On a New Simple Algorithm to Compute the Resolvents

B. K. Le., Michel Théra

Noname manuscript No.
(will be inserted by the editor)

On a New Simple Algorithm to Compute the Resolvents

B. K. Le **M. Théra**

Received: date / Accepted: date

Abstract Resolvents of operators are the core of many fundamental algorithms used in optimization. However their computation is in general difficult except for very particular operators. In the paper we provide a new simple algorithm with linear convergence rate to compute the resolvents for the class of operators which can be decomposed as a sum of a maximally monotone operator with a computable resolvent and a single-valued locally Lipschitz continuous mapping.

Keywords Maximally monotone operators, convex optimization, computation of resolvents

1 Introduction

Resolvents are an essential tool in the problem of finding the zeroes of a monotone inclusion

$$0 \in Ax$$

where A is a multivalued operator from a Hilbert space H to subsets of itself. As an example, let consider the case when A is the subdifferential ∂f of a lower semicontinuous convex function f. The problem of finding a zero of ∂f is equivalent to finding a minimizer of the function f, and we know since Moreau and Rockafellar that the resolvent of ∂f that is exactly the proximal function of f that plays a central role in the proximal point algorithm [6, 15, 16, 22].

Other well known approaches where resolvents are an important tool for finding a solution to monotone inclusions are the so-called splitting methods. They consist in splitting A additively ($A = B + C$) as a sum of two other monotone operators. Then the idea is to use the resolvents of B and C, to obtain a numerical method. This was the object of the Douglas-Rachford algorithm introduced by Lions and Mercier [13] and of its later extensions (see, e. g., [11, 12, 14]).

B. K. Le
Faculty of Mathematics and Statistics
Ton Duc Thang University, Ho Chi Minh City, Vietnam
E-mail: lebakhiet@tdtu.edu.vn

Michel Théra
ORCID 0000-0001-9022-6406
Laboratoire XLIM UMR-CNRS 7252, Université de Limoges, 87032 Limoges, France
Federation University Australia, Ballarat 3353, Australia
E-mail: michel.thera@unilim.fr, m.thera@federation.edu.au

Splitting algorithms are useful provided that the operators in the sum decomposition have resolvents that are easily computed. The main problem is that in general, the computation is difficult except for very particular operators.

In this paper, we consider the class of maximally monotone operators that have a decomposition of the form $A = B + C$, where $C : H \to H$ is Lipschitz continuous on bounded sets and $B : H \rightrightarrows H$ is maximally monotone with a computable resolvent. Most of the previous works in the literature provide algorithms to compute the resolvent J_A, provided the resolvents J_B and J_C are available (see, e. g., [16, 18]). Note that if one relies on the Douglas–Rachford algorithm for our class (see, e.g., [11, 13]), it is easy to see that the computation of J_C makes this approach less effective. The Douglas–Rachford algorithm is obviously a favorite choice if C is also set-valued.

On the other hand, the forward-backward algorithm, firstly introduced by Passty [21], is another important splitting method and requires only the computation of J_B [2–5, 10]. Using the forward-backward technique, we provide a direct simple but efficient algorithm to compute the resolvent of A. We show that our algorithm is preferable to the classical forward-backward approach. In [3], the authors also have used an approach based on the forward-backward algorithm to compute the resolvents. Here we provide explicitly the linear rate of our algorithm which helps us to decide a good choice of parameters easily.

The paper is structured as follows. In Section 2, we recall some useful definitions concerning maximally monotone operators. A new linear convergence algorithm to compute the resolvents is provided in Section 3. Finally, we end the paper with conclusions in Section 4.

2 Notations and preliminaries

Throughout this paper, we consider that H is a real Hilbert space with inner product $\langle \cdot, \cdot \rangle : H \times H \to \mathbb{R}$ and induced norm $\| \cdot | = \sqrt{\langle \cdot, \cdot \rangle}$. The closed unit ball is denoted by $\mathbb{B}$. Let K be a nonempty closedset of H. The distance from a point x to K is

$$d(x, K) := \inf_{u \in K} \|u - x\|$$

and the projector proj_K onto K is the mapping defined by

$$\mathrm{proj}_K(x) := \operatorname*{argmin}_{u \in K} \|x - u\|.$$

When K is convex and closed, then proj_K is single-valued and the least norm element of K is defined by $\mathrm{proj}_K(0)$.

The notation $A : H \rightrightarrows H$ is used to denote a set-valued mapping (operator), that is a mapping which assigns to every $x \in H$ a subset Ax (eventually empty) of H. The domain, the range, the graph and the inverse of A are defined respectively by

$$\mathrm{dom}\, A = \{x \in H : \ Ax \neq \emptyset\}, \ \ \mathrm{rge} A = \bigcup_{x \in H} Ax,$$

$$\mathrm{gph}\, A = \{(x, y) : x \in H, y \in Ax\}$$

and

$$A^{-1}(y) = \{x \in H : y \in Ax\}.$$

The operator $A : H \rightrightarrows H$ is said to be *monotone* if

$$\langle x^* - y^*, x - y \rangle \geq 0 \ \ \forall\, x, y \in H, x^* \in Ax \text{ and } y^* \in Ay.$$

In addition, it is called *maximally monotone* provided its graph $\{(x,y) \ : y \in Ax\}$ cannot be properly enlarged without destroying monotonicity. The set of maximally monotone operators on H includes subdifferential operators of proper lower semicontinuous convex functions as well as all square matrices with symmetric parts that are positive semidefinite (see e.g., [7, 22]). Let us remind that the *normal cone operator* N_C to a closed convex set C is defined by

$$N_C(y) := \{y \in C : \langle y, x-u \rangle \leq 0, \text{for all } y \in C\}.$$

It is the subdifferential of the *indicator function* of C which is equal to 0 in C and $+\infty$ outside C. Hence, the normal cone operator is also another important example of maximally monotone operators.

A is called μ*-strongly monotone* ($\mu > 0$) provided

$$\langle Ax - Ay, x-y \rangle \geq \mu \|x-y\|^2 \ \ \forall\, x,y \in H.$$

If $\mu \in \mathbb{R}$ then A is called μ*-monotone*.

The *resolvent* of index $\gamma > 0$ of A is defined as follows

$$J_{\gamma A} := (Id + \gamma A)^{-1}$$

where Id denotes the *identity* operator. A classical result which goes back to Minty [17] and Rockafellar [22] says that resolvents of maximally monotone operators are single-valued mappings with full domain and are *firmly nonexpansive*:

$$\|J_A x - J_A y\|^2 \leq \langle x-y, J_A x - J_A y \rangle.$$

In particular, they are *nonexpansive*, i.e., 1- Lipschitz continuous.

Let $B : H \rightrightarrows H$ be a maximally monotone operator and $C : H \to H$ be a monotone continuous mapping. Denoting by x_n the nth iterate and given a fixed step size $\alpha > 0$, we may consider the forward-backward algorithm:

$$x_0 \in H, x_{k+1} = J_{\alpha B} x_k - \alpha C x_k, \ k = 0, 1, 2 \ldots$$

used to find a zero x^* of the sum $B + C$, i.e., $0 \in B(x^*) + C(x^*)$. One has the following result (see, e.g., [10]).

Theorem 2.1 *If B is μ-strongly monotone and C is monotone, L-Lipschitz continuous then the sequence (x_k) generated by the forward-backward algorithm with $\alpha = \mu / L^2$ converges to the unique zero x^* of $B + C$ with the linear rate $r = L/\sqrt{L^2 + \mu^2}$.*

3 Main results

In this section, we propose a simple algorithm to compute resolvents of a monotone operator of the type $A = B + C$, where $C : H \to H$ is L-Lipschitz continuous and $B : H \rightrightarrows H$ is a maximally monotone operator with a computable resolvent J_B.

Precisely, let be given $y \in H$ and $\gamma > 0$, we want to find $J_{\gamma A}(y)$, i.e. the unique $x^* \in H$ such that

$$y \in (\gamma A + I)(x^*) = (\gamma B + \gamma C + I)(x^*).$$

We propose the following algorithm.

Algorithm 1 : $x_0 \in H, 0 < \alpha \leq 1, \ x_{k+1} = J_{\alpha \gamma B}(x_k - \alpha(x_k - y + \gamma C(x_k)), \ k \geq 0.$

Theorem 3.1 *If C is monotone then the sequence (x_k) generated by Algorithm 1 with $\alpha = \alpha^* := \frac{1}{1+\gamma^2L^2}$ converges to x^* with the linear rate $r_1 := \frac{\gamma L}{\sqrt{\gamma^2L^2+1}}$. If C is not monotone and $\gamma L < 1$, the best linear rate is $r_2 = \gamma L$ when $\alpha = 1$.*

Proof. Since $J_{\alpha\gamma B}$ is non-expansive, we have

$$\|x_{k+1} - x^*\| \le \|(1-\alpha)(x_k - x^*) - \alpha\gamma(C(x_k) - C(x^*))\|. \tag{3.1}$$

If C is monotone, one obtains

$$\begin{aligned}\|x_{k+1} - x^*\|^2 &\le [(1-\alpha)^2 + \alpha^2\gamma^2L^2]\|x_k - x^*\|^2 \\ &= [1 - 2\alpha + \alpha^2(1+\gamma^2L^2)]\|x_k - x^*\|^2. \end{aligned} \tag{3.2}$$

Consequently

$$\|x_{k+1} - x^*\| \le \frac{\gamma L}{\sqrt{\gamma^2L^2+1}}\|x_k - x^*\|$$

when $\alpha = \alpha^*$ and the conclusion follows.

If C is not monotone and $\gamma L < 1$, one has

$$\|x_{k+1} - x^*\| \le [(1-\alpha) + \alpha\gamma L]\|x_k - x^*\| = \gamma L\|x_k - x^*\|$$

when $\alpha = 1$. □

Remark 3.1 i) If $\alpha = 1$ then $x_{k+1} = J_{\gamma B}(y - \gamma C(x_k))$, which is very simple.

ii) Note that Algorithm 1 may be obtained by applying the forward-backward technique to $\mathscr{B} = \gamma B, \mathscr{C} = \gamma C + I - y$. However in this case $\mathscr{B}$ is not strongly monotone and $\mathscr{C}$ is Lipschitz continuous with constant $\gamma L + 1$. Thus one cannot obtain the linear rate convergence by using the classical analysis.

iii) Traditionally, one applies the forward-backward algorithm to $\mathscr{B} = I + \gamma B$ and $\mathscr{C} = y - \gamma C$. If C is monotone then the optimal linear rate is also $r_1 = \frac{\gamma L}{\sqrt{\gamma^2L^2+1}}$ with $\alpha = \tilde{\alpha} := \frac{1}{\gamma^2L^2}$. Note that if $\gamma L \to 0$, the behavior of α^* is very good: it tends to 1 while $\tilde{\alpha} \to \infty$. Furthermore, in Algorithm 1 we use directly the resolvent of B instead of the resolvent of $I + \gamma B$. It means that our approach is preferable.

Our analysis can be applied similarly if the monotonicity of C is replaced by the μ-monotonicity. Note that if C is L-Lipschitz continuous then C is also $(-L)$-monotone. Thus the only interesting case is when C is μ-monotone where $\mu \ge -L$.

Theorem 3.2 *Suppose that C is μ-monotone where $\mu \ge \max\{-L, -\gamma L^2\}$ and $\gamma\mu > -1$. Then the sequence (x_k) generated by Algorithm 1 with $\alpha = \alpha_m := \frac{1+\gamma\mu}{1+2\gamma\mu+\gamma^2L^2}$ converges to x^* with the linear rate $r_3 := \frac{\gamma\sqrt{L^2-\mu^2}}{\sqrt{1+2\gamma\mu+\gamma^2L^2}} < 1$.*

Proof. From (3.1) and the μ-monotonicity of C, we have

$$\begin{aligned}\|x_{k+1} - x^*\|^2 &\le [(1-\alpha)^2 - 2\alpha(1-\alpha)\gamma\mu + \alpha^2\gamma^2L^2]\|x_k - x^*\|^2 \\ &= [1 - 2\alpha(1+\gamma\mu) + \alpha^2(1+2\gamma\mu+\gamma^2L^2)]\|x_k - x^*\|^2.\end{aligned}$$

Thus

$$\|x_{k+1} - x^*\| \le \sqrt{1 - \frac{(1+\gamma\mu)^2}{1+2\gamma\mu+\gamma^2L^2}}\|x_k - x^*\|$$

when $\alpha = \alpha_m$ and the conclusion follows. □

Remark 3.2 The conditions $\mu \geq -\gamma L^2$ and $\gamma\mu > -1$ ensure that $0 < \alpha_m \leq 1$.

The following result is a consequence of Theorem 3.1 when C is only Lipschitz continuous on bounded sets.

Corollary 3.1 *Suppose that C is monotone and Lipschitz continuous on bounded sets. Let L be the Lipschitz constant of C on $\mathscr{K} := \mathbb{B}(x^*, \|x^* - x_0\|)$ and $\alpha = \alpha^* := \frac{1}{1+\gamma^2 L^2}$. Then the sequence (x_k) generated by Algorithm 1 converges to x^* with the linear rate $r_1 = \frac{\gamma L}{\sqrt{\gamma^2 L^2 + 1}}$.*

Proof. We can use arguments in the proof of Theorem 3.1. Note that $\|x_1 - x^*\| < \|x_0 - x^*\|$, which implies that $x_1 \in \mathscr{K}$. By using the induction we have $x_k \in \mathscr{K}$ for all $k \geq 0$ and the conclusion follows. $\square$

Remark 3.3 If $H = \mathbb{R}^n$ then the Lipschitz continuity on bounded sets of C becomes the local Lipschitz continuity.

4 Conclusions

Linear convergence of algorithms computing resolvents makes the resolvents-based approach more competitive in optimization (see also, e. g., [8, 9, 19, 20]). It would be interesting to compute the resolvent of a given continuous monotone function or a set-valued maximal monotone function where the Lipschitz continuity of the single-valued part is absent.

Acknowledgements

Research of the second author benefited from the support of the FMJH Program PGMO and from the support of EDF.

References

1. S. Adly, B. K. Le, *Douglas-Rachford splitting algorithm for solving state-dependent maximal monotone inclusions*, Optim Lett 15, 2861–2878 (2021)
2. V. Apidopoulos, J. Aujol, C. Dossal, *Convergence rate of inertial Forward-Backward algorithm beyond Nesterov's rule*, Math Program 180, 137–156, (2020)
3. F. J. A. Artacho, R. Campoy, M. K. Tam, *Strengthened Splitting Methods for Computing Resolvents*, arXiv:2011.01796
4. H. Attouch, A. Cabot, *Convergence Rates of Inertial Forward-Backward Algorithms*, SIAM J Optim, 28(1), 849–874 (2018)
5. H. Attouch, A. Cabot, *Convergence of a Relaxed Inertial Forward–Backward Algorithm for Structured Monotone Inclusions*. Appl Math Optim 80, 547–598 (2019)
6. H. Attouch, Z. Chbani, H. Riahi, *Fast Proximal Methods via Time Scaling of Damped Inertial Dynamics*, SIAM J. Optim., 29(3), 2227–2256 (2019)
7. H. H. Bauschke, P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. Springer, Berlin (2011)
8. A. Beck, M. Teboulle, *A fast iterative shrinkage-thresholding algorithm for linear inverse problems*, SIAM J. Imaging Sci. 2(1), 183–202 (2009)
9. A. Chambolle, Ch. Dossal, *On the convergence of the iterates of the "fast iterative shrinkage/ thresholding algorithm"* J. Optim. Theory Appl. 166, 968–982 (2015)
10. G. H. G. Chen, R. T. Rockafellar, *Convergence Rates in Forward-Backward Splitting*, SIAM J Optim, 7(2), 421–444 (1997)
11. P. Giselsson, *Tight global linear convergence rate bounds for Douglas-Rachford splitting*, J Fixed Point Theory Appl 19, 2241–2270 (2017)

12. B. He, X. Yuan, *On the convergence rate of Douglas–Rachford operator splitting method*. Math. Program. 153, 715–722 (2015)
13. P. L. Lions, B. Mercier, *Splitting algorithms for the sum of two nonlinear operators,* SIAM J Numer Anal 16, 964–979 (1979)
14. Y. Ma, L. Z. Liao, *The Glowinski–Le Tallec splitting method revisited: A general convergence and convergence rate analysis*, Journal of Industrial & Management Optimization 17(4): 1681-1711 (2021)
15. B. Martinet, *Régularisation d'inéquations variationnelles par approximations successives*. Revue Française D'automatique, Informatique, Recherche Opérationnelle, 3, 154–158 (1970)
16. C. A. Micchelli, L. Shen, Y. Xu, X. Zeng, *Proximity algorithms for the L1/TV image denoising model*. Adv Comput Math 38, 401–426 (2013)
17. G. J. Minty, *Monotone (nonlinear) operators in Hilbert space.* Duke Math. J. 29(3), 341–346 (1962)
18. A. Moudafi, *Computing the resolvent of composite operators*, CUBO, vol. 16, no. 3, pp. 87–96 (2014)
19. Y. Nesterov, *A method of solving a convex programming problem with convergence rate O(1/k2)*. Soviet Mathematics Doklady, 27(2):372–376 (1983)
20. Y. Nesterov, *Introductory lectures on convex optimization, volume 87 of Applied Optimization.* Kluwer Academic Publishers, Boston, MA, 2004. A basic course.
21. G. B. Passty, *Ergodic convergence to a zero of the sum of monotone operators in Hilbert space*, J Math Anal Appl 72(2), 383–390 (1979)
22. R. T. Rockafellar, *Monotone operators and the proximal point algorithm*. SIAM Journal on Control and Optimization, 14(5):877–898 (1976)