Journal of the Egyptian Mathematical Society (2013) **21**, 281–284

Egyptian Mathematical Society

Journal of the Egyptian Mathematical Society

www.etms-eg.org
www.elsevier.com/locate/joems

REVIEW PAPER

About the relaxed cocoercivity and the convergence of the proximal point algorithm

Abdellatif Moudafi [a,*], Zhenyu Huang [b]

[a] *French West Indies University, Ceregmia-Scientific Department, 97200 Schelcher, Martinique, France*
[b] *Department of Mathematics, Nanjing University, Nanjing 210093, PR China*

Received 6 February 2013; revised 14 March 2013; accepted 24 March 2013
Available online 8 May 2013

KEYWORDS

Relaxed cocoercivity;
Relaxed monotonicity;
Proximal point algorithm;
Strong convergence;
Variational inequalities

Abstract The aim of this paper is to study the convergence of two proximal algorithms via the notion of (α, r)-relaxed cocoercivity without Lipschitzian continuity. We will show that this notion is enough to obtain some interesting convergence theorems without any Lipschitz-continuity assumption. The relaxed cocoercivity case is also investigated.

MATHEMATICS SUBJECT CLASSIFICATION:
Primary 49J53 65K10 Secondary 49M37 90C25

Contents

1. Introduction and preliminaries

Throughout, H is a real Hilbert space, $\langle\cdot,\cdot\rangle$ denotes the associated scalar product and $\|\cdot\|$ stands for the corresponding norm.

To begin with, let us recall that an operator A is (α, r)-relaxed cocoercive if there exist constants $\alpha \geqslant 0$, $r > 0$, such that

$$\langle A(x) - A(y), x - y\rangle \geqslant -\alpha\|A(x) - A(y)\|^2 + r\|x - y\|^2 \quad \text{forall } x, \quad y \in H. \tag{1.1}$$

Recently, this notion was used to establish the convergence of algorithms for variational inequalities and systems of variational inequalities, see for instance [1–5] and the references therein. It is worth mentioning that this notion combined with the γ-Lipschitz continuity with positive constant γ, namely for all $x, y \in H$

$$\|A(x) - A(y)\| \leqslant \gamma\|x - y\|, \tag{1.2}$$

* Corresponding author. Tel.: +596 0596727363.
E-mail addresses: abdellatif.moudafi@martinique.univ-ag.fr (A. Moudafi), zhenyu@nju.edu.cn, profzhenyu@gmail.com (Z. Huang).
Peer review under responsibility of Egyptian Mathematical Society.

ELSEVIER Production and hosting by Elsevier

1110-256X © 2013 Production and hosting by Elsevier B.V. on behalf of Egyptian Mathematical Society. Open access under CC BY-NC-ND license.
http://dx.doi.org/10.1016/j.joems.2013.03.014

implies that the operator A is $(r - \alpha\gamma^2)$-strongly monotone, namely

$$\langle A(x) - A(y), x - y\rangle \geqslant (r - \alpha\gamma^2)\|x - y\|^2$$

provided that $(r - \alpha\gamma^2) > 0$. Consequently the convergence of gradient-projection type methods related to variational inequalities and systems of variational inequalities follows by virtue of the classical results.

Condition (1.2) combined with (1.1) is thus too strong since the resulting convergence results are very close to the classical ones.

The following question arises naturally:

Question: Could we obtain convergence results without the Lipschtzian continuity assumption (1.2)?

The purpose of this paper is to partially answer the question mentioned above by proving strong and weak convergence results for the celebrate implicit methods that are the proximal point algorithm and its relaxed version by Yosida approximation. It is well-known that the sequence generated by the proximal point algorithm converges in norm to the unique zero of A when A is strongly monotone. However, if A is monotone with a zero and the parameters are bounded away from zero, we only have weak convergence. Instead, hybrid proximal algorithms prevail, see for example [6]. We would like also to emphasize that very recently new convergence results of proximal point algorithms were obtained under generalized monotonicity notions, see for instance [7–9]. Proximal point algorithms were also successfully applied in various areas such as image restoration and signal recovery, see for instance [10]. It is worth mentioning that if $r = 0$ the operator is called relaxed cocoercive or cohypomonotone. Further, a relaxed cocoercive operator A is said to be maximal, if in addition its graph, *gph* $A := \{(x,y) \in H \times H : y \in A(x)\}$, is not properly contained in the graph of any other relaxed-cocoercive operator or in other words A^{-1} is maximal hypomonotone, see for example [11]. For $\lambda > 0$ the operator $J_\lambda^A := (I + \lambda A)^{-1}$ is called the resolvent operator of A of parameter λ and is related to its Yosida approximate, namely $A_\lambda(x) := \frac{x - J_\lambda^A(x)}{\lambda}$, by the relation $A_\lambda(x) \in A(J_\lambda^A(x))$. Finally, recall that the inverse A^{-1} of A is the operator defined by $x \in A^{-1}(y) \Longleftrightarrow y \in A(x)$.

2. The main convergence results

Variational inclusions of the form

$$\text{finding } \bar{x} \in H \text{ such that } 0 \in A(\bar{x}), \tag{2.3}$$

where $A : H \to 2^H$ is a set-valued operator on a Hilbert space H, providing a convenient form for many problems arising in practice. For instance, minimization problems can be written in this form by setting $A = \partial f$, where ∂f is the subdifferential of the objective function f. Other problems such as saddle point problems, variational inequalities and complementarity problem can be written in this form, see for instance [12]. Throughout this paper, we will consider the following notion of (α, r)-cocoercivity: a set-valued operator A will be said to be (α, r)-cocoercive, if there exist $\alpha \geqslant 0$, $r > 0$ such that for all $x, y \in H$,

$$\langle \eta - \xi, x - y\rangle \geqslant -\alpha\|\eta - \xi\|^2 + r\|x - y\|^2 \text{ for all } \eta \in A(x), \xi \in A(y). \tag{2.4}$$

It should be noticed that (2.4) coincides with (1.1) when the operator is single-valued.

Now, let us prove the following key property of the resolvent operator.

Proposition 2.1. *Assume that the operator A is (α,r)-relaxed cocoercive, then for all $x, y \in H$ one has*

$$\left\|J_\lambda^A(x) - J_\lambda^A(y)\right\|^2 \leqslant L(\lambda)\|x - y\|^2 - \kappa(\lambda)\left\|\left(I - J_\lambda^A\right)(x) - \left(I - J_\lambda^A\right)(y)\right\|^2, \tag{2.5}$$

where $L(\lambda) := \dfrac{1}{1 + 2\lambda r}$ *and* $\kappa(\lambda) := \dfrac{\lambda - 2\alpha}{\lambda(1 + 2\lambda r)}$.

Proof. Since $A_\lambda(x) \in A\left(J_\lambda^A(x)\right)$ and $A_\lambda(y) \in A\left(J_\lambda^A(y)\right)$, definition (2.4) yields

$$\langle A_\lambda(x) - A_\lambda(y), J_\lambda^A(x) - J_\lambda^A(y)\rangle \geqslant -\alpha\|A_\lambda(x) - A_\lambda(y)\|^2 + r\left\|J_\lambda^A(x) - J_\lambda^A(y)\right\|^2.$$

Using the fact that $A_\lambda(x) = \frac{x - J_\lambda^A(x)}{\lambda}$ and $A_\lambda(y) = \frac{y - J_\lambda^A(y)}{\lambda}$, we can write

$$\left\langle \frac{x - J_\lambda^A(x)}{\lambda} - \frac{y - J_\lambda^A(y)}{\lambda}, J_\lambda^A(x) - J_\lambda^A(y)\right\rangle \geqslant -\alpha\left\|\frac{x - J_\lambda^A(x)}{\lambda} - \frac{y - J_\lambda^A(y)}{\lambda}\right\|^2 + r\left\|J_\lambda^A(x) - J_\lambda^A(y)\right\|^2,$$

or equivalently

$$\left\langle x - y - \left(J_\lambda^A(x) - J_\lambda^A(y)\right), J_\lambda^A(x) - J_\lambda^A(y)\right\rangle \geqslant -\frac{\alpha}{\lambda}\left\|x - J_\lambda^A(x) - \left(y - J_\lambda^A(y)\right)\right\|^2 + \lambda r\left\|J_\lambda^A(x) - J_\lambda^A(y)\right\|^2.$$

Thus

$$\left\langle J_\lambda^A(x) - J_\lambda^A(y), x - y\right\rangle \geqslant -\frac{\alpha}{\lambda}\left\|\left(I - J_\lambda^A\right)(x) - \left(I - J_\lambda^A\right)(y)\right\|^2 + (1 + \lambda r)\left\|J_\lambda^A(x) - J_\lambda^A(y)\right\|^2. \tag{2.6}$$

On the other hand, we also have

$$2\left\langle J_\lambda^A(x) - J_\lambda^A(y), x - y\right\rangle = \|x - y\|^2 + \left\|J_\lambda^A(x) - J_\lambda^A(y)\right\|^2 - \left\|\left(I - J_\lambda^A\right)(x) - \left(I - J_\lambda^A\right)(y)\right\|^2.$$

We infer, by taking into account (2.6), that

$$\begin{aligned}&\|x - y\|^2 + \left\|J_\lambda^A(x) - J_\lambda^A(y)\right\|^2 - \left\|\left(I - J_\lambda^A\right)(x) - \left(I - J_\lambda^A\right)(y)\right\|^2\\ &= 2\left\langle J_\lambda^A(x) - J_\lambda^A(y), x - y\right\rangle\\ &\geqslant -\frac{2\alpha}{\lambda}\left\|\left(I - J_\lambda^A\right)(x) - \left(I - J_\lambda^A\right)(y)\right\|^2\\ &\quad + 2(1 + \lambda r)\left\|J_\lambda^A(x) - J_\lambda^A(y)\right\|^2.\end{aligned}$$

Thus

$$\|x - y\|^2 - \left(1 - \frac{2\alpha}{\lambda}\right)\left\|\left(I - J_\lambda^A\right)(x) - \left(I - J_\lambda^A\right)(y)\right\|^2 \geqslant (1 + 2\lambda r)\left\|J_\lambda^A(x) - J_\lambda^A(y)\right\|^2.$$

Dividing by $(1 + 2\lambda r)$, we finally obtain the desired result, namely

$$\left\|J_\lambda^A(x) - J_\lambda^A(y)\right\|^2 \leqslant \frac{1}{1+2\lambda r}\|x-y\|^2 - \frac{\lambda - 2\alpha}{\lambda(1+2\lambda r)} \times \left\|\left(I - J_\lambda^A\right)(x) - \left(I - J_\lambda^A\right)(y)\right\|^2. \quad \square$$

Now, we are in position to prove the following convergence theorems for the proximal point algorithm and its relaxed version.

Theorem 2.1. *Assume that the operator A is (α,r)-relaxed cocoercive. Then the sequence generated by the proximal point algorithm*

$$x_{n+1} = J_{\lambda_n}^A(x_n), \tag{2.7}$$

strongly converges to x^**, the unique solution of the problem of finding zeroes of A, provided that the regularized parameters* $\{\lambda_n\}$ *satisfy* $\lambda_n \geqslant 2\alpha$ *for all* $n \in I\ N$.

Proof. Since finding zeroes of A is equivalent to the problem of finding fixed points of the associated resolvent operator, it is easy to see using (2.5) and $\lambda \geqslant 2\alpha$ that the resolvent operator is a $\frac{1}{1+4\alpha r}$-contraction and thus problem (2.3) has a unique solution x^*. On the other hand using (2.7) and the fact that $\lambda_n \geqslant 2\alpha$, we can write

$$\|x_{n+1} - x^*\|^2 = \left\|J_{\lambda_n}^A(x_n) - J_{\lambda_n}^A(x^*)\right\|^2 \leqslant \frac{1}{1+4\alpha r}\|x_n - x^*\|^2,$$

from which we infer that $\{x_n\}$ generated by (2.7) strongly converges to x^*. $\square$

Now, the limit case $r = 0$, namely: there exists $\alpha \geqslant 0$ such that for all $x, y \in H$ one has

$$\langle \eta - \xi, x - y\rangle \geqslant -\alpha\|\eta - \xi\|^2 \text{ for all } \eta \in A(x), \xi \in A(y),$$

is very interesting and corresponds to relaxed cocoercivity. This case has been studied, for instance, [7,8] and the references therein, using a local maximal monotonicity of the Yosida approximate. In what follows, our approach is different and is based on the property (2.5) of the resolvent operator which reduced, in this context, to

$$\left\|J_\lambda^A(x) - J_\lambda^A(y)\right\|^2 \leqslant \|x-y\|^2 - \left(1 - \frac{2\alpha}{\lambda}\right) \times \left\|\left(I - J_\lambda^A\right)(x) - \left(I - J_\lambda^A\right)(y)\right\|^2. \tag{2.8}$$

The following theorem summarizes our convergence results.

Theorem 2.2. *Assume that the operator A is α-relaxed cocoercive with a zero, then the following assertions hold true:*

(i) The sequence generated by the proximal point algorithm (2.7), namely

$$x_{n+1} = J_{\lambda_n}^A(x_n)$$

weakly converges to, x^**, a solution of the problem of finding zeroes of A, provided that A is maximal relaxed cocoercive and the regularized parameters* $\{\lambda_n\}$ *satisfy* $\lambda_n \geqslant \frac{2\alpha}{1-\delta}$ *for all* $n \in IN$*, where* $\delta \in (0,1)$ *is a small enough constant.*

(ii) Assume that the regularized parameter satisfies now the condition $\lambda \in (\alpha, 2\alpha)$ and the control sequence $\{\alpha_n\}$ is chosen so that $\alpha_n \in (\kappa, 1)$ for all $n \in I\ N$ and $\sum_{n=0}^{\infty}(\alpha_n - \kappa)(1-\alpha_n) = \infty$ with $\kappa := \frac{2\alpha}{\lambda} - 1$. Then the sequence generated by the relaxed rule

$$x_{n+1} = \alpha_n x_n + (1-\alpha_n) J_\lambda^A(x_n) \text{ with } \alpha_n \in (0,1),$$

weakly converges to a zero of A.

Proof.

(i) Relation (2.8) combined with the condition on the parameters yields

$$\|x_{n+1} - x^*\|^2 \leqslant \|x_n - x^*\|^2 - \delta\|x_{n+1} - x_n\|^2.$$

From which we infer that the sequence $\{\|x_n - x^*\|^2\}$ converges in $I\ R$, $\{x_n\}$ is bounded and asymptotically regular, that is $\{\|x_{n+1} - x_n\|\}$ norm converges to 0. Now, let $\bar{x}$ be a weak cluster point of the sequence $\{x_n\}$, there exists a subsequence of $\{x_n\}$ (which we still noted $\{x_n\}$) that weakly converges to $\bar{x}$. By passing to the limit in the following equivalent formulation of (2.7)

$$x_{n+1} \in A^{-1}\left(\frac{x_n - x_{n+1}}{\lambda_n}\right),$$

and by taking into account the fact that the graph of A^{-1} is weakly-strongly closed (because A^{-1} is maximal hypomotone operator, see for instance [11, Proposition 4]), we obtain at the limit that $\bar{x}$ is a zero of A. The uniqueness of the weak-cluster point and thus the weak convergence of the whole sequence (x_n) follows by the same arguments as the classical result by Rockafellar [12].

(ii) In this context, relation (2.8) combined with the condition on the parameters yields

$$\left\|J_\lambda^A(x) - J_\lambda^A(y)\right\|^2 \leqslant \|x-y\|^2 + \kappa\left\|\left(I - J_\lambda^A\right)(x) - \left(I - J_\lambda^A\right)(y)\right\|^2, \tag{2.9}$$

with $\kappa = \frac{2\alpha}{\lambda} - 1 \in (0,1)$. Hence, the resolvent operator is a κ-strict pseudo contraction and the result follows by [13, Theorem 3.1], and the fact that fixed points of the resolvent operators are exactly the zeroes of A. $\square$

Remark 2.1. It should be noticed that, in the all cases considered in this paper, the resolvent operator is single-valued. This follows directly from (2.5) when $\lambda \geqslant 2\alpha$ and in the other cases from (2.9) and [13], Proposition 2.1(i) showing that every κ-strict pseudo-contraction is Lipschitzian continuous with constant $\frac{1+\kappa}{1-\kappa}$. $\square$

Acknowledgment

We would like to thank the reviewers for their careful reading of the paper. This work is supported by the French Ministry of Higher Education and Research, and by the National Natural Science Foundation of China (NSFC Grant No. 10871092).

References

[1] Zhenyu Huang, M. Aslam Noor, An explicit projection method for a system of nonlinear variational inequalities with different (γ, r)-cocoercive mappings, Appl. Math. Comput. 190 (2007) 356–361.
[2] Zhenyu Huang, M. Aslam Noor, Some new unified iteration schemes with errors for nonexpansive mappings and variational inequalities, Appl. Math. Comput. 194 (2007) 135–142.
[3] M. Aslam Noor, Zhenyu Huang, Wiener-Hopf equation technique for variational inequalities and nonexpansive mappings, Appl. Math. Comput. 191 (2007) 504–510.
[4] M. Aslam Noor, Zhenyu Huang, Some resolvent iterative methods for variational inclusions and nonexpansive mappings, Appl. Math. Comput. 194 (2007) 267–275.
[5] R.U. Verma, Generalized system for relaxed cocoercive variational inequalities and its projection methods, J. Optim. Theory Appl. 121 (2004) 203–210.
[6] M.V. Solodov, B.F. Svaiter, Forcing strong convergence of proximal point iterations in a Hilbert space, Math. Programm. 87 (2000) 189–202.
[7] R.P. Agarwal, R.U. Verma, Role of relative A-maximal monotonicity in overrelaxed proximal-point-algorithms with applications, J. Optim. Theory Appl. 143 (2009) 1–15.
[8] R.P. Agarwal, R.U. Verma, Inexact A-proximal-point-algorithm and applications to nonlinear variational inclusion problems, J. Optim. Theory Appl. (in press), DOI:10.1007/s10957-009-9615-3.
[9] R.U. Verma, A new relaxed proximal point procedure and applications to nonlinear variational inclusions, Comput. Math. Appl. 58 (2009) 1631–1635.
[10] P.L. Combettes, V.R. Wajs, Signal recovery by proximal forward–backward splitting, Multicale Model. Simul. 4 (2005) 1168–1200.
[11] A.N. Iusem, T. Pennanen, B.F. Svaiter, Inexact variants of the proximal point algorithm without monotonicity, SIAM J. Optim. 13 (2003) 1080–1097.
[12] R.T. Rockafellar, Monotone operators and the proximal point algorithm, SIAM J. Control Optim. 14 (1976) 877–898.
[13] G. Marino, H.K. Xu, Weak and strong convergence theorems for strict pseudo-contractions in Hilbert spaces, J. Math. Anal. Appl. 329 (1) (2007) 336–346.