Journal of Mathematical Research with Applications
Mar., 2022, Vol. 42, No. 2, pp. 199–205
DOI:10.3770/j.issn:2095-2651.2022.02.010
Http://jmre.dlut.edu.cn

On a New Analysis Framework for the Linear Convergence of Relaxed Operator Splitting Methods

Qingjing LI, Ke GUO*
School of Mathematics and Information, China West Normal University, Sichuan 637002*, P. R. China*

Abstract In this paper, we propose a new analysis framework to study the linear convergence of relaxed operator splitting methods, which can be viewed as an extension of the classic Krasnosel'skiĭ-Mann iteration and Banach-Picard contraction. As applications, we derive the linear convergence of the generalized proximal point algorithm and the relaxed forward-backward splitting method in a simple and elegant way.

Keywords averaged operator; negatively averaged operator; relaxed forward-backward splitting method; proximal point algorithm

MR(2020) Subject Classification 47H05; 47H09

1. Introduction

Let $\mathcal{H}$ be a real Hilbert space with inner product $\langle\cdot,\cdot\rangle:\mathcal{H}\times\mathcal{H}\to\mathcal{R}$ and its induced norm $\|x\|:=\sqrt{\langle x,x\rangle}$. A mapping $T:\mathcal{H}\to\mathcal{H}$ is nonexpansive if

$$\|T(x)-T(y)\|\le\|x-y\|,\quad \forall x,y\in\mathcal{H}.$$

Nonexpansive operators play a central role in modern optimization because the set of its fixed points often represents solutions to optimization or inclusion problems [1]. In the process of the research of iterative methods for finding fixed points of nonexpansive mappings, one of the most famous fixed point methods is the Krasnosel'skiĭ-Mann iteration [2, 3] and its iteration is as follows:

$$x_{n+1}=(1-\kappa_1)x_n+\kappa_1T(x_n), \tag{1.1}$$

which admits the weak convergence when $\kappa_1\in(0,1)$. If T is further assumed to be a firmly nonexpansive operator (also called averaged with constant $\frac{1}{2}$, see Definition 2.1), the weak convergence of $\{x_n\}$ generated by (1.1) can still be guaranteed when $\kappa_1\in(0,2)$. For the convergence and sublinear convergence rate of (1.1), one can consult [4–8].

We now turn to the linear convergence rate of (1.1). On one hand, by means of the well-known Banach contraction theorem [1], it is natural to assume T is contractive. In this situation,

Received March 20, 2021; Accepted October 16, 2021
Supported by the National Natural Science Foundation of China (Grant Nos. 11801455; 11871059; 11971238), China Postdoctoral Science Foundation (Grant Nos. 2019M663459; 2020T130081), the Applied Basic Project of Sichuan Province (Grant No. 2020YJ0111) and the Open Project of Key Laboratory of School of Mathematical Sciences, Chongqing Normal University (Grant No. CSSXKFKTM202004).
* Corresponding author
E-mail address: liqingjing0120@163.com (Qingjing LI); keguo2014@126.com (Ke GUO)

we know that $(1-\kappa_1)I+\kappa_1 T$ is contractive if $\kappa_1 \in (0, 1+\frac{\mu\lambda}{2+\mu\lambda})$, for example [9, Lemma 3.3]. Since contraction operator is firmly nonexpansive, there is a gap for the interval $[1+\frac{\mu\lambda}{2+\mu\lambda}, 2)$. On the other hand, if T is negatively averaged (see Definition 2.2), it is known from [9] that the sequence of $\{x_n\}$ generated by (1.1) is linearly convergent when $\kappa_1 \in (0,2)$. From the perspective of linear convergence, it is plausible that the negative averagedness assumption seems to be more suitable than the conventional contraction assumption. Indeed, we should emphasize that the recently proposed concept of negatively averaged operator plays a vital role in analyzing the linear convergence of many classic operator splitting methods [9–11].

In this paper, we propose and study the linear convergence of the following algorithm

$$x_{n+1} = (1-\kappa_1-\kappa_2)x_n + \kappa_1 T_1(x_n) + \kappa_2 T_2(x_n), \tag{1.2}$$

where $\kappa_1, \kappa_2 \in (0,1)$, T_1 is averaged with constant $\alpha_1 > 0$ and T_2 is negatively averaged with constant $\alpha_2 \in (0,1)$. Obviously, (1.2) can be viewed as an extension of (1.1). However, the main motivation for considering (1.2) roots in the relaxed forward-backward splitting method, whose linear convergence was established in [10] recently. Besides, the classic generalized proximal point algorithm (to be discussed in Section 4.1) can also be viewed as the special case of (1.2).

The rest of this paper is organized as follows. In Section 2, we recall some important definitions and some known results for further analysis. The linear convergence of (1.2) is established in Section 3. As applications, we recover the linear convergence of the generalized proximal point algorithm and the relaxed forward-backward splitting method in Section 4.

2. Preliminaries

This section contains some important definitions and basic results that will be used in our subsequent analysis.

Definition 2.1 ([1, Definition 4.33]) *Let D be a nonempty subset of $\mathcal{H}$, and $T: D \to \mathcal{H}$ be nonexpansive, and let $\alpha \in (0,1)$. Then T is averaged with constant α if there exists a nonexpansive S such that*

$$T = (1-\alpha)I + \alpha S.$$

Definition 2.2 ([9, Definition 3.7]) *Let D be a nonempty subset of $\mathcal{H}$, and $T: D \to \mathcal{H}$. Then $T: D \to \mathcal{H}$ is negatively averaged with constant $\theta \in (0,1)$, if $-T$ is averaged with constant θ, that is $T = (\theta-1)I + \theta S$ for some nonexpansive mapping S.*

Definition 2.3 ([1, Definition 2.31]) *Given $A: \mathcal{H} \rightrightarrows \mathcal{H}$, the resolvent and the reflected operators of A are denoted by*

$$J_A = (I+A)^{-1} \quad \text{and} \quad R_A = 2J_A - Id,$$

respectively.

Definition 2.4 ([1, Definition 4.10]) *Let D be a nonempty subset of $\mathcal{H}$ and $\beta > 0$. Then*

$B : D \to \mathcal{H}$ is cocoercive with constant β if

$$\langle x - y, B(x) - B(y)\rangle \geq \beta \|B(x) - B(y)\|^2, \quad \forall x, y \in D.$$

Lemma 2.5 ([1, Proposition 4.39]) *Let $B : \mathcal{H} \to \mathcal{H}$, $\beta > 0$ and $\lambda \in (0, 2\beta)$. Then B is cocoercive with constant β if and only if $I - \lambda B$ is averaged with constant $\frac{\lambda}{2\beta}$.*

Definition 2.6 ([1, Definition 22.1]) *Let $\mu > 0$. An operator $A : \mathcal{H} \rightrightarrows \mathcal{H}$ is strongly monotone with constant μ if*

$$\langle u - v, x - y\rangle \geq \mu \|x - y\|^2, \quad \forall\, x, y \in \mathcal{H},\ u \in A(x),\ v \in A(y).$$

Lemma 2.7 ([10, Lemma 3.1]) *Let $\lambda > 0$. Then $A : \mathcal{H} \rightrightarrows \mathcal{H}$ is strongly monotone with constant μ if and only if $R_{\lambda A}$ is negatively averaged with constant $\frac{1}{1+\lambda\mu}$.*

Theorem 2.8 ([1, Theorem 1.50]) *Let (X,d) be a complete metric space and let $T : X \to X$ be Lipschitz continuous with constant $\sigma \in (0, 1)$. Given $x_0 \in X$, set*

$$x_{n+1} = T(x_n).$$

Then there exists $x \in X$ such that x is the unique fixed point of T. Moreover, $\{x_n\}$ converges linearly to x.

3. Linear convergence for (1.2)

In this section, we prove the linear convergence of (1.2). For convenience, we set

$$T := (1 - \kappa_1 - \kappa_2)I + \kappa_1 T_1 + \kappa_2 T_2.$$

Next, we need to prove the contractiveness of the operator T.

Theorem 3.1 *Let T_1 be averaged with constant $\alpha_1 > 0$ and T_2 be negatively averaged with constant $\alpha_2 > 0$. Assume that*

$$\kappa_1 \in [0, 1), \kappa_2 \in (0, 1), \alpha_2 \in (0, 1), \kappa_2 + \kappa_1\alpha_1 < 1. \tag{3.1}$$

Then the operator T is contractive. More precisely,

- *If $1-2\kappa_2-\kappa_1\alpha_1+\kappa_2\alpha_2 \geq 0$, then the operator T is contractive with constant $1-2\kappa_2+2\kappa_2\alpha_2$.*
- *If $1-2\kappa_2-\kappa_1\alpha_1+\kappa_2\alpha_2 < 0$, then the operator T is contractive with constant $2\kappa_2-1+2\kappa_1\alpha_1$.*

Proof Note that, by definition we have

$$\begin{aligned} T &= (1 - \kappa_1 - \kappa_2)I + \kappa_1 T_1 + \kappa_2 T_2 \\ &= (1 - \kappa_1 - \kappa_2)I + \kappa_1((1 - \alpha_1)I + \alpha_1 S_1) + \kappa_2((\alpha_2 - 1)I + \alpha_2 S_2) \\ &= (1 - 2\kappa_2 - \kappa_1\alpha_1 + \kappa_2\alpha_2)I + \kappa_1\alpha_1 S_1 + \kappa_2\alpha_2 S_2, \end{aligned}$$

where S_1, S_2 are nonexpansive operators. Thus, for any $x, y \in \mathcal{H}$, it holds that

$$\begin{aligned} \|Tx - Ty\| &= \|(1 - 2\kappa_2 - \kappa_1\alpha_1 + \kappa_2\alpha_2)(x - y) + \kappa_1\alpha_1(S_1(x) - S_1(y)) + \kappa_2\alpha_2(S_2(x) - S_2(y))\| \\ &\leq |1 - 2\kappa_2 - \kappa_1\alpha_1 + \kappa_2\alpha_2| \cdot \|x - y\| + \kappa_1\alpha_1 \cdot \|S_1(x) - S_1(y)\| + \kappa_2\alpha_2 \cdot \|S_2(x) - S_2(y)\| \end{aligned}$$

$$\leq |1 - 2\kappa_2 - \kappa_1\alpha_1 + \kappa_2\alpha_2| \cdot \|x - y\| + \kappa_1\alpha_1 \cdot \|x - y\| + \kappa_2\alpha_2 \cdot \|x - y\|$$
$$= (|1 - 2\kappa_2 - \kappa_1\alpha_1 + \kappa_2\alpha_2| + \kappa_1\alpha_1 + \kappa_2\alpha_2) \cdot \|x - y\|.$$

Now, we divide the proof into two cases.

• If $1 - 2\kappa_2 - \kappa_1\alpha_1 + \kappa_2\alpha_2 \geq 0$, then we have

$$0 < 1 - 2\kappa_2 - \kappa_1\alpha_1 + \kappa_2\alpha_2 + \kappa_1\alpha_1 + \kappa_2\alpha_2 = 1 + 2\kappa_2(\alpha_2 - 1) < 1, \quad (3.2)$$

due to $\alpha_2 \in (0,1)$. This means T is contractive with $1 - 2\kappa_2 + 2\kappa_2\alpha_2$.

• If $1 - 2\kappa_2 - \kappa_1\alpha_1 + \kappa_2\alpha_2 < 0$, we know

$$0 < 2\kappa_2 - 1 + \kappa_1\alpha_1 - \kappa_2\alpha_2 + \kappa_1\alpha_1 + \kappa_2\alpha_2 = 2(\kappa_2 + \kappa_1\alpha_1) - 1 < 1, \quad (3.3)$$

due to $\kappa_2 + \kappa_1\alpha_1 < 1$. This means T is contractive with $2\kappa_2 - 1 + 2\kappa_1\alpha_1$.

Thus, the proof is completed. □

Now, the linear convergence rate for (1.2) is stated in the following theorem.

Theorem 3.2 *Let $\{x_n\}$ be the sequence generated by (1.2) and assume that (3.1) holds. Then $\{x_n\}$ converges to the unique fixed point of T linearly. More precisely,*

• *If $1 - 2\kappa_2 - \kappa_1\alpha_1 + \kappa_2\alpha_2 \geq 0$, then we have*

$$\|x_n - x^*\| \leq (1 - 2\kappa_2 + 2\kappa_2\alpha_2)^n \cdot \|x_0 - x^*\|.$$

• *If $1 - 2\kappa_2 - \kappa_1\alpha_1 + \kappa_2\alpha_2 < 0$, then we have*

$$\|x_n - x^*\| \leq (2\kappa_2 - 1 + 2\kappa_1\alpha_1)^n \cdot \|x_0 - x^*\|.$$

Proof We know from Theorem 3.1 that the operator $(1 - \kappa_1 - \kappa_2)I + \kappa_1 T_1 + \kappa_2 T_2$ is contractive. Then, it follows from Theorem 2.8 that the conclusions are true. The proof is thus completed. □

4. Applications

Next, we will give the applications as follows.

4.1. Generalized proximal point algorithm

Let $A : \mathcal{H} \rightrightarrows \mathcal{H}$ be a maximal monotone operator. A fundamental problem is finding a zero point of A:

$$0 \in A(x). \quad (4.1)$$

We assume that the solution set of problem (4.1) is nonempty. The proximal point algorithm (PPA), which traces back to [12, 13], has been playing an important role both theoretically and algorithmically for (4.1). Starting from an arbitrary point $x_0 \in \mathcal{H}$, the iterative scheme of PPA reads as

$$x_{n+1} := J_{\lambda A}(x_n), \quad (4.2)$$

where $J_{\lambda A} := (I + \lambda A)^{-1}$ denotes the resolvent operator of the maximal monotone operator A. Since $J_{\lambda A}$ is an averaged operator, it is known that sequence $\{x_n\}$ generated by (4.2) converges weakly to a point in $\mathrm{Fix}(J_{\lambda A}) = \mathrm{zero}(A)$. Under the assumption that A is strongly monotone with constant μ, the author in [13] showed that PPA (4.2) is linear convergence. This is clear because in this case $J_{\lambda A}$ is contractive in the sense that

$$\|J_{\lambda A}(x) - J_{\lambda A}(y)\| \leq \frac{1}{1+\mu\lambda} \cdot \|x-y\|, \quad \forall x, y \in \mathcal{H}.$$

As studied by [14, 15], the classic PPA (4.2) can be further relaxed to the following form:

$$x_{n+1} := (1-\gamma)x_n + \gamma J_{\lambda A}(x_n), \tag{4.3}$$

where $\gamma \in (0,2)$. Algorithm (4.3) is known as the generalized PPA (GPPA) in the literature. Recall that $J_{\lambda A}$ is averaged with constant $\frac{1}{2}$. For $\gamma \in (0,2)$, the sequence generated by (4.3) converges weakly to a zero point of A. It is natural to ask whether or not the sequence generated by (4.3) is linearly convergent under the assumption A is strongly monotone when $\gamma \in (0,2)$. If we set $T := (1-\gamma)I + \gamma J_{\lambda A}$, in view of $J_{\lambda A}$ is contractive with constant $\frac{1}{1+\mu\lambda}$, one has T is contractive if $\gamma \in (0, 1+\frac{\mu\lambda}{2+\mu\lambda})$, for example, see [9, Lemma 3.3]. Thus, for the case $\gamma \in [1+\frac{\mu\lambda}{2+\mu\lambda}, 2)$, it is not obvious to deduce the contraction of T if we directly use the contraction of $J_{\lambda A}$. In 2013, with the help of the properties of the Yosida approximation operator [1, Definition 23.1] of maximal monotone operator, Corman and Yuan [16] successfully showed the linear convergence for the GPPA (4.3) under the above mentioned assumptions. More precise, they proved the following results.

Theorem 4.1 ([16, Theorem 6.2]) *Let A be strongly monotone with constant μ. Let $\{x_n\}$ be the sequence generated by the GPPA (4.3) with $\gamma \in (0,2)$. Then $\{x_n\}$ converges to a root of A on a linear rate. More precisely, for x^* satisfying $0 \in A(x^*)$, we have the following:*

- *If $0 < \gamma \leq 1 + \frac{1}{1+2\mu\lambda}$, then $\|x_n - x^*\| \leq K^n \cdot \|x_0 - x^*\|$, where $K := \left|1 - \frac{\gamma\lambda\mu}{1+\lambda\mu}\right|$.*
- *If $1 + \frac{1}{1+2\mu\lambda} \leq \gamma < 2$, then $\|x_n - x^*\| \leq |\gamma - 1|^n \cdot \|x_0 - x^*\|$.*

Unlike the classic PPA (4.2), we see that the analysis for the linear convergence of GPPA (4.3) is more complicated. In essential, they divided the proof into three cases with careful analysis [16]. Now, we give a new simple and short proof of the linear convergence of GPPA (4.3). Indeed, the proof strategy relies on reformulating the GPPA (4.3) as the special case of (1.2).

Proof Based on the reflected resolvent operator $R_{\lambda A}$, we can rewrite the GPPA (4.3) as

$$x_{n+1} = \left(1 - \frac{\gamma}{2}\right) x_n + \frac{\gamma}{2} \cdot R_{\lambda A}(x_n). \tag{4.4}$$

Since A is strongly monotone with constant μ, it follows from Lemma 2.7 that $R_{\lambda A}$ is $\frac{1}{1+\lambda\mu}$ negatively averaged. Thus, (4.4) can be viewed as a special case of (1.2) with

$$T_1 = 0, \ T_2 = R_{\lambda A}, \ \kappa_1 = 0, \ \kappa_2 = \frac{\gamma}{2}.$$

In this setting, $\alpha_1 = 0$ and $\alpha_2 = \frac{1}{1+\lambda\mu}$. Clearly, the assumption (3.1) holds if $\gamma \in (0,2)$. By Theorem 3.2, we know $\{x_n\}$ is linearly convergent. Note that

$$1 - 2\kappa_2 - \kappa_1\alpha_1 + \kappa_2\alpha_2 = 1 - \gamma + \frac{\gamma}{2} \cdot \frac{1}{1+\lambda\mu}.$$

Then, $1 - 2\kappa_2 - \kappa_1\alpha_1 + \kappa_2\alpha_2 \geq 0$ is equivalent to $\gamma \leq 1 + \frac{1}{1+2\mu\lambda}$. In this case, the contraction parameter is $1 - 2\kappa_2 + 2\kappa_2\alpha_2 = 1 - \frac{\gamma\lambda\mu}{1+\lambda\mu}$. On the other hand, if $1 + \frac{1}{1+2\mu\lambda} \leq \gamma < 2$, the contraction parameter is $2\kappa_2 - 1 + 2\kappa_1\alpha_1 = \gamma - 1$. The proof is thus completed. □

4.2. Relaxed forward-backward splitting method

As another application of Theorem 3.1, we recover the linear convergence of the relaxed forward-backward splitting method (FBSM) recently established in [10] for solving the following zero point problem

$$0 \in A(x) + B(x),$$

where $A : \mathcal{H} \to \mathcal{H}$ is strongly monotone with constant μ and $B : \mathcal{H} \to \mathcal{H}$ is cocoercive with constant β. Recall that, the relaxed FBSM reads as,

$$x_{n+1} = (1-\gamma)x_n + \gamma J_{\lambda A}(I - \lambda B)(x_n). \tag{4.5}$$

The following theorem was established in [10].

Theorem 4.2 ([10, Theorem 3.2]) *Let A be strongly monotone with constant μ and B be cocoercive with constant β. Let $\{x_n\}$ be the sequence generated by the relaxed FBSM (4.5). Assume that*

$$0 < \gamma < 2,\ 0 < \lambda < \min\{2\beta, \frac{2\beta(2-\gamma)}{\gamma}\}.$$

Then $\{x_n\}$ converges linearly to the optimal solution.

Proof Based on the reflected resolvent operator $R_{\lambda A}$, we can rewrite the relaxed FBSM (4.5) as

$$x_{n+1} = (1-\gamma)x_n + \frac{\gamma}{2}(I - \lambda B)(x_n) + \frac{\gamma}{2}R_{\lambda A}(I - \lambda B)(x_n). \tag{4.6}$$

Note that, we know from Lemma 2.5 that $I - \lambda \mathrm{B}$ is averaged with constant $\frac{\lambda}{2\beta}$ if $\lambda < 2\beta$. Moreover, we know from [10, Theorem 3.1] that $R_{\lambda A}(I - \lambda B)$ is negatively averaged with constant $\frac{\kappa}{\kappa+1}$, where $\kappa = \frac{1}{\lambda\mu} + \frac{\lambda}{2\beta-\lambda}$. Thus, (4.6) can be viewed a special case of (1.2) with

$$\kappa_1 = \kappa_2 = \frac{\gamma}{2},\ T_1 = I - \lambda B,\ T_2 = R_{\lambda A}(I - \lambda B).$$

In this setting, $\alpha_1 = \frac{\lambda}{2\beta}$ and $\alpha_2 = \frac{\kappa}{\kappa+1}$. Then, assumption (3.1) becomes

$$\frac{\gamma}{2} \in (0,1),\ \frac{\gamma}{2} + \frac{\gamma}{2} \cdot \frac{\lambda}{2\beta} < 1, \quad \text{and} \quad \lambda < 2\beta,$$

which is equivalent to

$$0 < \gamma < 2,\ 0 < \lambda < \min\{2\beta, \frac{2\beta(2-\gamma)}{\gamma}\}.$$

Therefore, by assumption we know from Theorem 3.2 that T is contractive and the conclusion follows from Theorem 2.8 immediately. The proof is thus completed. □

References

[1] H. H. BAUSCHKE, P. L. COMBETTES. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces.* Second edition. Springer, Cham, 2017.

[2] M. A. KRASNOSEL'SKIǏ. *Two remarks on the method of successive approximations*. Uspehi Mat. Nauk (N.S.), 1955, **10**: 123–127. (in Russian)

[3] W. R. MANN. *Mean value methods in iteration*. Proc. Am. Math. Soc., 1953, **4**(3): 506–510.

[4] C. KANZOW, Y. SHEHU. *Generalized Krasnosel'skiǐ-Mann-type iterations for nonexpansive mappings in Hilbert spaces*. Comput. Optim. Appl., 2017, **67**(3): 593–620.

[5] W. A. KRIK. *Krasnosel'skiǐ's iteration process in hyperbolic space*. Numer. Funct. Anal. Optim., 1982, **4**(4): 371–381.

[6] S. REICH. *Weak convergence theorems for nonexpansive mappings in Banach spaces*. J. Math. Anal. Appl., 1979, **67**(2): 274–276.

[7] S. REICH, A. J. ZASLAVSKI. *Convergence of Krasnosel'skiǐ-Mann iterations of nonexpansive operators*. Math. Comput. Model., 2000, **32**(11): 1423–1431.

[8] Youcai ZHANG, Ke GUO, Tao WANG. *Generalized Krasnosel'skiǐ-Mann-type iteration for nonexpansive mappings in Banach spaces*. J. Oper. Res. Soc. China, 2021, **9**(1): 195–206.

[9] P. GISELSSON. *Tight global linear convergence rate bounds for Douglas-Rachford splitting*. J. Fixed Point Theory Appl., 2017, **19**(4): 2241–2270.

[10] Ke GUO. *On the linear convergence rate of a relaxed forward-backward splitting method*. Optimization, 2021, **70**(5-6): 1161–1170.

[11] W. M. MOURSI, L. VANDENBERGHE. *Douglas-Rachford splitting for the sum of a Lipschitz continuous and a strongly monotone operator*. J. Optim. Theory Appl., 2019, **183**(1): 179–198.

[12] B. MARTINET. *Regularization dinequations variationelles par approximation ssuccessives*. Revue Francaise dInformatiue et de Recherche Oprationelle, 1970, **4**: 154–159.

[13] R. T. ROCKAFELLAR. *Monotone operators and the proximal point algorithm*. SIAM J. Control Optim., 1976, **14**(5): 877–898.

[14] J. ECKSTEIN, D. P. BERTSEKAS. *On the Douglas-Rachford splitting method and the proximal points algorithm for maximal monotone operators*. Math. Programming, Ser. A, 1992, **55**(3): 293–318.

[15] E. G. GOLSHTEIN, N. V. TRETYAKOV. *Modied Lagrangian in convex programming and thier generalizations*. Math. Program. Study, 1979, **4**: 86–97.

[16] E. CORMAN, Xiaoming YUAN. *A generalized proximal point algorithm and its convergence rate*. SIAM J. Optim., 2014, **24**(4): 1614–1638.