STRONG CONVERGENCE OF A PROXIMAL POINT ALGORITHM WITH BOUNDED ERROR SEQUENCE

O. A. BOIKANYO AND G. MOROŞANU

Abstract. Given any maximal monotone operator $A : D(A) \subset H \to 2^H$ in a real Hilbert space H with $A^{-1}(0) \neq \emptyset$, it is shown that the sequence of proximal iterates $x_{n+1} = (I + \gamma_n A)^{-1}(\lambda_n u + (1 - \lambda_n)x_n + e_n)$ converges strongly to the metric projection of u on $A^{-1}(0)$ for (e_n) bounded, $\lambda_n \in (0, 1)$ with $\lambda_n \to 1$ and $\gamma_n > 0$ with $\gamma_n \to \infty$ as $n \to \infty$. In comparison with our previous paper [Optim. Lett. 4 (2010), 635-641], where the error sequence was supposed to converge to zero, here we consider the classical condition that errors be bounded. In the case when A is the subdifferential of a proper convex lower semicontinuous function $\varphi : H \to (-\infty, +\infty]$, the algorithm can be used to approximate the minimizer of φ which is nearest to u.

1. Introduction

Let H be a real Hilbert space with scalar product $\langle \cdot, \cdot \rangle$ and induced norm $\| \cdot \|$. Consider the following set valued problem:

$$\text{find an } x \in D(A) \text{ such that } 0 \in A(x), \tag{1}$$

where $A : D(A) \subset H \to 2^H$ is a maximal monotone operator. We remind the reader that an operator $A : D(A) \subset H \to 2^H$ is called monotone if its graph $G(A) = \{(x, y) \in H \times H : x \in D(A), y \in Ax\}$ is a monotone subset of $H \times H$, that is, it satisfies the monotonicity property

$$\langle x - x', y - y' \rangle \geq 0, \quad \forall\, (x, y), (x', y') \in G(A).$$

If in addition to being monotone, the graph of A is not properly contained in the graph of any other monotone operator, then A is said to be maximal monotone. An interesting result concerning such operators is due to Minty [10] and states that a monotone operator A is maximal if and only if (I+tA) is a surjection for every $t > 0$. This characterization allows one to define a single valued and nonexpansive mapping $J_t : H \to H$ by the rule $x \mapsto (I + tA)^{-1}x$ for all $t > 0$, see, e.g., [9]. The map J_t is called the resolvent of A.

One of the most effective iterative methods for solving problem (1) is the proximal point algorithm (PPA) which was initiated by Martinet [7] and further developed by Rockafellar [11]. The PPA according to Rockafellar generates a sequence (x_n) via the rule

$$x_{n+1} = J_{\gamma_n}(x_n + e_n), \quad \text{for all } n \geq 0 \tag{2}$$

where $x_0 \in H$ is a given starting point, $(\gamma_n) \subset (0, \infty)$ and (e_n) is a sequence of computational errors. Since the PPA converges only weakly in general, to a solution of problem (1), see [1, 4], modifying the PPA in order to generate strongly convergent sequences has

2000 *Mathematics Subject Classification.* 47J25, 47H05, 47H09.

Key words and phrases. maximal monotone operator, proximal point algorithm, nonexpansive map, resolvent operator, prox-Tikhonov method.

been the subject of intense research. Recently, Lehdili and Moudafi [6] combined the proximal method (2) with the Tikhonov regularization to obtain the iterative process

$$x_{n+1} = J_{\gamma_n}(\alpha_n x_n + e_n), \quad \text{for all } n \geq 0 \tag{3}$$

where $x_0 \in H$ is a given starting point, $\gamma_n \in (0, \infty)$ and $\alpha_n \in (0, 1)$ with the requirement $\alpha_n \to 1$ as $n \to \infty$, and (e_n) is a sequence of computational errors. Under the summability condition on $(\|e_n\|)$ and some additional conditions on α_n and γ_n, they showed that the sequence generated by (3) converges strongly to a point of $A^{-1}(0)$ nearest to u provided that this set is not empty. Xu [13] extended the prox-Tikhonov method (3) to

$$x_{n+1} = J_{\gamma_n}(\lambda_n u + (1 - \lambda_n)x_n + e_n), \quad \text{for all } n \geq 0 \tag{4}$$

for given $x_0, u \in H$, $\gamma_n \in (0, \infty)$, $\lambda_n \in (0, 1)$, and (e_n) has the same meaning as above. Worth mentioning is the fact that when $\lambda_n \to 0$ and $\|e_n\| \to 0$ as $n \to \infty$, algorithm (4) above is equivalent to an iterative process of the form

$$x_{n+1} = \lambda_n u + (1 - \lambda_n)J_{\gamma_n} x_n + e_n, \quad \text{for all } n \geq 0, \tag{5}$$

see [2] for details. The scheme (5) was independently shown by Xu [12] and Kamimura and Takahashi [5] to be strongly convergent to a point of $A^{-1}(0)$ (if this set is not empty) which is nearest to u, provided that $\lambda_n \to 0$ with $\sum_{n=0}^{\infty} \lambda_n = \infty$, $(\|e_n\|)$ is summable and $\gamma_n \to \infty$. In [2], the authors extended this result to general errors which only converge to zero in norm.

In this paper, we pursue a different direction to derive a strong convergence result associated with algorithm (4). More precisely, we will show that if $A^{-1}(0) \neq \emptyset$, $\lambda_n \to 1$, (e_n) is bounded and $\gamma_n \to \infty$, then the sequence generated by (4) converges strongly to an element of $A^{-1}(0)$ which is nearest to u.

2. Main Result

Our analysis relies on the following lemma which describes the asymptotic behavior of the resolvent of a maximal monotone operator. It was proved independently by Bruck Jr. [3] and Moroşanu [8].

Lemma 1. *Let $A : D(A) \subset H \to 2^H$ be a maximal monotone operator with $\emptyset \neq F := A^{-1}(0)$. Then for any $u \in H$, $(I + tA)^{-1}u \to P_F u$ as $t \to \infty$, where $P_F u$ denotes the projection of u on F.*

We will also need the following lemma in proving our main result

Lemma 2. *Assume that $A : D(A) \subset H \to 2^H$ is a maximal monotone operator with $F := A^{-1}(0) \neq \emptyset$. For any fixed $x_0, u \in H$, let (x_n) be generated by*

$$x_{n+1} = J_{\gamma_n}(\lambda_n u + (1 - \lambda_n)(x_n + e_n)), \quad \text{for all } n \geq 0, \tag{6}$$

where $\lambda_n \in (0, 1)$ and $\gamma_n \in (0, \infty)$. If $(1 - \lambda_n)\|e_n\| / \lambda_n \leq C$ for some $C > 0$, then (x_n) is bounded.

Proof. Denote $w_n := \lambda_n u + (1 - \lambda_n)(x_n + e_n)$. Then for each $p \in A^{-1}(0)$, we have from (6) and the fact that the resolvent operator is nonexpansive

$$\begin{aligned} \|x_{n+1} - p\| &= \|J_{\gamma_n} w_n - J_{\gamma_n} p\| \\ &\leq \|w_n - p\|. \end{aligned}$$

Therefore, it would suffice to show that the sequence (w_n) is bounded. Note that

$$w_{n+1} = \lambda_{n+1} u + (1 - \lambda_{n+1})e_{n+1} + (1 - \lambda_{n+1})J_{\gamma_n} w_n, \quad \text{for all } n \geq 0. \tag{7}$$

Let $M > 0$ be big enough such that for some $p \in F$

$$\|u-p\| + \frac{(1-\lambda_n)}{\lambda_n}\|e_n\| \le M \quad \text{and} \quad \|w_0 - p\| \le 2M, \quad \text{for all } n \ge 0.$$

Then from (7), we see that

$$\begin{aligned} \|w_{n+1}-p\|^2 &= \|\lambda_{n+1}(u-p) + (1-\lambda_{n+1})e_{n+1} + (1-\lambda_{n+1})(J_{\gamma_n}w_n - p)\|^2 \\ &\le (1-\lambda_{n+1})^2\|J_{\gamma_n}w_n - p\|^2 \\ &+ 2\lambda_{n+1}\left\langle u-p+\frac{(1-\lambda_{n+1})}{\lambda_{n+1}}e_{n+1}, w_{n+1}-p\right\rangle \\ &\le (1-\lambda_{n+1})^2\|w_n-p\|^2 + 2M\lambda_{n+1}\|w_{n+1}-p\|, \end{aligned} \tag{8}$$

where the first inequality follows from the subdifferential inequality

$$\|x+y\|^2 \le \|y\|^2 + 2\langle x, x+y\rangle \quad \text{for all } x, y \in H$$

and the last inequality follows from the fact that the resolvent operator is nonexpansive. Now, if we assume that

$$\|w_n - p\| \le 2M \tag{9}$$

for some $n \in \mathbb{N}_0$, then from (8), we have

$$\begin{aligned} (\|w_{n+1}-p\| - \lambda_{n+1}M)^2 &\le 4(1-\lambda_{n+1})^2M^2 + \lambda_{n+1}^2M^2 \\ &\le (2(1-\lambda_{n+1})M + \lambda_{n+1}M)^2 \\ &= (2M - \lambda_{n+1}M)^2, \end{aligned}$$

showing that (9) also holds true for $n+1$. Therefore, we conclude that (w_n) is bounded. □

We are now in a position to prove our main result

Theorem 3. *Assume that $A : D(A) \subset H \to 2^H$ is a maximal monotone operator with $F := A^{-1}(0) \neq \emptyset$. For any fixed $x_0, u \in H$, let the sequence (x_n) be generated by (6) where $\lambda_n \in (0,1)$ and $\gamma_n \in (0,\infty)$ for all $n \ge 0$. If (e_n) is bounded, $\lambda_n \to 1$ and $\gamma_n \to \infty$ as $n \to \infty$, then (x_n) converges strongly to $P_F u$, the metric projection of u on F.*

Proof. We know from Lemma 2 that (x_n) is bounded. Moreover, from (6), we have

$$\begin{aligned} \|x_{n+1} - P_F u\| &\le \|x_{n+1} - J_{\gamma_n}u\| + \|J_{\gamma_n}u - P_F u\| \\ &\le (1-\lambda_n)\|x_n - u + e_n\| + \|J_{\gamma_n}u - P_F u\| \end{aligned}$$

where the second inequality follows from the fact that the resolvent operator is nonexpansive. The result follows immediately on passing to the limit in the above inequality. □

Remark 4. We point out that for (e_n) bounded and $\lambda_n \to 1$ as $n \to \infty$, algorithms (6) and (7) are not equivalent. Note that under these assumptions and $A^{-1}(0) \neq \emptyset$, Lemma 2 guarantees that the sequence (w_n) defined above is bounded. Consequently, for any $u \notin A^{-1}(0) =: F$, (w_n) converges strongly to $u \neq P_F u$, whereas (x_n) generated from (6) converges strongly to $P_F u$ (see Theorem 3 above).

Remark 5. In [2], it was proved that if $F := A^{-1}(0) \neq \emptyset$, then the sequence generated by algorithm (6) converges strongly to $P_F u$, provided that $\lambda_n \to 0$ with $\sum_{n=0}^{\infty} \lambda_n = \infty$, $\gamma_n \to \infty$ and $\|e_n\|/\lambda_n \to 0$. We note that since $\lambda_n \to 0$ as $n \to \infty$ in [2, Theorem 1], for very large n, the argument of the resolvent operator in (6) becomes arbitrarily close to that of algorithm (2). In this paper, we have employed a different approach: instead of the assumption $\lambda_n \to 0$, we have used the condition $\lambda_n \to 1$ as $n \to \infty$ which allows the argument of the resolvent operator in (6) to tend to the point u as $n \to \infty$. This approach is based on the fact that for very large values of t, $(I + tA)^{-1}$ approximates the zero of A which is closest to u. Theorem 3 is a strong theoretical result, under the classic boundedness condition on errors. It remains to see whether computationally, the speed/rate of convergence of algorithm (6) is fast enough.

References

[1] H. H. Bauschke, J. V. Burke, F. R. Deutsch, H. S. Hundal and J. D. Vanderwerff, *A new proximal point iteration that convergences weakly but not in norm*, Proc. Amer. Math. Soc. 133 (2005), no. 6, 1829-1835.

[2] O. A. Boikanyo and G. Moroşanu, *A proximal point algorithm converging strongly for general errors*, Optim. Lett. 4 (2010), 635-641.

[3] R. E. Bruck Jr., *A strongly convergent iterative solution of* $0 \in U(x)$ *for a maximal monotone operator* U *in Hilbert space,* J. Math. Anal. Appl. 48 (1974), 114-126.

[4] O. Güler, *On the convergence of the proximal point algorithm for convex minimization*, SIAM J. Control Optim. 29 (1991), 403-419.

[5] S. Kamimura and W. Takahashi, *Approximating solutions of maximal monotone operators in Hilbert spaces*, J. Approx. Theory 106 (2000), 226-240.

[6] N. Lehdili and A. Moudafi, *Combining the proximal algorithm and Tikhonov regularization,* Optimization 37 (1996), 239-252.

[7] B. Martinet, *Régularisation d'inéquations variationnelles par approximations successives*, Rev. Française Informat. Recherche Opérationnelle 4 (1970), Ser. R-3, 154–158.

[8] G. Moroşanu, *Asymptotic behaviour of resolvent for a monotone set in a Hilbert space*, Atti Accad. Nsz. Lincei 61 (1977), 565-570.

[9] ———, *Nonlinear Evolution Equations and Applications*, Reidel, Dordrecht, 1988.

[10] G. J. Minty, *Monotone (nonlinear) operators in Hilbert space*, Duke Math. J. 29 (1962), 341-348.

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

[12] H. K. Xu, *Iterative algorithms for nonlinear operators,* J. London Math. Soc. (2) 66 (2002), 240-256.

[13] ———, *A regularization method for the proximal point algorithm,* J. Glob. Optim. 36 (2006), 115-125.

Department of Mathematics and its Applications, Central European University, Nador u. 9, H-1051 Budapest, Hungary

E-mail address: boikanyo_oganeditse@ceu-budapest.edu, and morosanug@ceu.hu