Noname manuscript No.
(will be inserted by the editor)

New efficient approach in finding a zero of a maximal monotone operator

B. K. Le

Received: date / Accepted: date

Abstract In the paper, we provide a new efficient approach to find a zero of a maximal monotone operator under very mild assumptions. Using a regularization technique and the proximal point algorithm, we can construct a sequence that converges strongly to a solution with at least linear convergence rate.

Keywords Maximal monotone operators, proximal point algorithm, regularization, fast convergence

1 Introduction

Our aim is to provide an efficient algorithm to solve numerically a zero of a set-valued maximal monotone operator $A : H \rightrightarrows H$, i. e., to find $x \in H$ such that

$$0 \in Ax, \tag{1}$$

where H is a given real Hilbert space. It is one of the fundamental problem in convex optimization which has been extensively studied (see, e. g., [1–3, 8, 13, 14] and the references therein). Indeed it is known that minimizing a proper convex lower semicontinuous function $f : H \to H$ can be reduced to this problem, i. e.,

$$0 \in \partial f(x). \tag{2}$$

The proximal point algorithm [10, 11, 13] is perhaps the most popular method to solve (1), which is given by

$$x_0 \in H, \;\; x_{k+1} = J_{\gamma A} x_k, \;\; k = 1, 2, 3 \dots \tag{3}$$

for some $\gamma > 0$ where $J_{\gamma A} := (I + \gamma A)^{-1}$ denotes the resolvent of A. Weak convergence of the proximal point algorithm follows by the firm non-expansiveness of $J_{\gamma A}$ and Opial's Lemma [8, 13]. Strong convergence can be obtained by combining

B. K. Le
Faculty of Mathematics and Statistics, Ton Duc Thang University, Ho Chi Minh City, Vietnam
(lebakhiet@tdtu.edu.vn)

with the Krasnoselskii–Mann iteration scheme or Halpern's procedure (see, e.g., [12, 14, 16]). Linear convergence of the algorithm is an attractive property, which was confirmed by Rockafellar [13] when A is strongly monotone or more weakly, A^{-1} is Lipschitz continuous at zero. However, the Lipschitz continuity of A^{-1} at zero in Rockafellar's sense (see [13, 15]) is quite restrictive since it requires that $A^{-1}(0)$ is singleton, i. e., the problem (1) must have a unique solution, which is not satisfied in many applications.

In this paper, we establish a new way to solve (1) with linear convergence rate under only the upper semicontinuity of A^{-1} at zero in the sense of set-valued analysis, which seems to be more natural than Rockafellar's sense. The idea is to solve the regularized problem

$$0 \in (A + \varepsilon Id)x_\varepsilon, \tag{4}$$

for $\varepsilon > 0$ small enough using the proximal point algorithm where Id denotes the identity mapping. We show that x_ε is closed to some $x^* \in A^{-1}(0)$ as ε tends to zero thanks to the upper semicontinuity of A^{-1} at zero. It is interesting for the practical view since we only need to solve one strongly monotone inclusions to obtain the approximate solution for a given accuracy.

The paper is organized as follows. First we recall some definitions and useful results concerning maximal monotone operators in Section 2. An efficient procedure for solving (1) and an application to convex optimization are presented in Section 3 and Section 4 respectively. Finally, some conclusions end the paper in Section 5.

2 Notations and preliminaries

Let H be a given real Hilbert space with its norm $\|\cdot\|$ and scalar product $\langle \cdot, \cdot \rangle$. The closed unit ball is denoted by $\mathbb{B}$. A mapping $f : H \to H$ is called *nonexpansive* if it is Lipschitz continuous with modulus 1. It is called *firmly nonexpansive* if

$$\|f(x) - f(y)\|^2 \le \|x - y\|^2 - \|(Id - f)(x) - (Id - f)(y)\|^2, \; \forall\, x, y \in H.$$

The domain, the range and the graph of a set-valued mapping $A : H \rightrightarrows H$ are defined respectively by

$$\mathrm{dom}(A) = \{x \in H : \; A(x) \neq \emptyset\}, \;\; \mathrm{rge}(A) = \bigcup_{x \in H} A(x)$$

and

$$\mathrm{gph}(A) = \{(x, y) : x \in H, y \in A(x)\}.$$

It is called *monotone* provided

$$\langle x^* - y^*, x - y \rangle \ge 0 \;\; \forall\, x, y \in H, x^* \in A(x) \text{ and } y^* \in A(y).$$

In addition, it is called *maximal monotone* if there is no monotone operator A' such that the graph of A is strictly contained in the graph of A'.

We have the following classical result (see, e. g.,[6]).

Proposition 1 *Let $A : H \rightrightarrows H$ be a maximal monotone operator. Then the graph of A is strongly-weakly closed, i. e., if the sequences (y_n) converges strongly to some y, (x_n) converges weakly to some x and $y_n \in A(x_n)$ then $y \in A(x)$.*

It is called *μ-strongly monotone* ($\mu > 0$) provided

$$\langle Ax - Ay, x - y\rangle \geq \mu\|x - y\|^2 \;\; \forall\, x, y \in H.$$

We can see that if A is μ-strongly monotone then A^{-1} is single-valued $\frac{1}{\mu}$-Lipschitz continuous.

The *resolvent* of A are defined respectively as follows

$$J_A := (Id + A)^{-1}.$$

Remark 1 The resolvents of maximal monotone operators are firmly nonexpansive (see, e.g., [13]).

The set-valued operator A is called upper semicontinuous at $0 \in \mathrm{dom}(A)$ (see also [4]) if there exist an increasing function $\rho : \mathbb{R}^+ \to \mathbb{R}^+$ with $\rho(0) = 0$ and $\epsilon > 0$ such that

$$A(x) \subset A(0) + \rho(\|x\|)\mathbb{B}, \forall x \in \epsilon\mathbb{B}. \tag{5}$$

It is called Lipschitz continuous at $0 \in \mathrm{dom}(A)$ if there exists $L > 0$ and $\epsilon > 0$ such that

$$A(x) \subset A(0) + L\|x\|\mathbb{B}, \forall x \in \epsilon\mathbb{B}.$$

Remark 2 i) Our Lipschitz continuity is weaker than the Lipschitz continuity in Rockafellar's sense [13, 15] since it allows $A(0)$ is a set. Two definitions coincides if $A(0)$ is singleton. The Sign mapping, defined by

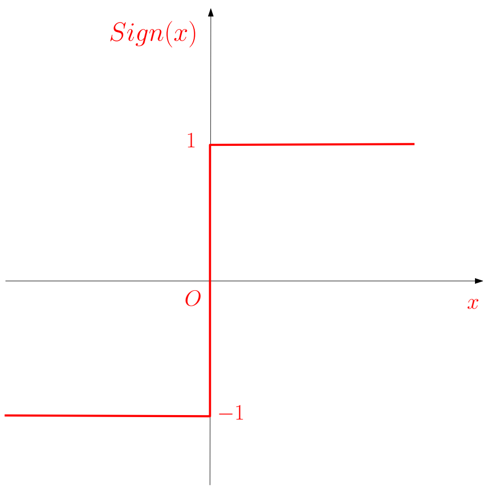

Fig. 1 The *Sign* function

$$Sign(x) = \begin{cases} -1 & if \;\; x < 0, \\ [-1,1] & if \;\; x = 0, \\ 1 & if \;\; x > 0, \end{cases}$$

is Lipschitz continuous at 0 *in our sense but not Lipschitz continuous at* 0 *in Rockafellar's sense. We also note that* $Sign = N^{-1}_{[-1,1]}$*, the inverse of the normal cone to the set* $[-1,1]$*. The mappings Sign,* $N_{[-1,1]}$ *and set-valued operators with similar form play an important role in engineerings and mechanics [7, 9].*

ii) If A is Lipschitz continuous at 0 *then it is upper semicontinuous at* 0 *with* $\rho(\|x\|) = L\|x\|$ *but the converse is not true, e. g., when* $\rho(\|x\|) = \sqrt{\|x\|}$.

3 Main result

In this section, we provide an efficient approach to solve (1) numerically. The only assumptions are that A is maximal monotone with non-empty set $A^{-1}(0)$ and A^{-1} is upper semicontinuous at zero in the set-valued sense. It is significantly weaker than the assumption used in [13, 15]. Let x^* be a fixed element of $A^{-1}(0)$. We have the following results.

Lemma 1 *Let* $\epsilon > 0$ *small enough and* $x_\epsilon := (A + \epsilon Id)^{-1}(0)$*. Then*

$$\|x_\epsilon\| \le \|x^*\|, \tag{6}$$

and

$$\|x_\epsilon - y^*\| \le \rho(\epsilon\|x^*\|) \tag{7}$$

for some $y^* \in A^{-1}(0)$ *where* $\rho(\cdot)$ *is defined in (5).*

Proof Note that x_ϵ is well-defined since $A + \epsilon Id$ is strongly monotone. We have

$$0 \in Ax^* \;\text{ and }\; -\epsilon x_\epsilon \in Ax_\epsilon.$$

The monotonicity of A implies that

$$\epsilon\langle x_\epsilon, x_\epsilon - x^*\rangle \le 0,$$

and (6) follows. On the other hand,

$$x_\epsilon \in A^{-1}(-\epsilon x_\epsilon) \subset A^{-1}(0) + \rho(\epsilon\|x_\epsilon\|)\mathbb{B} \subset A^{-1}(0) + \rho(\epsilon\|x^*\|)\mathbb{B}. \tag{8}$$

Then there exists $y^* \in A^{-1}(0)$ such that $x_\epsilon - y^* \in \rho(\epsilon\|x^*\|)\mathbb{B}$ and we obtain the conclusion.

Theorem 1 *For given* n*,* $c_n > 0$ *and* $z_0 \in H$*, we choose* $\varepsilon_n > 0$ *small enough such that*

$$\rho(\varepsilon_n\|x^*\|) \le \frac{1}{(1+c_n)^n}.$$

Let $A_n := A + \varepsilon_n Id$ *and* $\gamma_n := \frac{c_n}{\varepsilon_n}$. *Let* $(z_{k,n})_{k\geq 1}$ *be the sequence generated by the proximal point algorithm*

$$z_{k+1,n} := J_{\gamma_n A_n} z_{k,n},\ k = 0, 1, 2 \ldots \tag{9}$$

Let $w_n := z_{n,n}$. *Then we have*

$$\|z_{k,n} - y^*\| \leq \frac{1}{(1+c_n)^k}(\|x^*\| + \|z_0\|) + \frac{1}{(1+c_n)^n}, k = 1, 2, 3, \ldots$$

and thus

$$\|w_n - y^*\| \leq \frac{1}{(1+c_n)^n}(\|x^*\| + \|z_0\| + 1),$$

for some $y^* \in A^{-1}(0)$.

Proof The positive real number ε_n is well-defined since ρ is increasing and $\rho(0) = 0$ (for example if $\rho(\|x\|) = L\|x\|$ then $\varepsilon_n = \frac{1}{L\|x\|^*(1+c_n)^n}$). Since A_n is ε_n-strongly monotone, we obtain that $J_{\gamma_n A_n}$ is $\frac{1}{1+\gamma_n\varepsilon_n} = \frac{1}{1+c_n}$-Lipschitz continuous (see, e. g., [13]). Let $x_{\varepsilon_n} := A_n^{-1}(0)$. Then $x_{\varepsilon_n} = J_{\gamma_n A_n} x_{\varepsilon_n}$ and using Lemma 1, there exists some $y^* \in A^{-1}(0)$ such that

$$\begin{aligned}\|z_{k,n} - y^*\| &\leq \|z_{k,n} - x_{\varepsilon_n}\| + \|x_{\varepsilon_n} - y^*\| \leq \frac{1}{(1+c_n)^k}\|x_{\varepsilon_n} - z_0\| + \rho(\varepsilon_n\|x^*\|)\\ &\leq \frac{1}{(1+c_n)^k}(\|x^*\| + \|z_0\|) + \frac{1}{(1+c_n)^n}\end{aligned}$$

since $\|x_{\varepsilon_n}\| \leq \|y^*\|$. The conclusion thereby follows.

Remark 3 i) Using Theorem 1, we only have to solve one strongly monotone inclusion with at least linear convergence rate to obtain an approximate solution for a given accuracy, which is interesting for the practical view.

ii) If $A^{-1}(0) = \{x^*\}$ *is singleton, then* x_ε *converges to* x^* *due to (7). In general,* x_ε *is closed to one of the solutions when* ε *small enough and every weak limit point of* (x_ε) *is a solution of (1) (see Proposition 1 and Lemma 1).*

ii) The upper semicontinuous of A^{-1} *at zero is very mild and natural. Indeed* A^{-1} *is maximal monotone hence the graph of* A *is closed (Proposition 1). In particular, if* A^{-1} *is compact in a neighborhood of zero then* A^{-1} *is upper-semicontinuous at zero ([4, Theorem 1- p. 41]). In finite dimensional spaces, if the domain of* A *is bounded then this assumption is automatically satisfied.*

4 Application

Let us consider the unconstrained minimization problem of a convex, L-continuous function $f : H \to H$ with nonempty set $\partial f^{-1}(0)$ and ∂f^{-1} is upper semicontinuous at zero in the set-valued sense. Then for given a positive sequence (c_n), by using Theorem 1, we can find some constant $C > 0$ and a sequence (w_n) such that

$$f(w_n) - \min_{x\in H} f(x) \leq \frac{C}{(1+c_n)^n}. \tag{10}$$

5 Conclusions

The paper provides a new efficient approach in the finding a zero of a maximal monotone operator under very mild assumptions. An application to convex minimization with linear convergence is provided.

References

1. F. Alvarez, H. Attouch, *An inertial proximal method for maximal monotone operators via discretization of a nonlinear oscillator with damping*, Set-Valued Analysis 9, 3–11, 2001.
2. H. Attouch, J. Peypouquet,*Convergence of inertial dynamics and proximal algorithms governed by maximal monotone operators*, Math. Program. 174, 391–432, 2019.
3. H. Attouch, A. Cabot, *Convergence of a relaxed inertial proximal algorithm for maximally monotone operators*, Mathematical Programming (2019), DOI 10.1007/s10107-019-01412-0.
4. J. P. Aubin, A. Cellina, *Differential Inclusions*, vol. 264 of Grundlehren der Mathematischen Wissenschaften. Springer-Verlag, Berlin, 1984.
5. H. H. Bauschke, P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, Springer Berlin, 2011.
6. H. Brezis, *Opérateurs Maximaux Monotones et Semi-groupes de Contractions dans les Espaces de Hilbert*, Math. Studies 5, North-Holland American Elsevier (1973).
7. B. Brogliato, R. Lozano, B. Maschke, O. Egeland, *Dissipative Systems Analysis and Control*, Springer-verlag London, 2nd Edition, 2007.
8. J. Eckstein, *Splitting methods for monotone operators with applications to parallel optimization*, PhD thesis, MIT, 1989.
9. R. I. Leine, N. Wouw, *Stability and convergence of mechanical systems with unilateral constraints.* Springer Verlag, 2008.
10. B. Martinet, *Régularisation d'inéquations variationnelles par approximations successives.* Revue Française D'automatique, Informatique, Recherche Opérationnelle, 3, 154–158, 1970.
11. B. Martinet, *Determination approchée d'un point fixe d'une application pseudo-contractante.* In: Cas de l'application prox. CR Acad Sc Paris, ser A, 274:163–165, 1972.
12. K. Nakajo, W. Takahashi, *Strong convergence theorems for nonexpansive mappings and nonexpansive semigroups*, J. Math. Anal. Appl. 279, 372–379, 2003.
13. R. T. Rockafellar, *Monotone operators and the proximal point algorithm.* SIAM Journal on Control and Optimization, 14(5):877–898, 1976.
14. M.V. Solodov, B.F. Svaiter, *Forcing strong convergence of proximal point iterations in a Hilbert space*, Math. Programming Ser. A 87 (2000) 189–202.
15. M. Tao, X. Yuan, *On the Optimal Linear Convergence Rate of a Generalized Proximal Point Algorithm,* J Sci Comput 74, 826–850, 2018.
16. R. Wittmann, *Approximation of fixed points of nonexpansive mappings*, Arch. Math. 58, 486–491, 1992.