).

**Theorem 3.1.** Let \( X \) be a real Hilbert space, let \( A : X \to X \) be \((r, \eta)\)-strongly monotone, and let \( M : X \to 2^X \) be \((A, \eta)\)-maximal monotone. Then the following statements are mutually equivalent:

(i) An element \( u \in X \) is a solution to (1).

(ii) For an \( u \in X \), we have

\[
    u = J_{\rho, A}^{M, \eta}(A(u)),
\]

where

\[
    J_{\rho, A}^{M, \eta}(u) = (A + \rho M)^{-1}(u) \quad \text{for} \quad r - \rho m > 0.
\]

**Theorem 3.2** ([3]). Let \( X \) be a real Hilbert space, let \( H : X \to X \) be \((r, \eta)\)-strongly monotone, and let \( M : X \to 2^X \) be \((H, \eta)\)-maximal monotone. Then the following statements are mutually equivalent:

(i) An element \( u \in X \) is a solution to (1).

(ii) For an \( u \in X \), we have

\[
    u = J_{\rho, H}^{M, \eta}(H(u)),
\]

where

\[
    J_{\rho, H}^{M, \eta}(u) = (H + \rho M)^{-1}(u).
\]

In the following theorem, we apply the hybrid proximal point algorithm to approximate the solution of (1), and as a result, we end up showing linear convergence.

**Theorem 3.3.** Let \( X \) be a real Hilbert space, let \( A : X \to X \) be \((r, \eta)\)-strongly monotone, and let \( M : X \to 2^X \) be \((A, \eta)\)-maximal monotone. Let \( \eta : X \times X \to X \) be \((\tau)\)-Lipschitz continuous. Suppose further that \( AOJ_{\rho, A}^{M, \eta} \) is \((\lambda)\)-cocoercive for \( \lambda > 1 \), that is, for all \( u, v \in X \),

\[
    \langle A(u) - A(v), A(J_{\rho, A}^{M, \eta}(A(u))) - A(J_{\rho, A}^{M, \eta}(A(v))) \rangle \geq \lambda \| A(J_{\rho, A}^{M, \eta}(A(u))) - A(J_{\rho, A}^{M, \eta}(A(v))) \|^2.
\]

For an arbitrarily chosen initial point \( x^0 \), let the sequence \( \{x^k\} \) be generated by an iterative algorithm of the form

\[
    A(x^{k+1}) = (1 - \alpha_k)A(x^k) + \alpha_k y^k \quad \forall \, k \geq 0,
\]

and \( y^k \) satisfies

\[
    \| y^k - A(J_{\rho, A}^{M, \eta}(A(x^k))) \| \leq \delta_k \| y^k - A(x^k) \|,
\]

where \( J_{\rho, A}^{M, \eta} = (A + \rho_k M)^{-1} \) for \( r - \rho_k m > 0 \), and

\[
    \{\delta_k\}, \{\alpha_k\}, \{\rho_k\} \subseteq [0, \infty)
\]

are scalar sequences such that \( \sum_{k=0}^{\infty} \delta_k < \infty \), \( \delta_k \to 0 \), \( \alpha = \lim_{k \to \infty} \alpha_k \), \( \rho_k \uparrow \rho \leq \infty \), and \( \alpha_k < 1 \).
Then the sequence \( \{x^k\} \) converges linearly to a solution of (1) with convergence rate
\[
(4) \quad \sqrt{1 - 2\alpha \{1 - (1 - \alpha)^2 - \frac{1}{\lambda} - \frac{1}{2}(1 - \alpha^2)\}} < 1
\]
for \( \lambda > 1 \).

Proof. Suppose that \( x^* \) is a zero of \( M \). From Theorem 3.1, it follows that any solution to (1) is a fixed point of \( J_{\mu_k,\lambda}(A(x)) \). For all \( k \geq 0 \), we express
\[
A(x^{k+1}) = (1 - \alpha_k)A(x^k) + \alpha_kA(J_{\mu_k,\lambda}(A(x^k))).
\]
Next, we find the estimate using (2) and its implications that
\[
\|A(x^{k+1}) - A(x^*)\|^2 = \|(1 - \alpha_k)A(x^k) + \alpha_kA(J_{\mu_k,\lambda}(A(x^k))) - (1 - \alpha_k)A(x^*) - \alpha_kA(x^*)\|^2
\]
\[
\leq (1 - \alpha_k)^2\|A(x^k) - A(x^*)\|^2 + 2\alpha_k(1 - \alpha_k)\frac{1}{\lambda}\|A(x^k) - A(x^*)\|^2
\]
\[
\leq (1 - \alpha_k)^2\|A(x^k) - A(x^*)\|^2 + 2\alpha_k(1 - \alpha_k)\frac{1}{\lambda}\|A(x^k) - A(x^*)\|^2
\]
\[
= (1 - \alpha_k)^2\|A(x^k) - A(x^*)\|^2 + 2\alpha_k(1 - \alpha_k)\frac{1}{\lambda}\|A(x^k) - A(x^*)\|^2
\]
\[
\leq (1 - \alpha_k)^2\|A(x^k) - A(x^*)\|^2 + 2\alpha_k\|A(x^k) - A(x^*)\|^2
\]
\[
\leq (1 - \alpha_k)^2\|A(x^k) - A(x^*)\|^2 + 2\alpha_k\|A(x^k) - A(x^*)\|^2
\]
\[
\leq (1 - \alpha_k)^2\|A(x^k) - A(x^*)\|^2 + 2\alpha_k\|A(x^k) - A(x^*)\|^2
\]
\[
= \|A(x^{k+1}) - A(x^*)\| \leq \theta_k\|A(x^k) - A(x^*)\|,
\]
where
\[
\theta_k = \sqrt{1 - 2\alpha_k\{1 - (1 - \alpha_k)^2 - \frac{1}{\lambda} - \frac{1}{2}(1 - \alpha^2)\}}.
\]
Since \( A(x^{k+1}) = (1 - \alpha_k)A(x^k) + \alpha_ky^k \), we have \( A(x^{k+1}) - A(x^k) = \alpha_k(y^k - A(x^k)) \). It follows that
\[
\|A(x^{k+1}) - A(x^*)\|
\]
\[
= \|(1 - \alpha_k)A(x^k) + \alpha_ky^k - (1 - \alpha_k)A(x^k) - \alpha_kA(J_{\mu_k,\lambda}(A(x^k)))\|
\]
\[
= \|\alpha_k(y^k - A(J_{\mu_k,\lambda}(A(x^k))))\|
\]
Hence, we have
\[ r_k \frac{\|y^k - A(x^k)\|}{\|x^k - y^k\|}. \]
Next, we find the estimate
\[ \|A(x^{k+1}) - A(x^*)\| \leq \|A(x^{k+1}) - A(x^*)\| + \|A(x^{k+1}) - A(z^{k+1})\| \]
\[ \leq \|A(z^{k+1}) - A(x^*)\| + \alpha_k \delta_k \|y^k - A(x^k)\| \]
\[ \leq \|A(z^{k+1}) - A(x^*)\| + \delta_k \|A(x^{k+1}) - A(x^k)\| \]
\[ \leq \|A(z^{k+1}) - A(x^*)\| + \delta_k \|A(x^{k+1}) - A(x^k)\| + \delta_k \|A(x^k) - A(x^*\| \]
\[ \leq \theta_k \|A(x^k) - A(x^*\| + \delta_k \|A(x^{k+1}) - A(x^k)\| + \delta_k \|A(x^k) - A(x^*\|). \]
This implies that
\[ \|A(x^{k+1}) - A(x^*)\| \leq \frac{\theta_k + \delta_k}{1 - \delta_k} \|A(x^k) - A(x^*)\|, \]
where
\[ \limsup \frac{\theta_k + \delta_k}{1 - \delta_k} = \limsup \theta_k \]
\[ = \sqrt{1 - 2\alpha\left(1 - \frac{1}{\lambda} - \frac{1}{2\lambda^2} - \frac{1}{2}\right) < 1.} \]

It follows from (6) that \(\|A(x^k) - A(x^*)\| \to 0\) as \(k \to \infty\). On the other hand, \(A\) is \((r, \eta)\)-strongly monotone, and hence,
\[ \|A(x^k) - A(x^*)\| \geq \frac{r}{\tau} \|x^k - x^*\|. \]
Hence, we have
\[ \frac{r}{\tau} \|x^k - x^*\| \leq \|A(x^k) - A(x^*)\| \to 0, \]
and this concludes the proof. \(\square\)

**Theorem 3.4.** Let \(X\) be a real Hilbert space, let \(H : X \to X\) be \((r, \eta)\)-strongly monotone, and let \(M : X \to 2^X\) be \((H, \eta)\)-maximal monotone. Let \(\eta : X \times X \to X\) be \((\tau)\)-Lipschitz continuous. Suppose further that \(H_0J_{\rho_k, H}^{M, \eta}\) is \((\lambda)\)-cocoercive for \(\lambda > 1\), that is, for all \(u, v \in X\),
\[ \langle H(u) - H(v), J_{\rho_k, H}^{M, \eta}(u) - J_{\rho_k, H}^{M, \eta}(v) \rangle \geq \lambda \| J_{\rho_k, H}^{M, \eta}(u) - J_{\rho_k, H}^{M, \eta}(v) \|^2. \]
(7)

For an arbitrarily chosen initial point \(x^0\), let the sequence \(\{x^k\}\) be generated by another iterative algorithm
\[ H(x^{k+1}) = (1 - \alpha_k)H(x^k) + \alpha_k y^k \forall k \geq 0, \]
(8)
and \(y^k\) satisfies
\[ \|y^k - H(J_{\rho_k, H}^{M, \eta}(x^k))\| \leq \delta_k \|y^k - H(x^k)\|. \]
where $J_{\rho_k,H}^{M,\eta} = (H + \rho_k M)^{-1}$, and
\[
\{\delta_k, \{\alpha_k\}, \{\rho_k\} \subseteq [0, \infty)
\]
are scalar sequences such that $\sum_{k=0}^{\infty} \delta_k < \infty$, $\delta_k \to 0$, $\alpha = \lim \sup_{k \to \infty} \alpha_k$, $\rho_k \uparrow \rho \leq \alpha$, $\alpha_k < 1$ and $\lambda > 1$.

Then the sequence $\{x^k\}$ converges linearly to a solution of (1) for $\lambda > 1$.

**Proof.** The proof is similar to that of Theorem 3.3. Suppose that $x^*$ is a zero of $M$. From Theorem 3.2, it follows that any solution to (1) is a fixed point of $J_{\rho_k,H}^{M,\eta} \circ H$. For all $k \geq 0$, we express
\[
H(x^{k+1}) = (1 - \alpha_k)H(x^k) + \alpha_k H(J_{\rho_k,H}^{M,\eta}(H(x^k))
\]
Next, we find the estimate using (7) and its variants that
\[
\|H(z^{k+1}) - H(x^*)\|_2^2
= \|(1 - \alpha_k)H(x^k) + \alpha_k H(J_{\rho_k,H}^{M,\eta}(H(x^k))
- \|(1 - \alpha_k)H(x^*) + \alpha_k H(J_{\rho_k,H}^{M,\eta}(H(x^*)))\|_2^2
\]
\[
= \|(1 - \alpha_k)(H(x^k) - H(x^*)) + \alpha_k[H(J_{\rho_k,H}^{M,\eta}(H(x^k)) - H(J_{\rho_k,H}^{M,\eta}(H(x^*))))\|_2^2
\]
\[
= (1 - \alpha_k)^2\|H(x^k) - H(x^*)\|_2^2
+ 2\alpha_k(1 - \alpha_k)\|H(x^k) - H(x^*)\|_2^2
+ \alpha_k^2\|H(J_{\rho_k,H}^{M,\eta}(H(x^k)) - H(J_{\rho_k,H}^{M,\eta}(H(x^*))))\|_2^2
\]
\[
\leq (1 - \alpha_k)^2\|H(x^k) - H(x^*)\|_2^2 + 2\alpha_k(1 - \alpha_k)\frac{1}{\lambda}\|H(x^k) - H(x^*)\|_2^2
+ \alpha_k^2\|H(J_{\rho_k,H}^{M,\eta}(H(x^k)) - H(J_{\rho_k,H}^{M,\eta}(H(x^*))))\|_2^2
\]
\[
\leq (1 - \alpha_k)^2\|H(x^k) - H(x^*)\|_2^2 + 2\alpha_k(1 - \alpha_k)\frac{1}{\lambda}\|H(x^k) - H(x^*)\|_2^2
+ \alpha_k^2\frac{1}{\lambda^2}\|H(x^k) - H(x^*)\|_2^2
\]
\[
= \|H(z^{k+1}) - H(x^*)\| \leq \theta_k\|H(x^k) - H(x^*)\|
\]
where
\[
\theta_k = \sqrt{1 - 2\alpha_k(1 - \alpha_k)\frac{1}{\lambda} - \frac{1}{2}\alpha_k(\frac{1}{\lambda^2} - \frac{1}{2}\alpha_k)}
\]

Since $H(x^{k+1}) = (1 - \alpha_k)H(x^k) + \alpha_k y^k$, we have $H(x^{k+1}) - H(x^k) = \alpha_k(y^k - H(x^k))$. It follows that
\[
\|H(x^{k+1}) - H(z^{k+1})\|
= \|(1 - \alpha_k)H(x^k) + \alpha_k y^k - [(1 - \alpha_k)H(x^k) + \alpha_k H(J_{\rho_k,H}^{M,\eta}(H(x^k)))]\|
Next, we find the estimate

\[
\|H(x^{k+1}) - H(x^*)\| \\
\leq \|H(z^{k+1}) - H(x^*)\| + \|H(x^{k+1}) - H(z^{k+1})\| \\
\leq \|H(z^{k+1}) - H(x^*)\| + \alpha_k \delta_k \|y^k - H(x^k)\| \\
\leq \|H(z^{k+1}) - H(x^*)\| + \delta_k \|H(x^{k+1}) - H(x^k)\| \\
\leq \|H(z^{k+1}) - H(x^*)\| + \delta_k \|H(x^{k+1}) - H(x^k)\| + \delta_k \|H(x^{k+1}) - H(x^*)\| \\
\leq \theta_k \|H(x^k) - H(x^*)\| + \delta_k \|H(x^{k+1}) - H(x^*)\| + \delta_k \|H(x^k) - H(x^*)\|.
\]

This implies that

\[
\|H(x^{k+1}) - H(x^*)\| \leq \frac{\theta_k \delta_k}{1 - \delta_k} \|H(x^k) - H(x^*)\|,
\]

where

\[
\lim \sup \frac{\theta_k \delta_k}{1 - \delta_k} = \lim \sup \theta_k
\]

\[
= \sqrt{\left[1 - 2\alpha \{1 - (1 - \alpha) \frac{1}{\lambda} - \frac{1}{2} \alpha(1/\lambda^2) - \frac{1}{2} \alpha\}\right]} < 1.
\]

Thus, \(\|H(x^k) - H(x^*)\| \to 0\). Since \(H\) is \((r, \eta)\)-strongly monotone, it implies

\[
\|H(x^k) - H(x^*)\| \geq \frac{r}{\tau} \|x^k - x^*\|
\]

we have

\[
\frac{r}{\tau} \|x^k - x^*\| \leq \|H(x^k) - H(x^*)\| \to 0.
\]

\[\square\]

**Concluding Remark.** We observe that because of the linear convergence concerns, the applications of the most of the constants such as \(\alpha, \beta, m\) and \(\tau\) are pretty much limited during the proofs except for \(r - \rho m > 0\) applied in Proposition 2.1, and \(\frac{r}{\tau}\) is applied to the final stages of the proof, though it is crucial achieving linear convergence in a general setting. A reasonably dominant role is played by the \((\lambda)\)-cocoercivity of the composition \(A_0 J_{\rho_0, A}^{M, \eta}\) for \(\lambda > 1\), while Proposition 2.1 is turned out to be quite significant to the single-valuedness of the generalized resolvent. Moreover, the convergence analysis is consistent with the theory of classical resolvents in the sense that linear convergence is unachievable for a general setting of classical resolvent when \(A/ H\) is just the identity mapping.
References


Department of Mathematics
Seminoles State College of Florida
Sanford, Florida 32773, USA
E-mail address: verma99@msn.com