

ON THE CONVERGENCE OF THE FORWARD-BACKWARD ALGORITHM FOR NULL-POINT PROBLEMS

ABDELLATIF MOUDAFI

Aix Marseille Université CNRS-LSIS UMR 7296,13397, Marseille, France

Abstract. The aim of this paper is to investigate the asymptotic behavior of the forward-backward algorithm for solving null-point problems governed by two maximal monotone operators. An application to the split feasibility problem is also stated.

Keywords. Maximal monotone operator; Forward-backward algorithm; Split null point problem; Split feasibility problem; Error bound.

2010 Mathematics Subject Classification. 49J53, 65K10, 49M37.

1. INTRODUCTION AND PRELIMINARIES

Throughout, H is a real Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and induced norm $\| \cdot \|$. Notations not explicitly defined here are standard.

Recall that the graph, $gphA$, of a set-valued operator $A : H \rightarrow 2^H$ is given by

$$gphA = \{(x, y) \in H \times H; y \in A(x)\}.$$

Recall that the mapping A is monotone if $\langle x - x', y - y' \rangle \geq 0$, $\forall (x, y) \in gphA, \forall (x', y') \in gphA$. A is said to be maximal monotone if it cannot be properly extended without destroying monotonicity. The inverse of A is defined via its graph by $gphA^{-1} = \{(y, x) \in H \times H; y \in A(x)\}$ and the resolvent of A with parameter $\gamma > 0$ is $J_\gamma^A = (I + \gamma A)^{-1}$. This resolvent is not only always single-valued, but also firmly monotone (and thus Lipschitz continuous), namely

$$\langle J_\gamma^A(x) - J_\gamma^A(y), x - y \rangle \geq \|J_\gamma^A(x) - J_\gamma^A(y)\|^2, \quad \forall x, y \in H.$$

Moreover, the resolvent has full domain H precisely when A is maximal monotone.

Now, let A and B be two maximal monotone operators on H with A coercive. In this paper, we will be concerned by some convergence results of a split algorithm for solving the following null-point problem

$$\text{find } x^* \in H \text{ such that } 0 \in (A + B)(x^*). \quad (1.1)$$

An inclusion problem which subsumes a wide spectrum of problems in nonlinear analysis. Remember that an operator A is cocoercive if there exists $\tau > 0$ such that

$$\langle A(x) - A(y), x - y \rangle \geq \tau \|A(x) - A(y)\|^2, \quad \forall x, y \in H. \quad (1.2)$$

We recall the two following facts.

E-mail address: abdellatif.moudafi@univ-amu.fr.

Received May 25, 2018; Accepted June 4, 2018.

Fact 1. The resolvent of an operator B is firmly nonexpansive if and only if, for every $x, y \in H$ and $\gamma > 0$,

$$\|J_\gamma^B(x) - J_\gamma^B(y)\|^2 \leq \|x - y\|^2 - \|(I - J_\gamma^B)(x) - (I - J_\gamma^B)(y)\|^2. \quad (1.3)$$

Fact 2. Using the resolvent operator, we can rewrite problem (1.1) as a fixed point equation

$$x^* = J_\gamma^B(x^* - \gamma Ax^*), \quad (1.4)$$

for some $\gamma > 0$.

The forward-backward algorithm (or prox-gradient method) is to solve this fixed point equation via the iteration

$$x_{k+1} = J_{\gamma_k}^B(x_k - \gamma_k Ax_k), \quad (1.5)$$

where γ_k is a stepsize.

Fact 3. The resolvent of any maximal monotone operator B verifies the following so-called resolvent inequality

$$J_\alpha^B(x) = J_\alpha^B\left(\frac{\alpha}{\beta}x + \left(1 - \frac{\alpha}{\beta}\right)J_\beta^B(x)\right), \quad (1.6)$$

for every $x \in H$ and every $\alpha, \beta > 0$.

After revisiting some convergence properties of (1.5), we propose two convergence rate results based on bounded linear and bounded Hölder regular assumptions. An application to split feasibility null-point problems is stated and a remark on a possible extension to general null-point problems is also suggested.

2. THE MAIN CONVERGENCE RESULTS

Now, we are in a position to state and prove some convergence results. To begin with, let us establish the following key inequality that will be noted (\star) .

Proposition 2.1. *Let $\gamma_k \in]0, 2\tau[$, A, B be two maximal monotone operators with A a τ -cocoercive operator and assume that (1.1) possesses at least one solution. Let $(x_k)_{k \in \mathbb{N}}$ be the sequence defined by (1.5). Then, we have the following estimate*

$$\|x_{k+1} - x^*\|^2 \leq \|x_k - x^*\|^2 - 2\gamma_k \tau \|Ax_k - Ax^* - \frac{x_k - x_{k+1}}{2\tau}\|^2 - \frac{2\tau - \gamma_k}{2\tau} \|x_k - x_{k+1}\|^2. \quad (\star)$$

Proof. Let x^* be a solution of (1.1). In view of relations (1.3) and (1.5) and according to the firm nonexpansiveness property of the resolvent operators, we can successively write

$$\begin{aligned} \|x_{k+1} - x^*\|^2 &\leq \|(x_k - \gamma_k Ax_k) - (x^* - \gamma_k Ax^*), x_k - x^*\|^2 - \|(x_k - \gamma_k Ax_k - x_{k+1}) + \gamma_k Ax^*\|^2 \\ &\leq \|(x_k - x^*) - \gamma_k (Ax_k - Ax^*)\|^2 - \|(x_k - x_{k+1}) - \gamma_k (Ax_k - Ax^*)\|^2 \\ &\leq \|x_k - x^*\|^2 - 2\gamma_k \langle Ax_k - Ax^*, x_k - x^* \rangle - \|x_k - x_{k+1}\|^2 \\ &\quad + 2\gamma_k \langle Ax_k - Ax^*, x_k - x_{k+1} \rangle \\ &\leq \|x_k - x^*\|^2 - 2\gamma_k \tau \|Ax_k - Ax^*\|^2 - \|x_k - x_{k+1}\|^2 \\ &\quad + 2\gamma_k \langle Ax_k - Ax^*, x_k - x_{k+1} \rangle \\ &= \|x_k - x^*\|^2 - 2\gamma_k \tau \|Ax_k - Ax^* - \frac{x_k - x_{k+1}}{2\tau}\|^2 \\ &\quad - \frac{2\tau - \gamma_k}{2\tau} \|x_k - x_{k+1}\|^2. \end{aligned}$$

This completes the proof. \square

Throughout Ω will denote the solution set of (1.1).

Proposition 2.2. *If we suppose that $\gamma_k \in]\varepsilon, 2\tau - \varepsilon[$ for some given $\varepsilon > 0$ small enough, then we obtain the classical convergence results, namely, (x_k) is Fejer monotone with respect to Ω , bounded and asymptotically regular, and (Ax_k) norm converges to Ax^* . Moreover, (x_k) weakly converges to some solution of (1.1).*

Proof. From (\star) and the conditions on the sequence of parameters (γ_k) , we have

$$\|x_{k+1} - x^*\|^2 \leq \|x_k - x^*\|^2 - 2\varepsilon\tau \|Ax_k - Ax^* - \frac{x_k - x_{k+1}}{2\tau}\|^2 - \frac{\varepsilon}{2\tau} \|x_k - x_{k+1}\|^2.$$

This implies that the sequence $(\|x_k - x^*\|^2)$ is no increasing and thus it converges, which implies in turn the boundedness of the sequence (x_k) . We also directly obtain that $\sum_{k=0}^\infty \|x_k - x_{k+1}\|^2 < +\infty$ as well as that $\lim_{k \rightarrow +\infty} \|Ax_k - Ax^*\| = 0$. On the other hand, (1.5) can be rewritten as

$$\frac{x_k - x_{k+1}}{\gamma_k} - (Ax_k - Ax_{k+1}) \in (A + B)x_{k+1}. \tag{2.1}$$

Since A is Lipschitz continuous, we obtain by passing to the limit on a subsequence of (x_k) converging to a weak-cluster point \tilde{x} in (2.1) and by taking into account that $A + B$ is maximal monotone and thus its graph is weakly-strongly closed, that

$$0 \in (A + B)\tilde{x}.$$

Thus \tilde{x} belongs in Ω and the weak convergence of the whole sequence follows then by the celebrated Opial's Lemma. □

Now, let us now focus on the linear convergence of (1.5). To that end, we consider the bounded linear regular assumption, namely, for some $\sigma > 0$ and for every $x \in H$, we have

$$d(x, \Omega) \leq \sigma \|x - J_1^B(x - Ax)\|. \tag{2.2}$$

To begin with, let us prove the following key inequality: for all $x, z \in H$, $\alpha, \beta > 0$, we have

$$\|x - J_\alpha^A(x + \alpha z)\| \leq (1 + |1 - \frac{\alpha}{\beta}|) \|x - J_\beta^A(x + \beta z)\|. \tag{2.3}$$

Indeed, using the resolvent equation, we can write

$$\begin{aligned} \|J_\alpha^A(x + \alpha z) - J_\beta^A(x + \beta z)\| &= \|J_\alpha^A(x + \alpha z) - J_\alpha^A(\frac{\alpha}{\beta}(x + \beta z)) \\ &\quad + (1 - \frac{\alpha}{\beta})J_\beta^A(x + \beta z)\| \\ &\leq |1 - \frac{\alpha}{\beta}| \|x - J_\beta^A(x + \beta z)\|. \end{aligned}$$

On the other hand, we also have

$$\begin{aligned} \|J_\alpha^A(x + \alpha z) - J_\beta^A(x + \beta z)\| &= \|(x - J_\alpha^A(x + \alpha z)) - (x - J_\beta^A(x + \beta z))\| \\ &\geq \|x - J_\alpha^A(x + \alpha z)\| - \|x - J_\beta^A(x + \beta z)\|. \end{aligned}$$

Combining the two last inequalities, we obtain the desired inequality immediately.

Proposition 2.3. *Let A, B be two maximal monotone operators with A a τ -cocoercive operator. Let $\gamma_k \in]\varepsilon, 2\tau - \varepsilon[$ and assume that (2.2) is satisfied. Then, the sequence (x_k) generated by (1.5) converges linearly to some x^* which solves (1.1) and we have*

$$\|x_k - x^*\| \leq 2\chi^{\frac{k}{2}} \text{dist}(x_0, \Omega), \quad \forall k \in \mathbb{N}, \tag{2.4}$$

χ is defined in the proof.

Proof. From (\star) , we directly infer

$$\text{dist}^2(x_{k+1}, \Omega) \leq \text{dist}^2(x_k, \Omega) - \frac{2\tau - \gamma_k}{2\tau} \|x_k - x_{k+1}\|^2. \tag{2.5}$$

Using (2.3) with $x = x_k, z = -Ax_k, \alpha = 1$ and $\beta = \gamma_k$, we obtain relation $(\star\star)$ below

$$\begin{aligned} \|x_k - J_1^B(x_k - Ax_k)\| &\leq (1 + |1 - \frac{1}{\gamma_k}|) \|x_k - J_{\gamma_k}^B(x_k - \gamma_k Ax_k)\| \\ &= (1 + |1 - \frac{1}{\gamma_k}|) \|x_k - x_{k+1}\|. \end{aligned} \tag{\star\star}$$

The latter combined with (2.2) yields

$$\begin{aligned} \text{dist}^2(x_k, \Omega) &\leq \sigma^2 \|x_k - J_1^B(x_k - Ax_k)\|^2 \\ &= (\sigma(1 + |1 - \frac{1}{\gamma_k}|))^2 \|x_k - x_{k+1}\|^2 \\ &\leq (\sigma \max(2, \frac{1}{\gamma_k}))^2 \|x_k - x_{k+1}\|^2, \end{aligned}$$

which in the light of (2.5) and by taking into account conditions on γ_k yields

$$\text{dist}^2(x_{k+1}, \Omega) \leq \left(1 - \frac{\varepsilon}{2\tau(\sigma \max(2, \frac{1}{\varepsilon}))^2}\right) \text{dist}^2(x_k, \Omega).$$

Consequently, one has

$$\text{dist}(x_{k+1}, \Omega) \leq \chi^{\frac{1}{2}} \text{dist}(x_k, \Omega),$$

from which we infer, thanks to [[1], Theorem 5.12], that

$$\|x_k - x^*\| \leq 2\chi^{\frac{k}{2}} \text{dist}(x_0, \Omega),$$

with $\chi = \max\left(0, 1 - \frac{\varepsilon}{2\tau(\sigma \max(2, \frac{1}{\varepsilon}))^2}\right)$. □

Now, we will propose another convergence rate result based on the bounded Hölder regularity hypothesis. Namely, we assume that there exists $\eta \in (0, 1)$ such that, for every $x \in H$,

$$d(x, \Omega) \leq \sigma \|x - J_1^B(x - Ax)\|^\eta, \tag{2.6}$$

for some $\sigma > 0$.

We obtain the following result.

Proposition 2.4. *Let A, B be two maximal monotone operators with A a τ -cocoercive operator. Let $\gamma_k \in]\varepsilon, 2\tau - \varepsilon[$ and suppose that (2.6) is satisfied. Then, for all $k \in \mathbb{N}$, the following estimate holds*

$$\text{dist}^2(x_k, \Omega) \leq \text{dist}^2(x_0, \Omega) \left(1 + \eta \text{dist}^{2\eta}(x_0, \Omega) \left(\frac{\varepsilon}{2\tau\sigma^{\frac{2}{\eta}} \max(2, \frac{1}{\varepsilon})} k\right)\right)^{\frac{\eta}{\eta-1}}.$$

Proof. Similarly, by combining inequality $(\star\star)$ with the bounded Hölder regularity assumption (2.6), we obtain

$$\text{dist}^2(x_k, \Omega) \leq \sigma^2 \max^{2\eta} \left(2, \frac{1}{\gamma_k} \right) \|x_k - x_{k+1}\|^{2\eta}.$$

Tanking into account (2.5) and conditions on the parameters γ_k , we infer

$$\text{dist}^2(x_{k+1}, \Omega) \leq \text{dist}^2(x_k, \Omega) - \frac{\varepsilon}{2\tau\sigma^{\frac{2}{\eta}} \max^2 \left(2, \frac{1}{\varepsilon} \right)} \text{dist}^{\frac{2}{\eta}}(x_k, \Omega).$$

The claimed estimate follows by applying [[2], Lemma 6, page 46]. □

Remark 2.1. It is worth mentioning that our analysis can be applied to the problem of minimizing the sum of two proper convex lower semicontinuous functions f and g , g being differentiable with a L -Lipschitz gradient (which, in this case, is well-known to be co-coercive with constant $\frac{1}{L}$). Just take $B = \partial f$ and $A = \nabla g$, the subdifferential and the gradient of the functions f and g respectively. Likewise, by setting $B = N_C$ the normal cone of a closed convex set and $A = I - T$ with T , for instance, κ -strictly pseudocontractive or ξ -strongly pseudocontractive (the operator $I - T$ is $\frac{1-\kappa}{2}$ -cocoercive and ξ -cocoercive, respectively), we can apply our analysis to the variational inequality $(I - T)x^* + N_C(x^*) \ni 0$ which amount to finding x^* such that $x^* = P_C \circ T(x^*)$, P_C being the projection onto the set C .

3. FEASIBILITY NULL-POINT PROBLEMS

Feasibility null-point consists of finding $x^* \in H_1$ such that

$$0 \in B(x^*) \text{ and } 0 \in C(Tx^*), \tag{3.1}$$

where B , and C are two maximal monotone operators defined on two real Hilbert spaces H_1, H_2 , respectively, and $T : H_1 \rightarrow H_2$ a bounded linear operator. Problem (3.1) is equivalent to finding x^* such that

$$0 \in T^*(I - J_\lambda^C)Tx^* + Bx^*,$$

$\lambda > 0$ and T^* being the adjoint operator of T , see, for example, [[3], Lemma 3.3]. Moreover $A := T^*(I - J_\lambda^C)Tx^*$ is $\frac{1}{\|T\|^2}$ -cocoercive, see, for example, [4]. As a direct application, for instance, of Proposition 2.3 gives

$$\|x_k - x^*\| \leq 2\chi^{\frac{k}{2}} \text{dist}(x_0, \Omega), \quad \forall k \in \mathbb{N}, \tag{3.2}$$

with $\chi = \max \left(0, 1 - \frac{\varepsilon\|T\|^2}{2 \left(\sigma_{\max} \left(2, \frac{1}{\varepsilon} \right) \right)^2} \right)$.

Remark 3.1. Now, we can solve a general split null-point problem, namely,

$$0 \in (B + C)x^*, \tag{3.3}$$

with B, C two general maximal monotone operators. The key is the use of the smoothing of one of the two original operators by its Yosida approximate which is always cocoercive. Namely, we may consider the following approximate problem

$$0 \in (B + C_\lambda)x_\lambda^*, \tag{3.4}$$

where C_λ is the Yosida approximate of C . This is a relevant approximation, since in addition to the fact that C_λ is λ -cocoercive, we have that $B + C_\lambda$ graph converges to $B + C$ when $B + C$ is a maximal monotone operator and it was established, for example, that if x^* is a limit point of the family $\{x_\lambda^*, \lambda \rightarrow 0\}$ and if

we assume that $\{C_\lambda(x_\lambda^*), \lambda \rightarrow 0\}$ is bounded, then x^* solve (3.3), see [5]. If in addition C is strongly monotone and x^* the solution of (3.3), by [5]-Theorem 3, we also have the following estimate

$$\|x_\lambda^* - x^*\| \leq o(\sqrt{\lambda}).$$

Acknowledgement

This work was supported by I&M team of Systems and Computer Sciences Laboratory (LIS) at Aix Marseille University.

REFERENCES

- [1] H. H. Bauschke, P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, Springer, New York (2011).
- [2] B. T. Polyak, *Introduction to Optimization*, Optimization Software, New York, 1987.
- [3] S. M. Alsulami, W. Takahashi, The split common null point problem for maximal monotone mappings in Hilbert spaces and applications, *J. Nonlinear Convex Anal.* 15 (2014), 793-808.
- [4] H.-K. Xu, Iterative methods for the split feasibility problem in infinite-dimensional Hilbert spaces, *Inverse Probl.* 26 (2010), Article ID 105018.
- [5] P. Mahey, P.D. Tao, Partial regularization of the sum of two maximal monotone operators, *RAIRO Modél. Math. Anal. Numér* 27 (1993), 375-392.