Optim Lett (2013) 7:415–420
DOI 10.1007/s11590-011-0418-8

SHORT COMMUNICATION

# Strong convergence of a proximal point algorithm with bounded error sequence

**Oganeditse A. Boikanyo · Gheorghe Moroşanu**

Received: 21 August 2011 / Accepted: 17 October 2011 / Published online: 1 November 2011


**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 (Boikanyo and Moroşanu in Optim Lett 4(4):635–641, 2010), 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 $\varphi$ which is nearest to $u$.

**Keywords** Maximal monotone operator · Nonexpansive map · Proximal point algorithm · Prox-Tikhonov method · Resolvent operator

## 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:

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

O. A. Boikanyo (✉)
Department of Mathematics, University of Botswana, Private Bag 00704, Gaborone, Botswana
e-mail: boikanyoa@gmail.com

G. Moroşanu
Department of Mathematics and its Applications, Central European University, Nador u. 9, Budapest 1051, Hungary
e-mail: morosanug@ceu.hu

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 [12] 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 $t > 0$. The map $J_t$ is called the resolvent of $A$. For information on monotone operators, including subdifferentials of convex functions, and related problems, we refer the reader, e.g., to Borwein [4], Moroşanu [11] and Pardalos et al. [13].

One of the most effective iterative methods for solving problem (1) is the proximal point algorithm (PPA) which was initiated by Martinet [9] and further developed by Rockafellar [14]. 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 weakly converges in general [to a solution of problem (1)], see [1,6], modifying the PPA in order to make it strongly convergent has been the subject of intense research. Recently, Lehdili and Moudafi [8] 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 $\alpha_n$ and $\gamma_n$, 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 [16] 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 vectors $x_0, u \in H$, $\gamma_n \in (0, \infty)$, $\lambda_n \in (0, 1)$, and $(e_n)$ is a sequence of computational errors. 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 [3] for details. The scheme (5) (with $u := x_0$, the starting point of the PPA) was independently shown by Xu [15] and Kamimura and Takahashi [7] to be strongly convergent to a point of $A^{-1}(0)$ (if this set is not empty) which is nearest to $x_0$, provided that $\lambda_n \to 0$ with $\sum_{n=0}^{\infty} \lambda_n = \infty$, $(\|e_n\|)$ is summable and $\gamma_n \to \infty$. The case when $u$ is not necessarily the starting point of the PPA was first considered in [2]. In [3], 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 the algorithm (6) below. 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 algorithm (6) 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 [5] and Moroşanu [10].

**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 $u$ onto $F$.*

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 large enough such that for some $p \in F$

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

Then from (7), we see that

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

where the first inequality follows from the following obvious 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

$$\|w_n-p\| \le 2M, \qquad (9)$$

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 1** *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 $u$ onto $F$.*

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

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

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 1* 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 the aforementioned 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 1 above).

*Remark 2* In [3], 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 [3, Theorem 1], for 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 large values of $t$, $(I + tA)^{-1}$ approximates the zero of $A$ which is closest to $u$.

## 3 Final comments

Theorem 1 is a strong theoretical result, under the classic boundedness condition on errors. By performing simulations on particular examples, we can see that computationally the rate of convergence of algorithm (6) is fast enough. Let us explain in what follows how algorithm (6) functions when performing simulations.

Assume that we work under the assumptions of Theorem 1 and, for the sake of simplicity, $A$ is single-valued. Starting with a given $x_0$ we compute $x_1$ by solving inexactly for $x$ the equation

$$(I + \gamma_0 A)x = \lambda_0 u + (1 - \lambda_0)x_0$$

and get $x_1 + e_1$, instead of the exact solution $x_1$. We do not have any error for $x_0$ (i.e., $e_0 = 0$), but we have a computational error $e_1$ corresponding to $x_1$. Next, we solve the equation

$$(I + \gamma_1 A)x = \lambda_1 u + (1 - \lambda_1)(x_1 + e_1)$$

and get $x_2 + e_2$ instead of the exact solution $x_2$. The previous error $e_1$ is multiplied by $(1 - \lambda_1)$ as required by the iterative process, and at this step we have a new error $e_2$ associated with $x_2$. Using $x_2 + e_2$ we then compute approximately $x_3$, i.e., we get $x_3 + e_3$, and so on. In fact, on the computer we obtain the sequence $z_n = x_n + e_n$, satisfying

$$z_{n+1} = (I + \gamma_n A)^{-1}(\lambda_n u + (1 - \lambda_n)z_n) + e_{n+1}, \quad n = 0, 1, \ldots,$$

where $z_0 = x_0$. It is obvious that if $\|e_n\| \leq \varepsilon$ $(n = 0, 1, \ldots)$, for an $\varepsilon$ small enough, then $\|z_n - x_n\| \leq \varepsilon$, thus $z_n$ approximates $P_F u$ for $n$ large enough (cf. Theorem 1). This is the practical meaning of our iterative process.

**Acknowledgments** We thank the anonymous referees for their valuable comments and suggestions.

## References


1. Bauschke, H.H., Burke, F.R., Deutsch, J.V., Hundal, H.S., Vanderwerff, J.D.: A new proximal point iteration that convergences weakly but not in norm. Proc. Am. Math. Soc. **133**(6), 1829–1835 (2005)
2. Boikanyo, O.A., Moroşanu, G.: Modified Rockafellar's algorithms. Math. Sci. Res. J. **5**(13), 101–122 (2009)
3. Boikanyo, O.A., Moroşanu, G.: A proximal point algorithm converging strongly for general errors. Optim. Lett. **4**(4), 635–641 (2010)
4. Borwein, J.M.: Fifty years of maximal monotonicity. Optim. Lett. **4**(4), 473–490 (2010)
5. Bruck, R.E. 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**, 114–126 (1974)
6. Güler, O.: On the convergence of the proximal point algorithm for convex minimization. SIAM J. Control Optim. **29**, 403–419 (1991)
7. Kamimura, S., Takahashi, W.: Approximating solutions of maximal monotone operators in Hilbert spaces. J. Approx. Theory **106**, 226–240 (2000)
8. Lehdili, N., Moudafi, A.: Combining the proximal algorithm and Tikhonov regularization. Optimization **37**, 239–252 (1996)
9. Martinet, B.: Régularisation d'inéquations variationnelles par approximations successives. Rev. Française Informat. Recherche Opérationnelle **4**(Ser. R-3), 154–158 (1970)
10. Moroşanu, G.: Asymptotic behaviour of resolvent for a monotone set in a Hilbert space. Atti Accad. Nsz. Lincei **61**, 565–570 (1977)
11. Moroşanu, G.: Nonlinear Evolution Equations and Applications. Reidel, Dordrecht (1988)
12. Minty, G.J.: Monotone (nonlinear) operators in Hilbert space. Duke Math. J. **29**, 341–348 (1962)
13. Pardalos, P.M., Rassias, T.M., Khan, A.A.: Nonlinear Analysis and Variational Problems. Springer, Berlin (2010)
14. Rockafellar, R.T.: Monotone operators and the proximal point algorithm. SIAM J. Control Optim. **14**, 877–898 (1976)
15. Xu, H.K.: Iterative algorithms for nonlinear operators. J. Lond. Math. Soc. **2**(66), 240–256 (2002)
16. Xu, H.K.: A regularization method for the proximal point algorithm. J. Glob. Optim. **36**, 115–125 (2006)