Optim Lett (2010) 4:635–641
DOI 10.1007/s11590-010-0176-z

SHORT COMMUNICATION

A proximal point algorithm converging strongly for general errors

O. A. Boikanyo · G. Moroşanu

Received: 26 November 2009 / Accepted: 18 January 2010 / Published online: 5 February 2010

Abstract In this paper a proximal point algorithm (PPA) for maximal monotone operators with appropriate regularization parameters is considered. A strong convergence result for PPA is stated and proved under the general condition that the error sequence tends to zero in norm. Note that Rockafellar (SIAM J Control Optim 14:877–898, 1976) assumed summability for the error sequence to derive weak convergence of PPA in its initial form, and this restrictive condition on errors has been extensively used so far for different versions of PPA. Thus this Note provides a solution to a long standing open problem and in particular offers new possibilities towards the approximation of the minimum points of convex functionals.

Keywords Proximal point algorithm · Prox-Tikhonov algorithm · Monotone operator · Regularization parameters · Strong convergence

1 Introduction

Let H be a real Hilbert space with scalar product $\langle \cdot, \cdot \rangle$ and the Hilbertian norm $\| \cdot \|$. Recall that a mapping (possibly multivalued) $A : D(A) \subset H \to H$ is said to be a monotone operator if

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

O. A. Boikanyo · G. Moroşanu (✉)
Department of Mathematics and its Applications, Central European University,
Nador u. 9, Budapest 1051, Hungary
e-mail: morosanug@ceu.hu

O. A. Boikanyo
e-mail: boikanyo_oganeditse@phd.ceu.hu

In other words, A is a monotone subset of $H \times H$. If A is monotone, then its inverse, defined by $A^{-1} = \{(y, x) : (x, y) \in A\}$, is clearly a monotone operator too. A monotone operator A is said to be maximal monotone if, regarded as a subset of $H \times H$, is not properly contained in any other monotone subset of $H \times H$. It is clear that A is maximal monotone if and only if A^{-1} is so. It is well known that if A is maximal monotone and $\beta > 0$, then the resolvent of A, the operator $J_\beta^A : H \to H$ defined by $J_\beta^A(x) = (I + \beta A)^{-1}(x)$, is single-valued and nonexpansive; i.e, for all $x, y \in H$, $\|J_\beta^A(x) - J_\beta^A(y)\| \le \|x - y\|$.

Recently, Xu [7] investigated a modified version of the initial PPA studied by Rockafellar [4], in the sense that the $(n + 1)$th iterate is given by

$$x_{n+1} = \alpha_n x_0 + (1 - \alpha_n) J_{\beta_n}^A(x_n) + e_n, \quad n \ge 0, \tag{1}$$

where x_0 is the starting point of PPA and $\{e_n\}$ is the error sequence. For $\{e_n\}$ summable, it was proved that $\{x_n\}$ is strongly convergent if $\beta_n \to \infty$, and $\alpha_n \in (0, 1)$ with $\alpha_n \to 0$, $\sum_{n=0}^{\infty} \alpha_n = \infty$. In [1], we generalized this algorithm by noting that one does not necessarily have to take x_0 as the initial/starting point of PPA (i.e., any point of H works) and proved a strong convergence result associated with this slightly modified algorithm when the error sequence is in ℓ^p for $1 \le p < 2$. Although our condition on the errors is weaker than summability, it is still restrictive.

We seek conditions which will guarantee strong convergence of $\{x_n\}$ for more general errors. Lehdili and Moudafi [2] introduced the so called prox-Tikhonov method which generates the sequence $\{x_n\}$ defined by the algorithm

$$x_{n+1} = J_{\beta_n}^{A_n}(x_n), \quad n \ge 0, \tag{2}$$

where $\beta_n > 0$, and A_n is the (strongly monotone) operator defined by $A_n = \mu_n I + A$, for $\mu_n > 0$. This operator is usually regarded as a Tikhonov regularization of A. Xu [6] proposed the following regularization for the proximal point algorithm

$$x_{n+1} = J_{\beta_n}^A((1 - \alpha_n)x_n + \alpha_n u + e_n), \quad \text{for any } u \in H \text{ and } n \ge 0, \tag{3}$$

which essentially includes algorithm 2 as a special case, as noticed by Xu himself. It is also easy to see that algorithms (5) and (3) are in fact equivalent.

The main purpose of this paper is to improve our earlier result [1, Theorem 5] to cover general errors. In fact our main result shows that strong convergence of PPA is preserved under the assumption that the error sequence converges to zero in norm, if we choose appropriate regularization parameters (see Sect. 3). That is important from the computational point of view.

2 Main result

Consider again algorithm (3) with $\beta_n > 0$, $\alpha_n \in (0,1)$ and $\emptyset \neq F := A^{-1}(0)$. Denoting

$$y_n := (1-\alpha_n)x_n + \alpha_n u + e_n,$$

one can rewrite (3) as

$$y_{n+1} = (1-\alpha_{n+1})J_{\beta_n}^A(y_n) + \alpha_{n+1}u + e_{n+1}, \quad n \geq 0. \tag{4}$$

If we assume that $\alpha_n \to 0$ and $\|e_n\| \to 0$ then obviously $\{x_n\}$ is convergent if and only if $\{y_n\}$ is so and their limits coincide, showing that (3) and (4) are equivalent. In fact, re-denoting $x_n := y_n$, $\alpha_n := \alpha_{n+1}$ and $e_n := e_{n+1}$, algorithm (4) reads

$$x_{n+1} = \alpha_n u + (1-\alpha_n)J_{\beta_n}^A(x_n) + e_n, \quad n \geq 0, \tag{5}$$

which is exactly the algorithm proposed and studied in [1].

In order to formulate and prove our main result we recall the following lemma due to Xu [7].

Lemma 1 [7]. *Let $\{s_n\}$ be a sequence of non-negative real numbers satisfying*

$$s_{n+1} \leq (1-a_n)s_n + a_n b_n + c_n, \quad n \geq 0,$$

where $\{a_n\}$, $\{b_n\}$, $\{c_n\}$ satisfy the conditions:

(i) $\{a_n\} \subset [0,1]$, $\sum_{n=0}^{\infty} a_n = \infty$,
(ii) $\limsup_{n\to\infty} b_n \leq 0$,
(iii) $c_n \geq 0$ *for all* $n \geq 0$ *with* $\sum_{n=0}^{\infty} c_n < \infty$.

Then $\lim_{n\to\infty} s_n = 0$.

We now state our main result.

Theorem 1 *Assume that $\alpha_n \in (0,1)$ with $\alpha_n \to 0$, $\sum_{n=0}^{\infty} \alpha_n = \infty$ and*

(i) *either* $\sum_{n=0}^{\infty} \|e_n\| < \infty$,
(ii) *or* $\|e_n\|/\alpha_n \to 0$.

If A is maximal monotone and $F := A^{-1}(0) \neq \emptyset$, then the sequence $\{x_n\}$ generated by algorithm (5) converges strongly to $q = P_F u$ provided $\beta_n \to \infty$, where P_F denotes the projection of u on F.

Proof We divide the proof into three steps.

Step 1: Boundedness of $\{x_n\}$:

If $\sum_{n=0}^{\infty} \|e_n\| < \infty$, then as in [6], it can be shown by induction that

$$\|x_n - p\| \leq \max\{\|x_0 - p\|, \|u - p\|\} + \sum_{k=0}^{n-1} \|e_k\| \quad \text{for any } p \in F \text{ and } n \geq 0. \tag{6}$$

Hence $\{x_n\}$ is bounded (see also [1]).

Now assume that $\{e_n/\alpha_n\}$ is bounded. Then, there exists a positive constant M such that

$$\|u - p\| + \frac{\|e_n\|}{\alpha_n} \leq M \quad \text{for some } p \in F \text{ and } n \geq 0.$$

We assume M is big enough so that $\|x_0 - p\| \leq C := 2M$. Let us show that

$$\|x_n - p\| \leq C \quad \text{for all } n \geq 0. \tag{7}$$

Using (5) and applying the subdifferential inequality

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

we get

$$\begin{aligned}
\|x_{n+1} - p\|^2 &= \left\|\alpha_n \left(u - p + \frac{e_n}{\alpha_n}\right) + (1 - \alpha_n)(J_{\beta_n}^A(x_n) - p)\right\|^2 \\
&\leq (1 - \alpha_n)^2 \|J_{\beta_n}^A(x_n) - p\|^2 + 2\alpha_n \left\langle u - p + \frac{e_n}{\alpha_n}, x_{n+1} - p\right\rangle \\
&\leq (1 - \alpha_n)^2 \|x_n - p\|^2 + 2M\alpha_n \|x_{n+1} - p\|.
\end{aligned}$$

If $\|x_n - p\| \leq C$ for some $n \geq 0$, then the last estimate gives

$$\|x_{n+1} - p\|^2 \leq (1 - \alpha_n)^2 C^2 + 2M\alpha_n \|x_{n+1} - p\|.$$

Hence,

$$(\|x_{n+1} - p\| - M\alpha_n)^2 \leq M^2\alpha_n^2 + (1 - \alpha_n)^2 C^2,$$

which yields

$$\|x_{n+1} - p\| \leq M\alpha_n + \sqrt{M^2\alpha_n^2 + (1 - \alpha_n)^2 C^2}.$$

On the other hand, it is easy to check that

$$M\alpha_n + \sqrt{M^2\alpha_n^2 + (1 - \alpha_n)^2 C^2} \leq C.$$

Step 2: We want to show that

$$\overline{\lim_{n\to\infty}}\langle u-q, x_n-q\rangle \le 0.$$

Take a subsequence $\{x_{n_k}\}$ of $\{x_n\}$ so that

$$\overline{\lim_{n\to\infty}}\langle u-q, x_n-q\rangle = \lim_{k\to\infty}\langle u-q, x_{n_k}-q\rangle.$$

Since $\{x_n\}$ is bounded, $\{x_{n_k}\}$ converges weakly on a subsequence, again denoted by $\{x_{n_k}\}$, to some x_∞. Then it follows that

$$\overline{\lim_{n\to\infty}}\langle u-q, x_n-q\rangle = \langle u-q, x_\infty-q\rangle.$$

Since $q = P_F u$ it is sufficient to show that $x_\infty \in F$. Note that in both cases $\|e_n\| \to 0$ and

$$y_n := \frac{1}{1-\alpha_n}(x_{n+1}-\alpha_n u - e_n) = J^A_{\beta_n}(x_n)$$

implies that

$$A(y_{n_k-1}) \ni \frac{1}{\beta_{n_k-1}}(x_{n_k-1}-y_{n_k-1}) \to 0.$$

Since $y_{n_k-1} \rightharpoonup x_\infty$ and A is demiclosed, we have $x_\infty \in F$.

Step 3: Now we show that $\{x_n\}$ converges strongly to $q = P_F u$. We have as before

$$\|x_{n+1}-q\|^2 \le (1-\alpha_n)\|x_n-q\|^2 + 2\alpha_n\left\langle u-q+\frac{e_n}{\alpha_n}, x_{n+1}-q\right\rangle. \tag{8}$$

If $\sum_{n=0}^\infty \|e_n\| < \infty$, then we derive from (8)

$$\|x_{n+1}-q\|^2 \le (1-\alpha_n)\|x_n-q\|^2 + \alpha_n b_n + c_n$$

where $b_n = 2\langle u-q, x_{n+1}-q\rangle$ and $c_n = K\|e_n\|$ for some $K > 0$. In the case $\|e_n\|/\alpha_n \to 0$, we have

$$\|x_{n+1}-q\|^2 \le (1-\alpha_n)\|x_n-q\|^2 + \alpha_n b_n,$$

where $b_n = 2\langle u-q+e_n/\alpha_n, x_{n+1}-q\rangle$. In either case Step 2 and Lemma 1 implies that $x_n \to q$ as desired. □

3 Concluding remarks

If $\|e_n\|/\alpha_n \to 0$ and $\sum_{n=0}^{\infty}\|e_n\| = \infty$, then automatically $\sum_{n=0}^{\infty}\alpha_n = \infty$. This shows that assumption *(ii)* covers the case when the error sequence is not summable. More precisely, we have

Corollary 1 *Assume that A is a maximal monotone operator with $F := A^{-1}(0) \neq \emptyset$, $\|e_n\| \to 0$ and $\sum_{n=0}^{\infty}\|e_n\| = \infty$. Then one can choose an appropriate sequence $\{\alpha_n\} \subset (0,1)$ such that the sequence $\{x_n\}$ generated by algorithm (5) converges strongly to $q = P_F u$ provided $\beta_n \to \infty$, where P_F denotes the projection of u on F.*

Proof One can take, e.g., $\alpha_n = \sqrt{\|e_n\|}$ if $e_n \neq 0$ and n is large enough, and $\alpha_n = 1/(n+2)$ otherwise. Obviously, $\alpha_n \in (0,1)$ for all n, $\alpha_n \to 0$, $\sum_{n=0}^{\infty}\alpha_n = \infty$, and condition (ii) of Theorem 1 is satisfied. This concludes the proof. □

To summarize, let us point out that practically every sequence $\{e_n\}$ converging strongly to zero is good to obtain strong convergence of $\{x_n\}$. Indeed, according to Theorem 1, if $\sum_{n=0}^{\infty}\|e_n\| < \infty$, then we can choose freely $\alpha_n \in (0,1)$ with $\alpha_n \to 0$ and $\sum_{n=0}^{\infty}\alpha_n = \infty$. Otherwise (i.e., if $\sum_{n=0}^{\infty}\|e_n\| = \infty$), we can choose for example $\{\alpha_n\}$ as in the proof of the above corollary to obtain strong convergence of $\{x_n\}$. (Of course this choice is not unique.) This conclusion is extremely important from the computational point of view. Note that in fact in the case *(ii)* of Theorem 1, which is not covered by the existing related results, we can build a good PPA by choosing appropriate regularization parameters α_n to keep the strong convergence of $\{x_n\}$ under the general condition $\|e_n\| \to 0$. These α_n depend on errors, but this is acceptable from the numerical point of view.

It is worth pointing out that in the particular case when A is the subdifferential of a proper, convex, lower semicontinuous function $\phi : H \longrightarrow (-\infty, +\infty]$, our result offers a reliable algorithm generating a sequence which approximates a minimum point of ϕ, provided that the error sequence converges to zero in norm. Indeed, in this case any point of $F = A^{-1}0$, in particular $P_F u$, is a minimum point of ϕ. See, e.g., [3, pp. 32–46].

Note however that this approach works under the assumption $\beta_n \to \infty$. The so-called regularization method developed in [5,6] (which in fact leads to an algorithm equivalent to ours, as pointed out before) works under different assumptions on β_n, but at the expense of the restrictive summability condition for the error sequence.

Therefore, the following open question arises: can one design a PPA by choosing appropriate regularization parameters α_n such that strong convergence of $\{x_n\}$ is preserved, for $\|e_n\| \to 0$ and β_n bounded? The answer seems to be positive, but probably requires more advanced analysis.

References

1. Boikanyo, O.A., Moroşanu, G.: Modified Rockafellar's algorithms. Math. Sci. Res. J. **13**(5), 101–122 (2009)
2. Lehdili, N., Moudafi, A.: Combining the proximal point for convex optimization. Optimization **37**, 239–252 (1996)

3. Moroşanu, G.: Nonlinear Evolution Equations and Applications. D Reidel, Dordrecht (1988)
4. Rockafellar, R.T.: Monotone operators and the proximal point algorithm. SIAM J. Control Optim. **14**, 877–898 (1976)
5. Song, Y., Yang, C.: A note on the paper "A regularization method for the proximal point algorithm". J. Glob. Optim. **43**, 171–174 (2009)
6. Xu, H.K.: A regularization method for the proximal point algorithm. J. Glob. Optim. **36**, 115–125 (2006)
7. Xu, H.K.: Iterative algorithms for nonlinear operators. J. Lond. Math. Soc. **66**(2), 240–256 (2002)