A hybrid inertial projection-proximal method for variational inequalities

Abdellatif Moudafi

Journal of Inequalities in Pure and Applied Mathematics

A HYBRID INERTIAL PROJECTION-PROXIMAL METHOD FOR VARIATIONAL INEQUALITIES

A. MOUDAFI

Univeristé des Antilles et de la Guyane
DSI-GRIMAAG, B.P. 7209, 97275 Schoelcher,
Martinique, France.

EMail: abdellatif.moudafi@martinique.univ-ag.fr

ISSN (electronic): 1443-5756
170-03

Abstract

The hybrid proximal point algorithm introduced by Solodov and Svaiter allowing significant relaxation of the tolerance requirements imposed on the solution of proximal subproblems will be combined with the inertial method introduced by Alvarez and Attouch which incorporates second order information to achieve faster convergence. The weak convergence of the resulting method will be investigated for finding zeroes of a maximal monotone operator in a Hilbert space.

2000 Mathematics Subject Classification: Primary, 49J53, 65K10; Secondary, 49M37, 90C25.

Key words: Maximal monotone operators, Weak convergence, Proximal point algorithm, Hybrid projection-proximal method, Resolvent, inertial proximal point scheme.

Contents

1. Introduction and Preliminaries

The theory of maximal monotone operators has emerged as an effective and powerful tool for studying a wide class of unrelated problems arising in various branches of social, physical, engineering, pure and applied sciences in unified and general framework. In recent years, much attention has been given to develop efficient and implementable numerical methods including the projection method and its variant forms, auxiliary problem principle, proximal-point algorithm and descent framework for solving variational inequalities and related optimization problems. It is well known that the projection method and its variant forms cannot be used to suggest and analyze iterative methods for solving variational inequalities due to the presence of the nonlinear term. This fact motivated the development of another technique which involves the use of the resolvent operator associated with maximal monotone operators, the origin of which can be traced back to Martinet [4] in the context of convex minimization and Rockafellar [8] in the general setting of maximal monotone operators. The resulting method, namely the proximal point algorithm has been extended and generalized in different directions by using novel and innovative techniques and ideas, both for their own sake and for their applications relying on the Bregman distance or based on the variable metric approach.

To begin with let us recall the following concepts which are of common use in the context of convex and nonlinear analysis, see for example Brézis [3]. Throughout, $\mathcal{H}$ is a real Hilbert space, $\langle \cdot, \cdot \rangle$ denotes the associated scalar product and $\| \cdot \|$ stands for the corresponding norm. An operator is said to be monotone if

$$\langle u - v, x - y \rangle \geq 0 \quad \text{whenever} \quad u \in A(x), v \in A(y).$$

It is said to be maximal monotone if, in addition, the graph, $\{(x,y) \in \mathcal{H} \times \mathcal{H} : y \in A(x)\}$, is not properly contained in the graph of any other monotone operator. It is well-known that for each $x \in \mathcal{H}$ and $\lambda > 0$ there is a unique $z \in \mathcal{H}$ such that $x \in (I + \lambda A)z$. The single-valued operator $J_\lambda^A := (I + \lambda A)^{-1}$ is called the resolvent of A of parameter λ. It is a nonexpansive mapping which is everywhere defined and satisfies: $z = J_\lambda^A z$, if and only if, $0 \in Az$.

In this paper we will focus our attention on the classical problem of finding a zero a maximal monotone operators A on a real Hilbert space $\mathcal{H}$

$$\text{find} \quad x \in \mathcal{H} \quad \text{such that} \quad A(x) \ni 0. \tag{1.1}$$

One of the fondamental approaches to solving (1.1) is the proximal method proposed by Rockafellar [8]. Specifically, having $x_n \in \mathcal{H}$ a current approximation to the solution of (1.1), the proximal method generated the next iterate by solving the proximal subproblem

$$0 \in A(x) + \mu_n(x - x_n), \tag{1.2}$$

where $\mu_n > 0$ is a regularization parameter.

Because solving (1.2) exactly can be as difficult as solving the original problem itself, it is of practical relevance to solve the subproblems approximately, that is find $x_{n+1} \in \mathcal{H}$ such that

$$0 = v_{n+1} + \mu_n(x_{n+1} - x_n) + \varepsilon_n, \quad v_{n+1} \in A(x_{n+1}), \tag{1.3}$$

where $\varepsilon_n \in \mathcal{H}$ is an error associated with inexact solution of subproblem (1.2).

In many applications proximal point methods in the classical form are not very efficient. Developments aimed at speeding up the convergence of proximal methods focus, among other approaches, on the ways of incorporating second order information to achieve faster convergence. To this end, Alvarez and Attouch proposed an inertial method obtained by discretization of a second-order (in time) dissipative dynamical system. Also, it is worth developing new algorithms which admit less stringent requirements on solving the proximal subproblems. Solodov and Zvaiter followed suit and showed that the tolerance requirements for solving the subproblems can be significantly relaxed if the solving of each subproblem is followed by a projection onto a certain hyperplane which separates the current iterate from the solution set of the problem.

To take advantage of the two approaches, we propose a method obtained by coupling the two previous algorithms.

Specifically, we introduce the following method.

Algorithm 1.1. *Choose any* $x_0, x_1 \in \mathcal{H}$ *and* $\sigma \in [0,1[$. *Having* x_n, *choose* $\mu_n > 0$ *and*

$$\text{(1.4)} \quad \textit{find } y_n \in \mathcal{H} \quad \textit{such that} \quad 0 = v_n + \mu_n(y_n - z_n) + \varepsilon_n, \quad v_n \in A(y_n),$$

where

$$\text{(1.5)} \quad z_n := x_n + \alpha_n(x_n - x_{n-1}) \quad \textit{and} \quad \|\varepsilon_n\| \leq \sigma \max\{\|v_n\|, \mu_n\|y_n - z_n\|\}.$$

Stop if $v_n = 0$ *or* $y_n = z_n$. *Otherwise, let*

$$x_{n+1} = z_n - \frac{\langle v_n, z_n - y_n\rangle}{\|v_n\|^2} v_n. \tag{1.6}$$

Note that the last equation amounts to

$$x_{n+1} = proj_{H_n}(z_n),$$

where

$$H_n := \{z \in \mathcal{H}, \langle v_n, z - y_n \rangle = 0\}. \tag{1.7}$$

Throughout we assume that the solution set of the problem (1.1) is nonempty.

2. Convergence Analysis

To begin with, let us state the following lemma which will be needed in the proof of the main convergence result.

Lemma 2.1. *([9, Lemma 2.1]). Let $x, y, v, \bar{x}$ be any elements of $\mathcal{H}$ such that*

$$\langle v, x - y\rangle > 0 \quad \textit{and} \quad \langle v, \bar{x} - y\rangle \leq 0.$$

Let $z = \mathit{proj}_H(x)$, where

$$H := \{s \in \mathcal{H}, \langle v, s - y\rangle = 0\}.$$

Then

$$\|x - \bar{x}\|^2 \leq \|x - \bar{x}\|^2 - \left(\frac{\langle v, x - y\rangle}{\|v\|}\right)^2.$$

We are now ready to prove our main convergence result.

Theorem 2.2. *Let $\{x_n\}$ be any sequence generated by our algorithm, where $A : \mathcal{H} \to \mathcal{P}(\mathcal{H})$ is a maximal monotone operator, and the parameters α_n, μ_n satisfy*

1. $\exists \bar{\mu} < +\infty$ *such that* $\mu_n \leq \bar{\mu}$.
2. $\exists \alpha \in [0, 1[$ *such that* $\forall k \in \mathbb{N}^*$ $0 \leq \alpha_k \leq \alpha$.

If the following condition holds

$$\sum_{n=1}^{\infty} \alpha_n \|x_{n-1} - x_n\|^2 < +\infty, \tag{2.1}$$

then, there exists $\bar{x} \in S := A^{-1}(0)$ such that the sequence $\{v_n\}$ strongly converges to zero and the sequence $\{x_n\}$ weakly converges to $\bar{x}$.

Proof. Suppose that the algorithm terminates at some iteration n. It is easy to check that $v_n = 0$ in other words $y_n \in S$. From now on, we assume that an infinite sequence of iterates is generated. It is also easy to see, using the monotonicity of A and the Cauchy-Schwarz inequality, that the hyperplane H_n, given by (1.7), strictly seperates x_n from any solution $\bar{x} \in S$. We are now in a position to apply Lemma 2.1, which gives

$$\|x_{n+1} - \bar{x}\|^2 \leq \|z_n - \bar{x}\|^2 - \frac{\langle v_n, z_n - y_n\rangle^2}{\|v_n\|^2}. \tag{2.2}$$

Setting $\varphi_n = \frac{1}{2}\|x_n - \bar{x}\|^2$ and taking in account the fact that

$$\frac{1}{2}\|z_n - \bar{x}\|^2 = \frac{1}{2}\|x_n - \bar{x}\|^2 + \alpha_n\langle x_n - \bar{x}, x_n - x_{n-1}\rangle + \frac{\alpha_n^2}{2}\|x_n - x_{n-1}\|^2,$$

and that

$$\langle x_n - \bar{x}, x_n - x_{n-1}\rangle = \varphi_n - \varphi_{n-1} + \frac{1}{2}\|x_n - x_{n-1}\|^2,$$

we derive

$$\varphi_{n+1} - \varphi_n \leq \alpha_n(\varphi_n - \varphi_{n-1}) + \frac{\alpha_n + \alpha_n^2}{2}\|x_n - x_{n-1}\|^2 - \frac{1}{2}\frac{\langle v_n, z_n - y_n\rangle^2}{\|v_n\|^2}.$$

On the other hand, using the same arguments as in the proof of Theorem 2.2 ([9]), we obtain that

$$\frac{\langle v_n, z_n - y_n\rangle}{\|v_n\|} \geq \frac{(1-\sigma)^2}{(1+\sigma)^4\mu_n^2}\|v_n\|^2. \tag{2.3}$$

Hence, from (2.2) it follows that

$$\varphi_{n+1} - \varphi_n \le \alpha_n(\varphi_n - \varphi_{n-1}) + \frac{\alpha_n + \alpha_n^2}{2}\|x_n - x_{n-1}\|^2 - \frac{1}{2}\cdot\frac{(1-\sigma)^2}{(1+\sigma)^4\mu_n^2}\|v_n\|^2,$$

from which we infer that

$$\varphi_{n+1} - \varphi_n \le \alpha_n(\varphi_n - \varphi_{n-1}) + \alpha_n\|x_n - x_{n-1}\|^2 - \frac{1}{2}\cdot\frac{(1-\sigma)^2}{(1+\sigma)^4\bar{\mu}^2}\|v_n\|^2. \tag{2.4}$$

Setting $\theta_n := \varphi_n - \varphi_{n-1}$, $\delta_n := \alpha_n\|x_n - x_{n-1}\|^2$ and $[t]_+ := \max(t, 0)$, we obtain

$$\theta_{n+1} \le \alpha_n\theta_n + \delta_n \le \alpha_n[\theta_n]_+ + \delta_n,$$

where $\alpha \in [0, 1[$.

The rest of the proof follows that given in [1] and is presented here for completeness and to convey the idea in [1]. The latter inequality yields

$$[\theta_{n+1}]_+ \le \alpha^n[\theta_1]_+ + \sum_{i=0}^{n-1}\alpha^i\delta_{n-i},$$

and therefore

$$\sum_{n=1}^{\infty}[\theta_{n+1}] \le \frac{1}{1-\alpha}\left([\theta_1]_+ + \sum_{n=1}^{+\infty}\delta_n\right),$$

which is finite thanks to the hypothesis of the theorem. Consider the sequence defined by $t_n := \varphi_n - \sum_{i=1}^{n}[\theta_i]_+$. Since $\varphi_n \ge 0$ and $\sum_{i=1}^{n}[\theta_i]_+ < +\infty$, it follows that $\{t_n\}$ is bounded from below. But

$$t_{n+1} = \varphi_{n+1} - [\theta_{n+1}]_+ - \sum_{i=1}^{n}[\theta_i]_+ \le \varphi_{n+1} - \varphi_{n+1} + \varphi_n - \sum_{i=1}^{n}[\theta_i]_+ = t_n,$$

so that $\{t_n\}$ is nonincreasing. We thus deduce that $\{t_n\}$ is convergent and so is $\{\varphi_n\}$. On the other hand, from (2.4), we obtain the following estimate

$$\frac{1}{2}\frac{(1-\sigma)^2}{(1+\sigma)^4\bar{\mu}^2}\|v_n\|^2 \leq \varphi_n - \varphi_{n+1} + \alpha[\theta_n]_+ + \delta_n.$$

Passing to the limit in the last inequality and taking into account that $\{\varphi_n\}$ converges, $[\theta_n]_+$ and δ_n go to zero as n tends to $+\infty$, we obtain that the sequence $\{v_n\}$ strongly converges to 0. Since, by (1.4),

$$\bar{\mu}^{-1}\|v_n\| \geq \|z_n - y_n\|,$$

we also have that the sequence $\{z_n - y_n\}$ strongly converges to 0.

Now let x^* be a weak cluster point of $\{x_n\}$. There exists a subsequence $\{x_\nu\}$, which weakly converges to x^*. According to the fact that

$$\lim_{\nu\to+\infty}\|z_\nu - y_\nu\| = 0 \quad \text{with} \quad z_\nu = x_\nu + \alpha_\nu(x_\nu - x_{\nu-1})$$

and in the light of assumption (2.1), it is clear that the sequences $\{z_\nu\}$ and $\{y_\nu\}$ also weakly converge to the weak cluster point x^*. By the monotonicity of A, we can write

$$\forall z \in \mathcal{H}\ \forall w \in A(z) \quad \langle z - y_\nu, w - v_\nu\rangle \geq 0,$$

Passing to the limit, as $\nu \to +\infty$, we obtain

$$\langle z - x^*, w\rangle \geq 0,$$

this being true for any $w \in A(z)$. From the maximal monotonicity of A, it follows that $0 \in A(x^*)$, that is $x^* \in S$. The desired result follows by applying the well-known Opial Lemma [7]. $\square$

3. Conclusion

In this paper we propose a new proximal algorithm obtained by coupling the hybrid proximal method with the inertial proximal scheme. The principal advantage of this algorithm is that it allows a more constructive error tolerance criterion in solving the inertial proximal subproblems. Furthermore, its second-order nature may be exploited in order to accelerate the convergence. It is worth mentioning that if $\sigma = 0$, the proposed algorithm reduces to the classical exact inertial proximal point method introduced in [2]. Indeed, $\sigma = 0$ implies that $\varepsilon_n = 0$, and consequently $x_{n+1} = y_n$. In this case, the presented analysis provides an alternative proof of the convergence of the exact inertial proximal method that permits an interesting geometric interpretation.

References

[1] F. ALVAREZ, On the minimizing property of a second order dissipative system in Hilbert space, *SIAM J. of Control and Optim.*, **14**(38) (2000), 1102–1119.

[2] F. ALVAREZ AND H. ATTOUCH, An inertial proximal method for maximal monotone operators via discretization of a non linear oscillator with damping, *Set Valued Analysis*, **9** (2001), 3–11.

[3] H. BRÉZIS, *Opérateurs maximaux monotones et semi-groupes des contractions dans les espaces de Hilbert*, North Holland, Amsterdam, 1973.

[4] B. MARTINET, Algorithmes pour la résolution des problèmes d'optimisation et de minmax, Thèse d'état Université de Grenoble, France (1972).

[5] A. MOUDAFI AND M. OLINY, Convergence of a splitting inertial proximal method for monotone operators, *Journal of Computational and Applied Mathematics*, **155** (2003), 447–454.

[6] A. MOUDAFI AND E. ELISABETH, Approximate inertial proximal methods using the enlargement of maximal monotone operators, *International Journal of Pure and Applied Mathematics*, **5**(3) (2003), 283–299.

[7] Z. OPIAL, Weak convergence of the sequence of successive approximations for nonexpansive mappings, *Bulletin of the American Mathematical Society*, **73** (1967), 591–597.

[8] R.T. ROCKAFELLAR, Monotone operator and the proximal point algorithm, *SIAM J. Control. Optimization,* **14**(5) (1976), 877–898.

[9] M.V. SOLODOV AND B.F. SVAITER, A hybrid projection-proximal point algorithm, *Journal of Convex Analysis*, **6**(1) (1999), 59–70.

[10] M.V. SOLODOV AND B.F. SVAITER, An inexact hybrid generalized proximal point algorithm and some new results on the theory of Bregman functions, *Mathematics of Operations Research*, to appear.