Generalized KM theorems and their applications

Qingzhi Yang and Jinling Zhao
School of Mathematics and LPMC, Nankai University
300071, Tianjin, P.R.China

Abstract

Some algorithms in signal processing and imagine reconstruction may be formulated as Krasnoselski-Mann (KM) iteration form and KM theorem asserts the convergence of this iteration under certain assumptions. We give more general iterative schemes which include KM iteration as a special case and establish the convergence of extended iterations. Based on the generalized KM theorems, some algorithms in more wide scopes are analyzed and treated in new settings.

Key words: fixed point, maximal monotone, variational inequality, split feasibility problem, convex feasibility problem

1 Introduction and preliminary

It is well-known that some algorithms in signal processing and image reconstruction may be written as following form ([8]):

$$x^{k+1} = (1-\alpha_k)x^k + \alpha_k N x^k, \quad k = 0, 1, \cdots \tag{1.1}$$

where N is a *nonexpansive (ne)* operator on a real Hilbert space $\mathcal{H}$ and $\{\alpha_k\}$ is a sequence of real numbers in $(0,1]$. This scheme is called Krasnoselski-Mann (KM) iteration. The Krasnoselski-Mann theorem associated with KM iteration asserts the $\{x^k\}$ generated by (1.1) converges weakly to a fixed point of N provided such fixed points exist and otherwise $\{x^k\}$ is unbounded [6].

The main feature of KM theorem is that it provides a unified frame for analyzing various concrete algorithms. We are stimulated by this idea and try to extend KM iteration in order to analyze and deal with various different algorithms in more broad settings. Precisely speaking, We generalize KM iteration in three directions so that they contain three classes of important algorithms: variable-step (or factor), inexact and perturbed schemes, which are popular and important in many fields such as the zero problem of maximal monotone operator, variational inequality (VI) problem, but can't be cast as KM iteration form. And three corresponding theorems, called generalized KM theorems, are proposed to ensure the convergence of extended KM iterations. Then we apply generalized KM theorems to several algorithms for solving zero of maximal monotone operator, variational inequality problem, split feasibility problem and convex feasibility problem respectively. Not only the convergence results of algorithms discussed are establish in new settings, but also some of them are naturally improved in either iterative form or requirement of relaxed factors.

We begin with some well-known definitions for operators. An operator G on $\mathcal{H}$ is called ν-*co-coercive* [13] (also called *inverse strongly monotone* in [8]), if there is $\nu > 0$ such that

$$\langle Gx - Gy, x - y\rangle \geq \nu \|Gx - Gy\|^2$$

An operator N on $\mathcal{H}$ is called *nonexpansive* (*ne*) if, for all x and y in $\mathcal{H}$, $\|Nx - Ny\| \leq \|x - y\|$; and an operator F on $\mathcal{H}$ is called *firmly nonexpansive* (*fne*, see [5, 8, 12]) if it is 1-*co-coercive*, i.e. for all $x, y \in \mathcal{H}$, $\langle Fx - Fy, x - y\rangle \geq \|Fx - Fy\|^2$.

Lemma 1.1 *[8] An operator N is ne if and only if its complement $I - N$ is $\frac{1}{2}$ -co-coercive; an operator F is fne if and only if its complement $I - F$ is fne; an operator A is av, i.e. $A = (1-\alpha)I + \alpha N$, where $\alpha \in (0,1)$ and N is a ne operator, if and only if its complement $I - A$ is $\frac{1}{2\alpha}$ -co-coercive.*

The KM [8, 15, 17] approach is remarkably useful for finding fixed points of a *ne* operator N. And the KM theorem associated with (1.1) can be stated as follows[8, 9].

Theorem 1.1 *Let N be a ne operator on $\mathcal{H}$. Then the sequence $\{x^k\}$ defined by the iterative step (1.1) converges weakly to a fixed point of N, provided $\sum_{k=0}^{+\infty} \alpha_k(1-\alpha_k) = +\infty$, whenever such fixed points exist.*

In this paper we will furthermore generalize the KM theorem in different cases, and apply the generalized theorems to some important algorithms.

The following lemma is very useful for our later analysis.

Lemma 1.2 *[4, 19] Let $\{\delta_k\}$ and $\{\gamma_k\}$ be nonnegative sequences satisfying $\sum\limits_{k=0}^{+\infty} \delta_k < +\infty$ and $\gamma_{k+1} \leq \gamma_k + \delta_k,\ k = 0, 1, \cdots$. Then $\{\gamma_k\}$ is a convergent sequence.*

This paper is organized as follows. In section 2, we first present generalized KM theorems associated with the extended KM iterations in three directions. In section 3, we apply the generalized theorems to three classes of important algorithms: variable-step(or factor), inexact and perturbed schemes. Finally in section 4, we conclude with further discussions.

2 Generalized KM theorems

In this section , we generalize the KM theorem in three directions and three theorems are presented and then they are applied to three classes of algorithms respectively in the following section.

Throughout this section we assume that $\{N_k\}$ is a family of *ne* operators on a Hilbert space.

Theorem 2.1 *Assume that the nonempty solution set of the concerned problem is the fixed point set of N_k for each $k = 0, 1, \cdots$, and for any "cluster operator" of $\{N_k\}$, the fixed points set of it coincides with that of N. Given any $x^0 \in \mathcal{H}$, the sequence $\{x^k\}$ generated by following iteration*

$$x^{k+1} = (1 - \alpha_k)x^k + \alpha_k N_k x^k, \quad k = 0, 1, \cdots$$

converges weakly to a solution point of N, provided $\alpha_k \in (0, 1)$ and $\sum\limits_{k=0}^{+\infty} \alpha_k(1 - \alpha_k) = +\infty$. Furthermore, if $\{N_k\}$ are fne operators on $\mathcal{H}$, then the sequence $\{x^k\}$ converges weakly to a solution point, provided $\alpha_k \in (0, 2)$ and $\sum\limits_{k=0}^{+\infty} \alpha_k(2 - \alpha_k) = +\infty$.

Proof. Let z be a solution point, then $N_k z = z$. Set $G_k = I - N_k$ and we have $G_k z = 0$. We know from (1.1) that

$$\begin{aligned} \|x^{k+1} - z\|^2 &= \|(1 - \alpha_k)x^k + \alpha_k N_k x^k - z\|^2 \\ &= \|x^k - z\|^2 - 2\alpha_k \langle x^k - z, G_k x^k - G_k z\rangle + \alpha_k^2 \|G_k x^k\|^2 \end{aligned} \quad (2.1)$$

If N_k is a *ne* operator, then from Lemma 1.1, G_k is $\frac{1}{2}$ -*co-coercive*. Consequently

$$\|x^{k+1} - z\|^2 \leq \|x^k - z\|^2 - \alpha_k(1 - \alpha_k)\|G_k x^k\|^2 \quad (2.2)$$

It follows from $\alpha_k \in (0,1)$ that the sequence $\{\|x^k - z\|\}$ is convergent since it is decreasing and bounded from below and hence the sequence $\{x^k\}$ is bounded. Furthermore, as a result of (2.2)

$$\alpha_k(1-\alpha_k)\|G_kx^k\|^2\|G_kx^k\|^2 - \|x^{k+1} - z\|^2$$

hence

$$\sum_{k=0}^{+\infty} \alpha_k(1-\alpha_k)\|G_kx^k\|^2 \leq \|x^0 - z\|^2 < +\infty$$

As $\sum_{k=0}^{+\infty} \alpha_k(1-\alpha_k) = +\infty$, we conclude

$$\liminf_{k\to+\infty} \|G_kx^k\| = \liminf_{k\to+\infty} \|x^k - N_kx^k\| = 0$$

Hence there must be a cluster point x^* of the bounded sequence $\{x^k\}$ and a cluster operator $\bar{N}$ of $\{N_k\}$ such that $x^* = \bar{N}x^*$, which implies that $\{x^*\}$ is a fixed point of N from the assumptiom. Then we may use x^* in place of the arbitrary solution point z. It follows that the sequence $\{\|x^k - x^*\|\}$ converges to zero since that the entire sequence converges and a subsequence converges to zero.

In the case that $\{N_k\}$ are *fne* operators, by Lemma 1.1, it follows from (2.1) that

$$\|x^{k+1} - z\|^2 \leq \|x^k - z\|^2 - \alpha_k(2-\alpha_k)\|G_kx^k\|^2 \tag{2.3}$$

since G_k is *fne*. Thus we can obtain the conclusion of this part from similar analysis as in the first part. Now the proof is complete. □

Remark. In this theorem the sequence of operators $\{N_k\}$ needn't converge to some operator N entirely.

Theorem 2.2 *Let N be a ne operator on $\mathcal{H}$. If $\alpha_k \in (0,1)$ satisfies $\sum_{k=0}^{+\infty} \alpha_k(1-\alpha_k) = +\infty$, then for any $x^0 \in \mathcal{H}$, the sequence $\{x^k\}$ defined by*

$$x^{k+1} = (1-\alpha_k)x^k + \alpha_k y^k \tag{2.4}$$

with

$$\|y^k - Nx^k\| \leq \varepsilon_k \tag{2.5}$$

converges weakly to a fixed point of N, provided $\sum_{k=0}^{+\infty} \alpha_k\varepsilon_k < +\infty$, whenever such fixed points exist. Furthermore, if the operator N is fne on $\mathcal{H}$, then the result is still valid under the assumption that $\alpha_k \in (0,2)$ satisfies $\sum_{k=0}^{+\infty} \alpha_k(2-\alpha_k) = +\infty$.

Proof. Let z be a fixed point of N, i.e. $Nz = z$. Let $G = I - N$ and we have $Gz = 0$.

Set $\bar{x}^{k+1} = (1-\alpha_k)x^k + \alpha_k N x^k$, then we get

$$\|x^{k+1} - \bar{x}^{k+1}\| \leq \alpha_k \varepsilon_k \tag{2.6}$$

by considering the iteration (2.4) and (2.5). It holds that

$$\begin{aligned} \|\bar{x}^{k+1} - z\|^2 &= \|(1-\alpha_k)x^k + \alpha_k N x^k - z\|^2 \\ &= \|x^k - z\|^2 - 2\alpha_k \langle x^k - z, Gx^k - Gz\rangle + \alpha_k^2 \|Gx^k\|^2 \end{aligned} \tag{2.7}$$

If N is a *ne* operator, since, by Lemma 1.1, G is $\frac{1}{2}$ -*co-coercive*, then

$$\|\bar{x}^{k+1} - z\|^2 \leq \|x^k - z\|^2 - \alpha_k(1-\alpha_k)\|G_k x^k\|^2 \tag{2.8}$$

As a result

$$x^{k+1} - z \leq \|x^{k+1} - \bar{x}^{k+1}\| + \|\bar{x}^{k+1} - z\| \leq \alpha_k \varepsilon_k + \|x^k - z\|$$

Combining $\sum_{k=0}^{+\infty} \alpha_k \varepsilon_k < +\infty$, we conclude from Lemma 1.2 that the sequence $\{\|x^k - z\|\}$ is convergent and the sequence $\{x^k\}$ is bounded. Then there is an $M > 0$, such that $\|x^k - z\| \leq M$ for all k.

We see from (2.8) that

$$\begin{aligned} \alpha_k(1-\alpha_k)\|G_k x^k\|^2 &\leq \|x^k - z\|^2 - \|\bar{x}^{k+1} - z\|^2 \\ &= \|x^k - z\|^2 - \|\bar{x}^{k+1} - x^{k+1} + x^{k+1} - z\|^2 \\ &\leq \|x^k - z\|^2 - \|x^{k+1} - z\|^2 + 2\langle \bar{x}^{k+1} - x^{k+1}, x^{k+1} - z\rangle \\ &\leq \|x^k - z\|^2 - \|x^{k+1} - z\|^2 + 2M\alpha_k \varepsilon_k \end{aligned}$$

This leads to

$$\sum_{k=0}^{+\infty} \alpha_k(1-\alpha_k)\|G_k x^k\|^2 \leq \|x^0 - z\|^2 + 2M \sum_{k=0}^{+\infty} \alpha_k \varepsilon_k < +\infty$$

As $\sum_{k=0}^{+\infty} \alpha_k(1-\alpha_k) = +\infty$, we can make a conclusion that

$$\liminf_{k\to+\infty} \|Gx^k\| = \liminf_{k\to+\infty} \|x^k - Nx^k\| = 0$$

Taking into account that the sequence $\{x^k\}$ is bounded, there must be a cluster point x^* such that $x^* = Nx^*$. Then we can use x^* in stead of the fixed point z.

Since the entire sequence $\{\|x^k - x^*\|\}$ is convergent and a subsequence converges to zero, it follows that the sequence $\{\|x^k - x^*\|\}$ converges to zero.

If the operator N is *fne*, then, by Lemma 1.1, we get from (2.7)

$$\|\bar{x}^{k+1} - z\|^2 \leq \|x^k - z\|^2 - \alpha_k(2 - \alpha_k)\|G_k x^k\|^2 \tag{2.9}$$

then we can prove the result of this case by following a similar way as above. □

Remark. In some cases it is impossible or requires too much work to exactly compute Nx^k for some operator N (see [5, 14, 20]), and consequently the efficiency of the method will be seriously affected. It is well know that inexact technology plays an important role in designing efficient, easily implemented algorithms for optimization problems, variational inequality problem and so on. So we give an inexact version of KM iteration which is suitable to analyze some inexact schemes. In addition, this theorem is still valid if we suppose that the solution set of the concerned problem is the fixed point set of N_k for all k, and replace the operator N with N_k in 2.5.

In the following we present a generalized theorem for a sequence of operators $\{N_k\}$ approaching to N.

Theorem 2.3 *Let N and N_k be ne operators on Hilbert space $\mathcal{H}$, for $k = 0, 1, \cdots$, $N_k \to N$ and $\alpha_k \in (0, 1)$ satisfy $\sum\limits_{k=0}^{+\infty} \alpha_k(1 - \alpha_k) = +\infty$. Then the sequence $\{x^k\}$ defined by the iterative step*

$$x^{k+1} = (1 - \alpha_k)x^k + \alpha_k N_k x^k \tag{2.10}$$

converges weakly to a fixed point of N, provided $\sum\limits_{k=0}^{+\infty} \alpha_k D_\rho(N_k, N) < +\infty$ for any given $\rho > 0$, whenever such fixed points exist, where $D_\rho(N_k, N)$ [3] is defined as

$$D_\rho(N_k, N) \triangleq \sup_{\|x\| \leq \rho} \|N_k x - Nx\|$$

Proof. Let z be a fixed point of *ne* operator N, i.e. $Nz = z$. Set $G = I - N$, then G is $\frac{1}{2}$ *-co-coercive* by Lemma 1.1 and $Gz = 0$. Since $\{N_k\}$ are *ne* operators, we have

$$\begin{aligned} \|x^{k+1} - z\| &= \|(1 - \alpha_k)x^k + \alpha_k N_k x^k - (1 - \alpha_k)z - \alpha_k Nz\| \\ &\leq (1 - \alpha_k)\|x^k - z\| + \alpha_k\|N_k x^k - Nz\| \end{aligned}$$

$$
\begin{aligned}
&\leq (1-\alpha_k)\|x^k - z\| + \alpha_k\|N_k x^k - N_k z\| + \alpha_k\|N_k z - Nz\| \\
&\leq \|x^k - z\| + \alpha_k\|N_k z - Nz\| \\
&\leq \|x^k - z\| + \alpha_k D_{\bar\rho}(N_k, N),\ \ for\ \bar\rho \geq \|z\|
\end{aligned}
$$

Taking into account $\sum_{k=0}^{+\infty} \alpha_k D_\rho(N_k, N) < +\infty$, we get from Lemma 1.2 that the sequence $\{\|x^k - z\|\}$ is convergent, and the sequence $\{x^k\}$ is bounded. Then there must be an $M > 0$ and a $\rho \geq \bar\rho > 0$ such that $\|x^k - z\| \leq M$ and $\|x^k\| \leq \rho$, for all k.

Set $\bar{x}^{k+1} = (1-\alpha_k)x^k + \alpha_k N x^k$. Considering the boundedness of $\{x^k\}$ and it follows that

$$
\|x^{k+1} - \bar{x}^{k+1}\| \leq \alpha_k D_\rho(N_k, N) \tag{2.11}
$$

Since G is $\frac{1}{2}$ *-co-coercive*

$$
\begin{aligned}
\|\bar{x}^{k+1} - z\|^2 &= \|x^k - z\|^2 - 2\alpha_k\langle x^k - z, Gx^k - Gz\rangle + \alpha_k^2\|Gx^k\|^2 \\
&\leq \|x^k - z\|^2 - \alpha_k(1-\alpha_k)\|Gx^k\|^2
\end{aligned} \tag{2.12}
$$

hence one get from (2.12) that

$$
\begin{aligned}
\alpha_k(1-\alpha_k)\|Gx^k\|^2 &\leq \|x^k - z\|^2 - \|\bar{x}^{k+1} - z\|^2 \\
&= \|x^k - z\|^2 - \|\bar{x}^{k+1} - x^{k+1} + x^{k+1} - z\|^2 \\
&\leq \|x^k - z\|^2 - \|x^{k+1} - z\|^2 - 2\langle \bar{x}^{k+1} - x^{k+1}, x^{k+1} - z\rangle \\
&\leq \|x^k - z\|^2 - \|x^{k+1} - z\|^2 + 2M\alpha_k D_\rho(N_k, N)
\end{aligned}
$$

Therefore, by combining the last part of the proof of Theorem 2.2, we can obtain the convergence, and we omit it here for brevity. □

3 Applications

3.1 Applications in algorithms for maximal monotone operators

A set valued operator T on a Hilbert space $\mathcal{H}$ is a maximal monotone operator [1, 12, 13], that is, T is monotone, i.e. $\forall x, y \in \mathcal{H}$, $\forall v \in T(x)$, $\forall w \in T(y)$, $\langle v - w, x - y\rangle \geq 0$, and the graph $Gr\ T = \{(x, v) \in \mathcal{H} \times \mathcal{H} | v \in T(x)\}$ is not properly contained in the graph of any other monotone operator. We are interested in the resolution of the inclusion problem

$$
\text{Find } \ x \in \mathcal{H}, \ \text{ such that } \ 0 \in T(x) \tag{3.1}
$$

which appears in a wide variety of equilibrium problems such as convex programming and monotone variational inequalities.

Given any positive scalar c and operator T, $J_{cT} = (I+cT)^{-1}$ is called a *resolvent* of T. We know that an operator T on $\mathcal{H}$ is monotone if and only if its resolvent $J_{cT} = (I+cT)^{-1}$ is *fne*, moreover T is maximal monotone if and only if J_{cT} is *fne* and *dom* $J_{cT} = \mathcal{H}$ (see Theorem 2, [12]).

In fact the zeroes of a monotone operator precisely coincide with the fixed points of its *resolvent*:

Lemma 3.1 *[12] Given any maximal monotone operator T, real number $c > 0$, and $x \in \mathcal{H}$, we have, $0 \in T(x)$ if and only if $J_{cT}(x) = x$.*

Let T be a maximal monotone set-valued operator on $\mathcal{H}$, then this lemma suggests a iterative method called Proximal Point Algorithm (PPA):

$$x^{k+1} = J_{c_k T}(x^k) \tag{3.2}$$

as named by Rockafellar[18].

In order to accelerate the standard PPA, the following Relaxed Proximal Point Algorithm (RPPA) is proposed in [12]

$$x^{k+1} = (1-\rho_k)I + \rho_k J_{c_k T}(x^k) \tag{3.3}$$

where $\bar{c} = \inf\limits_{k\geq 0} c_k > 0$ and $\rho_k \in (0,2)$ is a relaxation factor which is supposed to satisfy

$$R_1 = \inf_{k\geq 0} \rho_k > 0 \quad and \quad R_2 = \sup_{k\geq 0} \rho_k < 2 \tag{3.4}$$

Besides, Eckstein and Bertsekas' convergence theorem also allows the *resolvent* $J_{c_k T}$ to be evaluated approximately so long as the sum of all errors is finite, that is,

$$x^{k+1} = (1-\rho_k)I + \rho_k w_k \tag{3.5}$$

with

$$\|w_k - J_{c_k T}(x^k)\| \leq \varepsilon_k \tag{3.6}$$

where $\sum_{k=0}^{+\infty} \varepsilon_k < +\infty$ besides the above assumptions for c_k and ρ_k.

From Lemma 3.1, we know that for $c_k > 0$ with $\bar{c} = \inf\limits_{k\geq 0} c_k > 0$, $0 \in T(x)$ if and only if $J_{c_k T}(x) = x$ for all k. Combining the fact that $J_{c_k T}$ are *fne* operators, we can straightly get from Theorem 2.1 that the sequence $\{x^k\}$ defined by (3.3) converges weakly to a zero of T under the assumption that $\rho_k \in (0,2)$ and $\sum\limits_{k=0}^{+\infty} \rho_k(2-\rho_k) = +\infty$, which is much more relaxed than (3.4).

Furthermore, by considering Theorem 2.2 and the Remark, we can also affirm that the sequence $\{x^k\}$ generated from (3.5) and (3.6) converges weakly to a zero of T, provided $\rho_k \in (0,2)$, $\sum_{k=0}^{+\infty} \rho_k(2-\rho_k) = +\infty$, and $\sum_{k=0}^{+\infty} \rho_k \varepsilon_k < +\infty$.

For that the factor ρ_k is to accelerate the convergence, our relaxation of restrictions on ρ_k is not nonsense.

3.2 Applications in solving variational inequality problem

A variational inequality (VI) problem is to find $x \in C$, such that

$$\langle f(x), y-x \rangle \geq 0, \ \forall y \in C \tag{3.7}$$

where C is a closed convex subset of R^n and $f(x)$ is a mapping from C to R^n.

The following lemma is well-known in VI field.

Lemma 3.2 *x^* is a solution of (3.7) if and only if $x^* = P_C(x^* - \gamma f(x^*))$ for any given $\gamma > 0$*

where P_C is an orthogonal projection onto C.

If we set $N(x) \stackrel{\triangle}{=} P_C(x - \gamma f(x))$ for $\gamma > 0$, then by considering the properties of orthogonal projections, we can easily verify that

Lemma 3.3 *If f is a mapping that is co-coercive on C with modulus $c > 0$, then the operator N is ne for any $\gamma \in (0, 2c)$.*

Proof. By the non-expansivity property of the orthogonal projection operators, we have that $\forall x, y \in R^n$

$$\begin{aligned} \|Nx - Ny\|_2^2 &= \|P_C(x - \gamma f(x)) - P_C(y - \gamma f(y))\|_2^2 \\ &\leq \|x - \gamma f(x) - (y - \gamma f(y))\|_2^2 \\ &= \|x-y\|_2^2 - 2\gamma\langle x-y, f(x)-f(y)\rangle + \gamma^2 \|f(x)-f(y)\|_2^2 \\ &\leq \|x-y\|_2^2 - \gamma(2c-\gamma)\|f(x)-f(y)\|_2^2 \end{aligned}$$

where the last inequality follows from the co-coercivity of f. When $\gamma \in (0, 2c)$, it obvious that the operator N is *ne*. □

Let$f$ be a mapping that is co-coercive on C with modulus $c > 0$. Assuming that $\gamma_k \in (0, 2c)$, with $\inf_{k \geq 0} \gamma_k > 0$, one can conclude from Theorem 2.1 that the sequence $\{x^k\}$ defined by the following iterative step

$$x^{k+1} = (1 - \alpha_k)x^k + \alpha_k P_C(x^k - \gamma_k f(x^k)) \tag{3.8}$$

converges to a solution of (3.7), provided that $\alpha_k \in (0,1)$ satisfies $\sum_{k=0}^{+\infty} \alpha_k(2-\alpha_k) = +\infty$, whenever the solution set is nonempty.

While in the case of $\alpha_k \equiv 1$ in (3.8), the convergence theorem was established under the assumption that

$$0 < \inf_{k\geq 0} \gamma_k \leq \sup_{k\geq 0} \gamma_k < 2c$$

(see [12]).

We know that in some cases it costs too much work to exactly compute the orthogonal projections, so we present an inexact algorithm as follows

$$x^{k+1} = (1-\alpha_k)x^k + \alpha_k y^k \tag{3.9}$$

with

$$\|y^k - P_C(x^k - \gamma_k f(x^k))\|_2 \leq \varepsilon_k \tag{3.10}$$

We obtain from Theorem 2.2 and its remark that this inexact algorithm generates a sequence converging to a solution of (3.7) as long as $\alpha_k \in (0,1)$ satisfies $\sum_{k=0}^{+\infty} \alpha_k(2-\alpha_k) = +\infty$, and $\sum_{k=0}^{+\infty} \alpha_k \varepsilon_k < +\infty$.

Sometimes we prefer a sequence of nonempty closed convex sets $\{C_k\}$ approaching to C for solving the VI problem. For this purpose, we introduce the following notion of convergence, called Mosco-convergence [2, 16], for sequences of sets in a reflexive Banach space. Let N be the set of natural numbers.

Definition 3.1 *(see Definition-Proposition 3.21, [2]) Let X be a reflexive Banach space and $\{C_k\}_{k\in N}$, C a sequence of subsets of X. The sequence $\{C_k\}_{k\in N}$ is said to be Mosco-convergent to C, i.e. $C_k \overset{M}{\to} C$, if*

$$\begin{cases} (i) & \forall x \in C,\ \exists \{x^k\}_{k\in N},\ x^k \in C_k \ for\ every\ k \in N\ such\ that\ x^k \overset{X_s}{\to} x; \\ (ii) & \forall \{k_j\}_{j\in N},\ \forall \{x^j\}_{j\in N},\ x^j \in C_{k_j}\ for\ every\ j \in N,\ x^j \overset{X_w}{\to} x \ \Rightarrow\ x \in C \end{cases}$$

where X_s and X_w denote the strong and weak topology, respectively. Especially, if $\{C_k\}$ and C are in R^n, then $C_k \overset{M}{\to} C$ is equivalent to

$$\begin{cases} (i) & \forall x \in C,\ \exists \{x^k\}_{k\in N},\ x^k \in C_k \ for\ every\ k \in N\ such\ that\ x^k \to x; \\ (ii) & \forall \{k_j\}_{j\in N},\ \forall \{x^j\}_{j\in N},\ x^j \in C_{k_j}\ for\ every\ j \in N,\ x^j \to x \ \Rightarrow\ x \in C \end{cases}$$

The ρ-distance for closed convex sets $C_1, C_2 \in R^n$ is defined as

$$d_\rho(C_1, C_2) \stackrel{\triangle}{=} \sup_{\|x\|_2 \le \rho} \|P_{C_1}(x) - P_{C_2}(x)\|_2$$

Let C and C_k be nonempty closed convex sets in R^n, then $C_k \stackrel{M}{\to} C$ is equivalent to $d_\rho(C_k, C) \to 0$, for all $\rho \ge 0$ ([3, 19]). The following result is not hard to verify (see Proposition 2.5, [19]).

Proposition 3.1 *Let C and C_k be nonempty closed convex subsets in R^n, for $k = 0, 1, \cdots$. If the sequence $\{C_k\}$ is Mosco-convergent to C and the sequence $\{x^k\}$ converges to x, then it holds*

$$\lim_{k \to +\infty} P_{C_k}(x^k) = P_C(x)$$

We can straightly get from Proposition 3.1 that, if the nonempty closed convex set sequence $\{C_k\}$ Mosco-converges to C in R^n, i.e. $C_k \stackrel{M}{\to} C$, then $P_{C_k} \to P_C$ as $k \to +\infty$.

Suppose $\alpha_k \in (0, 1)$, $\sum_{k=0}^{+\infty} \alpha_k(2 - \alpha_k) = +\infty$, and $\sum_{k=0}^{+\infty} \alpha_k d_\rho(C_k, C) < +\infty$ for any given $\rho > 0$. Then Theorem 2.3 assures the sequence $\{x^k\}$ generated from the following perturbed algorithm

$$x^{k+1} = (1 - \alpha_k)x^k + \alpha_k P_{C_k}(x^k - \gamma_k f(x^k)) \tag{3.11}$$

converges to a solution of (3.7) for $\gamma_k \in (0, 2c)$ and $\inf_{k \ge 0} \gamma_k > 0$.

3.3 Applications for the split feasibility problem

The split feasibility problem (SFP)[7, 8, 10] is to find $x \in C$ with $Ax \in Q$, if such points exist, where A ia a real $M \times N$ matrix, and C and Q are nonempty closed convex sets in R^N and R^M, respectively. Byrne proposed the CQ algorithm in [7] for solving the SFP, and the CQ algorithm is as follows

$$x^{k+1} = P_C(x^k + \gamma A^T(P_Q - I)Ax^k) \tag{3.12}$$

where $\gamma \in (0, 2/\rho(A^T A))$, for $\rho(A^T A)$ the spectral radius of the matrix $A^T A$.

The CQ algorithm converges to a solution of the SFP for any $x^0 \in R^N$, whenever the SFP has solutions. When the SFP has no solutions, the CQ algorithm converges to a minimizer of the function

$$f(x) = \frac{1}{2}\|P_Q Ax - Ax\|_2^2 \tag{3.13}$$

over the set C, provided such constrained minimizers exist.

Lemma 3.4 *[8] Let $K \in R^M$ be a closed convex set. The operator $A^T(I - P_K)A$ is co-coercive with modulus $\nu = 1/\rho(A^TA)$.*

Combining this lemma and Lemma 3.3, we know that if $\gamma \in (0, 2/\lambda)$, with $\lambda = \rho(A^TA)$, then the operator $P_C(I + \gamma A^T(P_Q - I)A)$ is *ne*, hence we get the following iterative step for solving the SFP from Theorem 2.2.

$$x^{k+1} = (1 - \alpha_k)x^k + \alpha_k y^k \tag{3.14}$$

with

$$\|y^k - P_C(x^k + \gamma A^T(P_Q - I)Ax^k)\|_2 \le \varepsilon_k \tag{3.15}$$

where $\gamma \in (0, 2/\lambda)$.

If $\alpha_k \in (0,1)$ satisfies $\sum\limits_{k=0}^{+\infty} \alpha_k(1 - \alpha_k) = +\infty$ and $\sum\limits_{k=0}^{+\infty} \alpha_k\varepsilon_k < +\infty$, then the sequence $\{x^k\}$ converges to a fixed point of the operator $P_C(I + \gamma A^T(P_Q - I)A)$. Moreover, if we take $\gamma_k \in (0, 2/\lambda)$ in place of $\gamma \in (0, 2/\lambda)$ in (3.15), with $\inf\limits_{k\ge 0} \gamma_k > 0$, the sequence $\{x^k\}$ still converges to a fixed point of the operator $P_C(I + \gamma^* A^T(P_Q - I)A)$, where $\gamma^* > 0$ is a cluster point of the bounded sequence $\{\gamma_k\}$.

On the other hand, from Lemma 3.4 we can see that the co-coercive modulus ν has nothing to do with the set K. Let $\{C_k\} \subset R^N$ and $\{Q_k\} \subset R^M$ be nonempty closed convex sets, with $C_k \xrightarrow{M} C$ and $Q_k \xrightarrow{M} Q$. Then $d_\rho(C_k, C) \to 0$ and $d_\rho(Q_k, Q) \to 0$ for all $\rho > 0$. If we denote

$$Tx = P_C(x + \gamma A^T(P_Q - I)Ax)$$

$$T_k x = P_{C_k}(x + \gamma A^T(P_{Q_k} - I)Ax)$$

then, by Proposition 3.1, $T_k \to T$ as $k \to +\infty$, and by Lemma 3.3 and Lemma 3.4, we have that the operators T and T_k are *ne* for $\gamma \in (0, 2/\lambda)$.

For $\forall x \in R^N$, $\|x\|_2 \le \rho$ with $\rho > 0$, considering the non-expansivity of projections, we have

$$\begin{aligned}
\|T_k x - Tx\|_2 &= \|P_{C_k}(x - \gamma A^T(I - P_{Q_k})Ax) - P_C(x - \gamma A^T(I - P_Q)Ax)\|_2 \\
&\le \|P_{C_k}(x - \gamma A^T(I - P_{Q_k})Ax) - P_{C_k}(x - \gamma A^T(I - P_Q)Ax)\|_2 \\
&\quad + \|P_{C_k}(x - \gamma A^T(I - P_Q)Ax) - P_C(x - \gamma A^T(I - P_Q)Ax)\|_2 \\
&\le \gamma\|A^T(I - P_{Q_k})Ax - A^T(I - P_Q)Ax\|_2 \\
&\quad + \|P_{C_k}(x - \gamma A^T(I - P_Q)Ax) - P_C(x - \gamma A^T(I - P_Q)Ax)\|_2
\end{aligned}$$

$$
\begin{aligned}
&\leq \gamma\lambda^{1/2}\|P_{Q_k}Ax - P_Q Ax\|_2 \\
&\quad +\|P_{C_k}(x-\gamma A^T(I-P_Q)Ax) - P_C(x-\gamma A^T(I-P_Q)Ax)\|_2 \\
&\leq 2\lambda^{-1/2}\|P_{Q_k}Ax - P_Q Ax\|_2 \\
&\quad +\|P_{C_k}(x-\gamma A^T(I-P_Q)Ax) - P_C(x-\gamma A^T(I-P_Q)Ax)\|_2
\end{aligned}
$$

It follows that

$$D_\rho(T_k, T) \leq d_{\bar{\rho}}(C_k, C) + 2\lambda^{-1/2} d_{\bar{\rho}}(Q_k, Q) \tag{3.16}$$

where $\bar{\rho} \geq \max\{\|Ax\|_2, \|x-\gamma A^T(I-P_Q)Ax\|_2\}$.

Let's see the following iterative step

$$x^{k+1} = (1-\alpha_k)x^k + \alpha_k P_{C_k}(x^k + \gamma A^T(P_{Q_k} - I)Ax^k) \tag{3.17}$$

where $\gamma \in (0, 2/\lambda)$. If $\alpha_k \in (0,1)$, and $\sum_{k=0}^{+\infty} \alpha_k(1-\alpha_k) = +\infty$, then we obtain from Theorem 2.3 and (3.16) that the sequence $\{x^k\}$ defined by (3.17) converges to a fixed point of the operator T, provided

$$\sum_{k=0}^{+\infty} \alpha_k (d_\rho(C_k, C) + 2\lambda^{-1/2} d_\rho(Q_k, Q)) < +\infty$$

for any given $\rho > 0$.

The special case of $\alpha_k \equiv 1$ is to be considered below. Let z be a fixed point of the operator $T = P_C(I + \gamma A^T(P_Q - I)A)$, then $Tz = z$.

Let $\alpha_k \equiv 1$ in (3.14), then we get an inexact form of the CQ algorithm

$$\|x^{k+1} - P_C(x^k + \gamma A^T(P_Q - I)Ax^k)\|_2 \leq \varepsilon_k \tag{3.18}$$

where $\gamma \in (0, 2/\lambda)$, and we assume $\sum_{k=0}^{+\infty} \varepsilon_k < +\infty$.

Since T is *ne*, we have

$$
\begin{aligned}
\|x^{k+1} - z\|_2 &\leq \|x^{k+1} - Tx^k\|_2 + \|Tx^k - Tz\|_2 \\
&\leq \varepsilon_k + \|x^k - z\|_2
\end{aligned}
$$

Combining $\sum_{k=0}^{+\infty} \varepsilon_k < +\infty$ with Lemma 1.2, we know that the sequence $\{\|x^k - z\|_2\}$ is convergent, and thereby $\{x^k\}$ is bounded.

On the other hand, because T is an *av* operator, i.e. $T = (1-\alpha)I + \alpha N$ for some $\alpha \in (0,1)$ and *ne* operator N. Set $G = I - T$, then $Gz = 0$. Since, by Lemma 1.1, G is $\frac{1}{2\alpha}$ -co-coercive, we have

$$\begin{aligned} \|x^k - z\|_2^2 - \|Tx^k - z\|_2^2 &= 2\langle Gx^k - Gz, x^k - z\rangle - \|Gx^k - Gz\|_2^2 \\ &\geq (\frac{1}{2\alpha} - 1)\|x^k - Tx^k\|_2^2 \end{aligned}$$

Hence

$$\begin{aligned} (\frac{1}{2\alpha} - 1)\|x^k - Tx^k\|_2^2 &\leq \|x^k - z\|_2^2 - \|Tx^k - z\|_2^2 \\ &= \|x^k - z\|_2^2 - \|Tx^k - x^{k+1} + x^{k+1} - z\|_2^2 \\ &\leq \|x^k - z\|_2^2 - \|x^{k+1} - z\|_2^2 + 2\|x^{k+1} - z\|_2\|Tx^k - x^{k+1}\|_2 \\ &\leq \|x^k - z\|_2^2 - \|x^{k+1} - z\|_2^2 + 2M\varepsilon_k \end{aligned}$$

where M is assumed to be the bound of the convergent sequence $\{\|x^k - z\|_2\}$.

Consequently, since $\sum\limits_{k=0}^{+\infty} \varepsilon_k < +\infty$, it is obvious that

$$\sum_{k=0}^{+\infty} \|x^k - Tx^k\|_2 < +\infty$$

As a result

$$\lim_{k\to+\infty} \|x^k - Tx^k\|_2 = 0$$

Let x^* be a cluster point of $\{x^k\}$. Then we have $x^* = Tx^*$. So we can use x^* in place of the arbitrary fixed point z. Since we have proved the entire sequence $\{\|x^k - x^*\|_2\}$ converges and a subsequence converges to zero, we can conclude that $\{x^k\}$ defined by (3.18) converges to a fixed point of T.

Now we consider the case of $\alpha_k \equiv 1$ in (3.17), which is a perturbed form of the CQ algorithm, i.e.

$$x^{k+1} = P_{C_k}(x^k + \gamma A^T(P_{Q_k} - I)Ax^k) \tag{3.19}$$

where $\gamma \in (0, 2/\lambda)$, and we assume

$$\sum_{k=0}^{+\infty}(d_\rho(C_k, C) + 2\lambda^{-1/2} d_\rho(Q_k, Q)) < +\infty$$

for any given $\rho > 0$.

Since $\{T_k\}$ are *ne* operators, we have form (3.19) that

$$\begin{aligned}\|x^{k+1} - z\|_2 &= \|T_k x^k - T_k z + T_k z - Tz\|_2 \\ &\leq \|x^k - z\|_2 + \|T_k - Tz\|_2 \\ &\leq \|x^k - z\|_2 + D_{\bar{\beta}}(T_k, T)\end{aligned}$$

where we take $\bar{\beta} \geq \|z\|_2$.

From (3.16) we know that $\sum\limits_{k=0}^{+\infty}(d_\rho(C_k, C)+2\lambda^{-1/2}d_\rho(Q_k, Q)) < +\infty$ for any given $\rho > 0$ implies $\sum\limits_{k=0}^{+\infty} D_\beta(T_k, T) < +\infty$ for any given $\beta > 0$.

Therefore, combining $\sum\limits_{k=0}^{+\infty} D_{\bar{\beta}}(T_k, T) < +\infty$ with Lemma 1.2, we conclude that the sequence $\{\|x^k - z\|_2\}$ converges and $\{x^k\}$ is bounded. So there must be a sufficiently large number $\beta \geq \bar{\beta} > 0$, such that for any k,

$$\|x^{k+1} - Tx^k\|_2 = \|T_k x^k - Tx^k\|_2 \leq D_\beta(T_k, T)$$

Setting $\varepsilon_k = D_\beta(T_k, T)$ and following the analysis of the inexact form of the CQ algorithm, we can prove that the sequence $\{x^k\}$ defined by (3.19) converges to a fixed point of the operator T.

Therefore, just as the CQ algorithm, all of the inexact and perturbed algorithms converges to a solution of the SFP for any starting point, as long as the SFP has solutions; when the SFP has no solutions, they all converges to a minimizer of the function defined in (3.13) over the set C, provided such constrained minimizers exist.

3.4 Application for the convex feasibility problem

Let $C_1,\ C_2, \cdots,\ C_m$ be closed nonempty convex subsets of Hilbert space $\mathcal{H}$. The convex feasibility problem (CFP) is to find a member of their intersection, if such elements exist [5, 8, 11].

A relaxed method has been proposed by De Pierro and Iusem in [11] for solving the CFP. Given $\lambda_i > 0$ and $\sum\limits_{i=1}^{m} \lambda_i = 1$, let

$$x^{k+1} = (1 - \alpha_k)x^k + \alpha_k Ax^k \tag{3.20}$$

$$Ax \triangleq \sum_{i=1}^{m} \lambda_i P_i x \tag{3.21}$$

where P_i denotes the orthogonal projection onto C_i and

$$0 < \inf_{k\geq 0} \alpha_k \leq \sup_{k\geq 0} \alpha_k < 2 \tag{3.22}$$

Since $\forall x, y \in \mathcal{H}$, $\langle P_i x - P_i y, x - P_i x\rangle \geq 0$ and $\langle P_i y - P_i x, y - P_i y\rangle \geq 0$, we have

$$\langle P_i x - P_i y, x - y - P_i x + P_i y\rangle \geq 0$$

Hence

$$\begin{aligned} & \langle P_i x - P_i y, x - y\rangle \\ = \; & \|P_i x - P_i y\|^2 + \langle P_i x - P_i y, x - y - P_i x + P_i y\rangle \\ \geq \; & \|P_i x - P_i y\|^2 \end{aligned}$$

It follows that the operator P_i is *fne* for $i = 1, \cdots, m$. Combining this fact with $\lambda_i > 0$ and $\sum_{i=1}^{m} \lambda_i = 1$, we obtain

$$\begin{aligned} & \langle Ax - Ay, x - y\rangle \\ = \; & \sum_{i=1}^{m} \lambda_i \langle P_i x - P_i y, x - y\rangle \\ \geq \; & \sum_{i=1}^{m} \lambda_i \|P_i x - P_i y\|^2 \\ = \; & (\sum_{i=1}^{m} \lambda_i)(\sum_{i=1}^{m} \lambda_i \|P_i x - P_i y\|^2) \\ \geq \; & (\sum_{i=1}^{m} \lambda_i^{1/2} \lambda_i^{1/2} \|P_i x - P_i y\|)^2 \\ \geq \; & \|Ax - Ay\|^2 \end{aligned}$$

from which we know that the operator A is *fne*.

Therefore, we can use Theorem 2.1 and loosen the relaxation parameters α_k in (3.22) to

$$\alpha_k \in (0, 2) \;\; \textit{and} \;\; \sum_{k=0}^{+\infty} \alpha_k (2 - \alpha_k) = +\infty \tag{3.23}$$

4 Furthermore discussions

In this paper we extend KM theorem in three directions, then apply the generalized theorems to three classes of algorithms. Several algorithms discussed come from the zero problem of maximal monotone operators, variational inequality problem, the split feasibility problem and convex feasibility problem. We believe that more algorithms can be covered in the new setting and it is also possible to extend KM iteration and corresponding KM theorem to other fields, such as in the case of N being a point-to-set mapping.

References

[1] Alvarez F 2004 Weak convergence of a relaxed and inertial hybrid projection-proximal point algorithm for maximal monotone operators in Hilbert space *SIAM J. Optim.* **14** 773-82

[2] Attouch H 1984 Variational convergence for functions and operaters *Applicable Mathematics Series* Pitman (Boston)

[3] Attouch H and Wets R B 1986 Isometries for the Legendre-Fenchel transform textitTrans. Amer. Math. Soc. **296** 33-60

[4] Auslender A, Teboulle M and Ben-Tiba S 1999 A logarithmic-quadratic proximal method for variational inequalities *Computational Optimization and Applications* **12** 31-40

[5] Bauschke H and Borwein J 1996 On projection algorithms for solving convex feasibility problems *SIAM Rev.* **38** 367-426

[6] Borwein J, Reich S and Shafrir I 1992 Krasnoselski-Mann iterations in normed spaces *Canad. Math. Bull.* **35** 21-8

[7] Byrne C L 2002 Iterative oblique projection onto convex sets and the split feasibility problem *Inverse Problems* **18** 441-53

[8] Byrne C L 2004 A unified treatment of some iterative algorithms in signal processing and image reconstruction *Inverse Problems* **20** 103-20

[9] Combettes P 2000 Fejér monotonicity in convex optimization *Encyclopedia of Optimization* ed C A Floudas and P M Pardalos (Boston, MA: Kluwer)

[10] Censor Y and Elfving T 1994 A multiprojection algorithm using Bregman projections in a product space *Numerical Algorithms* **8** 221-39

[11] De Pierro A R and Iusem A N 1985 Asilmultaneous projections method for linear inequalities *Lin. Alg. Appl.* **64** 243-53

[12] Eckstein J and Bertsekas D P 1992 On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators *Math. Prog.* **55** 293-318

[13] Facchinei F and Pang J S 2003 *Finite-dimensional varia tional inequality and complementarity problems* Vols 1 and 2 (New York: Springer-Verlag)

[14] He B 1999 Inexact implicit methods for monotone general variational inequalities *Mathematical Programming* **86** 199-217

[15] Mann W 1953 Mean value methods in iteration *Proc. Am. Math. Soc.* **4** 506-10

[16] Mosco U 1973 *An introduction to the approximate solution of variational inequalities* C.I.M.E. II Circlo Erice 1971 ed. Cremonese (Rome)

[17] Reich S and Zaslavski A J 2000 Convergence of Krasnoselskii-Mann iterations of nonexpansive operators *Math. Comput. Modelling* **32** 1423-31

[18] Rockafellar R T 1976 Monotone operators and the proximal point algorithm *SIAM Journal on Control and Optimization* **14** 877-98

[19] Santos P D and Scheimberg S *Perturbed projection method for general variational inequalities* http://www.cos.ufrj.br/publicacoes/reltec

[20] Yang Q 2004 The relaxed CQ algorithm solving the split feasibility problem *Inverse problems* **20** 1261-6