RESEARCH

An improved method for solving multiple-sets split feasibility problem

Yunfei Du and Rudong Chen[*]

[*]Correspondence:
tjcrd@yahoo.com.cn
Department of Mathematics, Tianjin
Polytechnic University, Tianjin,
300387, China

Abstract

The multiple-sets split feasibility problem (MSSFP) has a variety of applications in the real world such as medical care, image reconstruction and signal processing. Censor *et al.* proposed solving the MSSFP by a proximity function, and then developed a class of simultaneous methods for solving split feasibility. In our paper, we improve a simultaneous method for solving the MSSFP and prove its convergence.

Keywords: multiple-sets split feasibility problem; convergence

1 Introduction

Throughout this paper, let H be a Hilbert space, $\langle \cdot, \cdot \rangle$ denote the inner product and $\| \cdot \|$ denote the corresponding norm. The multiple-sets split feasibility problem (MSSFP) is a generalization of the split feasibility problem (SFP) and the convex feasibility problem (CFP); see [1]. Let C_i and Q_j be closed convex sets in the N-dimensional and M-dimensional Euclidean spaces, respectively. The MSSFP is to find a vector x^* satisfying

$$x^* \in C := \bigcap_{i=1}^{t} C_i \quad \text{such that } Ax^* \in Q := \bigcap_{j=1}^{r} Q_j, \tag{1}$$

where A is a matrix of $M \times N$, and $t, r > 1$ are integers. When $t = r = 1$, the problem becomes to find a vector $x^* \in C$, such that $Ax^* \in Q$, which is just the two-sets split feasibility problem (SFP) introduced in [2]. The MSSFP has many applications in our real life such as image restoration, signal processing and medical care (*e.g.*, [3–5]). In order to solve the MSSFP, Censor *et al.* considered the MSSFP in the following form:

$$x^* \in C := X \cap \left(\bigcap_{i=1}^{t} C_i \right) \quad \text{and} \quad Ax^* \in Q := \bigcap_{j=1}^{r} Q_j, \tag{2}$$

$X \subseteq R^n$ is a nonempty closed convex set. In fact, (2) is equivalent to (1). Many methods have been developed to solve the SFP or MSSFP. The basic CQ algorithm was proposed by Byrne [6], then it was generalized to MSSFP by Censor [4]. The relaxed CQ algorithm was proposed by Yang [7], the half-space relaxation projection method was proposed by Qu and Xiu [8], and the variable Krasnosel'skii-Mann algorithm was proposed by Xu [9]. These algorithms first converted the problem to an equivalent optimization problem and then solved it *via* some technique from numerical optimization. It is easy to see that the

MSSFP (1) is equivalent to the minimization problem

$$\min\left\{\frac{1}{2}\|x - P_C x\|^2 + \frac{1}{2}\|x - P_Q Ax\|^2\right\},$$

where P_C and P_Q denote the orthogonal projections onto C and Q, respectively. The projections of a point onto C and Q are difficult to compute. In practical applications, however, the projections onto individual sets C_i are more easily calculated than the projection onto the intersection C. For this purpose, Censor *et al.* [4] defined a proximity function $p(x)$ to measure the distance of a point to all sets

$$p(x) := \frac{1}{2}\sum_{i=1}^{t} a_i \|P_{C_i}x - x\|^2 + \frac{1}{2}\sum_{j=1}^{r} b_j \|P_{Q_j}Ax - Ax\|^2, \tag{3}$$

where $a_i > 0$, $b_j > 0$ for all i and j, respectively, and

$$\sum_{i=1}^{t} a_i + \sum_{j=1}^{r} b_j = 1.$$

With the proximity function (3), they proposed an optimization model

$$\min\{p(x) | x \in X\} \tag{4}$$

to approach the (2) and exerted the projection gradient method to solve it with

$$x^{k+1} = P_C\{x^k - s\nabla p(x^k)\},$$

where ∇p denotes the gradient of $p(x)$ and can be shown as follows (see [4]):

$$\nabla p(x) = \sum_{i=1}^{t} a_i(x - P_{C_i}(x)) + \sum_{j=1}^{r} b_j(Ax - P_{Q_j}(Ax)).$$

In this paper, we continue the algorithmic improvement on the constrained MSSFP. More specifically, the constrained MSSFP [10] is to find x^* such that

$$x^* \in X \cap \left(\bigcap_{i=1}^{t} C_i\right) \quad \text{and} \quad Ax^* \in Y \cap \left(\bigcap_{j=1}^{r} Q_j\right). \tag{5}$$

By the same idea of approaching (2) *via* the model (4), we define $p_1 : R^n \to R$ and $p_2 : R^m \to R$ as follows:

$$p_1(x) = \frac{1}{2}\sum_{i=1}^{t} a_i \|x - P_{C_i}x\|^2, \tag{6}$$

$$p_2(y) = \frac{1}{2}\sum_{j=1}^{r} b_j \|y - P_{Q_j}y\|^2. \tag{7}$$

Then we get the following optimization model which can solve (5):

$$\min\{p_1(x) + p_2(Ax)|x \in X, y \in Y\}. \tag{8}$$

It is easy to see that model (8) is nonnegative and with the minimal value zero. So, we can further reformulate (8) into the following separable form:

$$\min\{p_1(x) + p_2(y)|Ax - y = 0, x \in X, y \in Y\}. \tag{9}$$

2 Preliminaries

In this section, we present some concepts and properties of the MSSFP.

Let M be a positive definite matrix. We denote the M-norm by $\|v\|_M = \sqrt{\langle v, Mv \rangle}$. In particular, $\|v\| = \sqrt{\langle v, v \rangle}$ is the Euclidean norm of the vector $v \in R^n$.

Lemma 1 *Let S be a nonempty closed convex subset of R^n. We denote $P_S(\cdot)$ as the projection onto S, i.e.,*

$$P_S(z) = \arg\min\{\|z - x\| | x \in S\}.$$

Then the following properties hold:
 (1) $x \in S \Leftrightarrow P_S(x) = x$;
 (2) $\langle x - P_S(x), z - P_S(x) \rangle \leq 0$, $\forall x \in R^n$ and $\forall z \in S$;
 (3) $\langle x - y, P_S(x) - P_S(y) \rangle \geq \|P_S(x) - P_S(y)\|^2$, $\forall x, y \in R^n$;
 (4) $\|P_S(x) - z\|^2 \leq \|x - z\|^2 - \|P_S(x) - x\|^2$, $\forall x \in R^n$ and $\forall z \in S$;
 (5) $\|P_S(x) - P_S(y)\| \leq \|x - y\|$, $\forall x, y \in R^n$.

Proof See Facchinei and Pang [11, 12]. □

Definition 1 Let F be a mapping from $S \subseteq R^n$ into R^n, then
 (a) F is called monotone on S if

$$\langle F(x) - F(y), x - y \rangle \geq 0, \quad \forall x, y \in S.$$

 (b) F is called strongly monotone on S if there is a $\mu > 0$ such that

$$\langle F(x) - F(y), x - y \rangle \geq \mu \|x - y\|^2, \quad \forall x, y \in S.$$

 (c) F is called co-coercive (or ν-inverse strongly monotone) on S if there is a $\nu > 0$ such that

$$\langle F(x) - F(y), x - y \rangle \geq \nu \|F(x) - F(y)\|^2, \quad \forall x, y \in S.$$

 (d) F is called pseudo-monotone on S if

$$\langle F(y), x - y \rangle \geq 0 \quad \Rightarrow \quad \langle F(x), x - y \rangle \geq 0, \quad \forall x, y \in S.$$

(e) *F* is called Lipschitz continuous on *S* if there exists a constant $L > 0$ such that

$$\big\|F(x) - F(y)\big\| \le L\|x - y\|, \quad \forall x, y \in S$$

and *F* is called nonexpansive iff $L = 1$.

Remark 1 From Lemma 1 and the above definition, we can infer that a monotone mapping is a pseudo-monotone mapping. An inverse strongly monotone mapping is monotone and Lipschitz continuous. A Lipschitz continuous and strongly monotone mapping is a strongly monotone mapping. The projection operator is 1-ism and nonexpansive.

Lemma 2 *A mapping F is* 1*-ism if and only if the mapping* $I - F$ *is* 1*-ism, where I is the identity operator.*

Proof See [3, Lemma 2.3]. □

Remark 2 If *F* is an inverse strongly monotone mapping, then *F* is a nonexpansive mapping.

Definition 2 Let *S* be a nonempty closed convex subset of *H* and x_n be a sequence in *H*, then the sequence x_n is called *Fejér* monotone with respect to *S* if

$$\|x_{n+1} - z\| \le \|x_n - z\|, \quad \forall n \ge 0, \forall z \in S.$$

Lemma 3 *Let* $p_1(x)$ *and* $p_2(x)$ *be defined in* (6)-(7), *then* $\nabla p_1(x)$ *and* $\nabla p_2(y)$ *are both Lipschitz continuous and inverse strongly monotone on X and Y, respectively.*

Proof From the definition (6), $p_1(x)$ is differentiable on *X* and

$$\nabla p_1(x) = \sum_{i=1}^{t} a_i \big(x - P_{C_i}(x)\big).$$

Since the projection operator P_{C_i} is 1-ism (from Remark 1), then from Lemma 2, the operator $I - P_{C_i}$ is 1-ism and is also nonexpansive. So, we have

$$\begin{aligned}
\big\|\nabla p_1(x_1) - \nabla p_1(x_2)\big\| &= \left\|\sum_{i=1}^{t} a_i (I - P_{C_i})(x_1 - x_2)\right\| \\
&\le \sum_{i=1}^{t} a_i \big\|(I - P_{C_i})(x_1 - x_2)\big\| \\
&\le \left(\sum_{i=1}^{t} a_i\right)\|x_1 - x_2\|,
\end{aligned}$$

therefore, $\nabla p_i(x)$ is Lipschitz continuous on *X*, and the Lipschitz constant is $L_1 = \sum_{i=1}^{t} a_i$. It also follows from [13, Corollary 10] that $\nabla p_1(x)$ is $\frac{1}{L_1}$-ism. Similarly, we can prove that $\nabla p_2(y)$ is Lipschitz continuous on *Y*, and the Lipschitz constant is $L_2 = \sum_{j=1}^{r} b_j$, furthermore, it is $\frac{1}{L_2}$-ism. □

For notational convenience, let

$$M = \begin{pmatrix} \tau I_n & & \\ & \sigma I_m & \\ & & \frac{1}{\beta} I_m \end{pmatrix},$$

where τ, σ and β are given positive scalars.

$$F(\omega) = \begin{pmatrix} \nabla p_1(x) \\ \nabla p_2(y) \\ Ax - y \end{pmatrix},$$

$$\xi(\omega, \hat{\omega}) = \begin{pmatrix} \xi_x(\omega, \hat{\omega}) \\ \xi_y(\omega, \hat{\omega}) \\ 0 \end{pmatrix} = \begin{pmatrix} \nabla p_1(x) - \nabla p_1(\hat{x}) \\ \nabla p_2(y) - \nabla p_2(\hat{y}) \\ 0 \end{pmatrix}, \tag{10}$$

$$\eta(\omega) = \begin{pmatrix} \eta_x(\omega) \\ \eta_y(\omega) \\ 0 \end{pmatrix} = \begin{pmatrix} \beta A^T(Ax - y) \\ -\beta(Ax - y) \\ 0 \end{pmatrix}, \tag{11}$$

$$d(\omega, \hat{\omega}) = M(\omega - \hat{\omega}) - \xi(\omega, \hat{\omega}),$$

$$q(\omega, \hat{\omega}) = F(\hat{\omega}) + \eta(\omega),$$

$$\varphi(\omega, \hat{\omega}) = \langle \omega - \hat{\omega}, d(\omega, \hat{\omega}) \rangle + \langle z - \hat{z}, Ax - y \rangle.$$

Furthermore, we let $\omega = (x, y, z)$. Suppose that (x^*, y^*) is an optimal solution of the problem (9). Then the constrained MSSFP (5) is equivalent to finding $\omega^* = (x^*, y^*, z^*) \in W = X \times Y \times R^n$ such that for any $\hat{\omega} = (\hat{x}, \hat{y}, \hat{z}) \in W$, we have

$$\begin{cases} \langle \hat{x} - x^*, \nabla p_1(x^*) \rangle \geq 0; \\ \langle \hat{y} - y^*, \nabla p_1(y^*) \rangle \geq 0; \\ \langle \hat{z} - z^*, Ax^* - y^* \rangle \geq 0. \end{cases} \tag{12}$$

3 Main results

In this section, we will present our method for solving the MSSFP and prove its convergence. Our algorithm is defined as follows:

Algorithm 3.1 Step 1. Give arbitrary $\nu \in (0,1)$, $\beta > 0$, $\gamma \in (0,2)$, $\mu > 1$, $\tau_0 > 0$, $\sigma_0 > 0$ and x_0, y_0, z_0. Let $\varepsilon > 0$ be the error tolerance for an approximate solution and set $k = 1$.

Step 2.

(1) Find the smallest nonnegative integer l_k such that $\tau_k = \mu^{l_k} \tau_{k-1}$ and

$$\hat{x}^k = P_X \left\{ x^k - \frac{1}{\tau_k} \left[\nabla p_1(x^k) + \beta A^T(Ax^k - y^k) \right] \right\}, \tag{13}$$

which satisfies

$$\langle x^k - \hat{x}^k, \nabla p_1(x^k) - \nabla p_1(\hat{x}^k) \rangle + \beta \| Ax^k - A\hat{x}^k \|^2 \leq \tau_k \nu \| x^k - \hat{x}^k \|^2; \tag{13'}$$

(2) Find the smallest nonnegative integer m_k such that $\sigma_k = \mu^{m_k}\sigma_{k-1}$ and

$$\hat{y}^k = P_Y\left\{y^k - \frac{1}{\sigma_k}\left[\nabla p_2(y^k) - \beta(Ax^k - y^k)\right]\right\}, \tag{14}$$

which satisfies

$$\langle y^k - \hat{y}^k, \nabla p_2(y^k) - \nabla p_2(\hat{y}^k)\rangle + \beta\|y^k - \hat{y}^k\|^2 \le \sigma_k v\|y^k - \hat{y}^k\|^2; \tag{14'}$$

(3) Then we define $\hat{z}^k$ by

$$\hat{z}^k = z^k - \beta(A\hat{x}^k - \hat{y}^k). \tag{15}$$

Step 3.
Let $\omega^k = (x^k, y^k, z^k)$, then we get the new iterate ω^{k+1} *via*

$$\omega^{k+1} = \omega^k - \gamma\alpha_k^* d(\omega^k, \hat{\omega}^k), \tag{16}$$

where

$$\alpha_k^* = \frac{\varphi(\omega^k, \hat{\omega}^k)}{\|d(\omega^k, \hat{\omega}^k))\|^2}. \tag{17}$$

Step 4.
If $\frac{p(x^{k+1})}{p(x^0)} \le \varepsilon$, stop. Otherwise, set $k = k + 1$ and go to Step 2.

Remark 3 In fact, from Lemma 3, we know that $\nabla p_1(x)$ is Lipschitz continuous with a constant L_1, then the left-hand side of (13′) satisfies

$$\langle x^k - \hat{x}^k, \nabla p_1(x^k) - \nabla p_1(\hat{x}^k)\rangle + \beta\|Ax^k - A\hat{x}^k\|^2 \le (L_1 + \beta\|A^TA\|)\|x^k - \hat{x}^k\|^2.$$

So, (13′) holds as long as $\tau_k \ge \frac{(L_1 + \beta\|A^TA\|)}{v}$. Since $\frac{(L_1 + \beta\|A^TA\|)}{v} > 0$, it has $\inf\{\tau_k\} > 0$ denoted by $\tau_0 = \inf\{\tau_k\}$. On the other hand, by a similar analysis as in [14, Lemma 3.3], it indicates that $\tau_k \le \frac{(L_1 + \beta\|A^TA\|)}{v^2}$, so we have

$$\tau_0 \le \tau_k \le \tau_{\max} = \frac{(L_1 + \beta\|A^TA\|)}{v^2}, \quad \forall k > 1. \tag{18}$$

Similarly, we can also have

$$\sigma_0 \le \sigma_k \le \sigma_{\max} = \frac{(L_2 + \beta)}{v^2}, \quad \forall k > 1. \tag{19}$$

Next, we analyze the convergence of Algorithm 3.1:

Lemma 4 *Suppose ω^k and $\hat{\omega}^k$ are generated by Algorithm 3.1, and $\omega^* = (x^*, y^*, z^*)$ is a solution of* (12). *Then there exits $m > 0$ for any $k \ge 1$ such that*

$$\langle \omega^k - \omega^*, d(\omega^k, \hat{\omega}^k)\rangle \ge \varphi(\omega^k, \hat{\omega}^k) \ge m\|(\omega^k - \hat{\omega}^2)\|^2.$$

Proof First, we prove $\langle \omega^k - \omega^*, d(\omega^k, \hat{\omega}^k) \rangle \geq \varphi(\omega^k, \hat{\omega}^k)$.

From the property of the projection operator in Lemma 1,

$$\langle x - P_S(x), z - P_S(x) \rangle \leq 0, \quad \forall x \in R^n \text{ and } \forall z \in S.$$

Combining it with (13), we have

$$\left\langle x' - \hat{x}^k, \hat{x}^k - x^k + \frac{1}{\tau_k} \left[\nabla p_1(x^k) + \beta A^T(Ax^k - y^k) \right] \right\rangle \geq 0, \quad \forall x' \in X.$$

Multiplying by τ_k, we get

$$\left\langle x' - \hat{x}^k, \tau_k(\hat{x}^k - x^k) + (\nabla p_1(x^k) - \nabla p_1(\hat{x}^k)) + \nabla p_1(\hat{x}^k) + \beta A^T(Ax^k - y^k) \right\rangle \geq 0, \quad \forall x' \in X.$$

And from the definitions of $\xi_x(\omega^k, \hat{\omega}^k)$, $\eta_x(\omega^k)$ in (10) and (11), it is equivalent to

$$\left\langle x' - \hat{x}^k, \tau_k(\hat{x}^k - x^k) + \xi_x(\omega^k, \hat{\omega}^k) + \nabla p_1(\hat{x}^k) + \eta_x(\omega^k) \right\rangle \geq 0, \quad \forall x' \in X. \tag{20}$$

Similarly, from (14) and (15), we can also get

$$\left\langle y' - \hat{y}^k, \sigma_k(\hat{y}^k - y^k) + \xi_y(\omega^k, \hat{\omega}^k) + \nabla p_2(\hat{y}^k) + \eta_y(\omega^k) \right\rangle \geq 0, \quad \forall y' \in Y \tag{21}$$

and

$$\left\langle z' - \hat{z}^k, A\hat{x}^k - \hat{y}^k + \frac{1}{\beta}(\hat{z}^k - z^k) \right\rangle \geq 0, \quad \forall z' \in R^n. \tag{22}$$

Using the notation defined above, from (20)-(22), we have

$$\left\langle \omega' - \hat{\omega}^k, F(\hat{\omega}^k) + \eta(\omega^k) + \xi(\omega^k, \hat{\omega}^k) - M_k(\omega^k - \hat{\omega}^k) \right\rangle \geq 0, \quad \forall \omega' \in W,$$

namely

$$\left\langle \omega' - \hat{\omega}^k, q(\omega^k, \hat{\omega}^k) - d(\omega^k, \hat{\omega}^k) \right\rangle \geq 0, \quad \forall \omega' \in W. \tag{23}$$

Note that $F(\omega)$ is monotone on W because of the monotonicity of $\nabla p_1(x)$ and $\nabla p_2(y)$. From (12), we have

$$\left\langle \hat{\omega}^k - \omega^*, F(\hat{\omega}^k) \right\rangle \geq \left\langle \hat{\omega}^k - \omega^*, F(\hat{\omega}^*) \right\rangle \geq 0. \tag{24}$$

Consequently,

$$\begin{aligned}
\left\langle \hat{\omega}^k - \omega^*, q(\omega^k, \hat{\omega}^k) \right\rangle &= \left\langle \hat{\omega}^k - \omega^*, F(\hat{\omega}^k) \right\rangle + \left\langle \hat{\omega}^k - \omega^*, \eta(\omega^k) \right\rangle \\
&\geq \left\langle \hat{\omega}^k - \omega^*, \eta(\omega^k) \right\rangle \\
&= \left\langle \hat{x}^k - x^*, \beta A^T(Ax^k - y^k) \right\rangle + \left\langle \hat{y}^k - y^k, -\beta(Ax^k - y^k) \right\rangle \\
&= \left\langle A\hat{x}^k - Ax^*, \beta(Ax^k - y^k) \right\rangle + \left\langle \hat{y}^k - y^k, -\beta(Ax^k - y^k) \right\rangle
\end{aligned}$$

$$
\begin{aligned}
&= \langle \beta(A\hat{x}^k - Ax^* - \hat{y}^k + y^*), Ax^k - y^k \rangle \\
&= \langle \beta(A\hat{x}^k - \hat{y}^k), Ax^k - y^k \rangle \\
&= \langle z^k - \hat{z}^k, Ax^k - y^k \rangle \\
&= \varphi(\omega^k, \hat{\omega}^k) - \langle \omega^k - \hat{\omega}^k, d(\omega^k, \hat{\omega}^k) \rangle,
\end{aligned}
\tag{25}
$$

where the first inequality follows from (24), the second equality follows from the definition of $\hat{\omega}^k$, ω^* and $\eta(\omega^k)$.

Setting $\omega' = \omega^*$ in (23), since $\omega^* \in W$ is a solution, we get

$$
\langle \hat{\omega}^k - \omega^*, d(\omega^k, \hat{\omega}^k) \rangle \geq \langle \hat{\omega}^k - \omega^*, q(\omega^k, \hat{\omega}^k) \rangle.
$$

Then

$$
\begin{aligned}
\langle \omega^k - \omega^*, d(\omega^k, \hat{\omega}^k) \rangle &= \langle \omega^k - \hat{\omega}^k, d(\omega^k, \hat{\omega}^k) \rangle + \langle \hat{\omega}^k - \omega^*, d(\omega^k, \hat{\omega}^k) \rangle \\
&\geq \langle \omega^k - \hat{\omega}^k, d(\omega^k, \hat{\omega}^k) \rangle + \langle \hat{\omega}^k - \omega^*, q(\omega^k, \hat{\omega}^k) \rangle \\
&\geq \varphi(\omega^k, \hat{\omega}^k),
\end{aligned}
$$

where the last inequality follows from (25).

Next, we prove

$$
\varphi(\omega^k, \hat{\omega}^k) \geq m \| \omega^k - \hat{\omega}^k \|^2.
$$

From the definition of $\varphi(\omega^k, \hat{\omega}^k)$ and $d(\omega^k, \hat{\omega}^k)$, we obtain

$$
\begin{aligned}
\varphi(\omega^k, \hat{\omega}^k) &= \langle \omega^k - \hat{\omega}^k, d(\omega^k, \hat{\omega}^k) \rangle + \langle z^k - \hat{z}^k, Ax^k - y^k \rangle \\
&= \langle \omega^k - \hat{\omega}^k, M(\omega^k - \hat{\omega}^k) \rangle - \langle \omega^k - \hat{\omega}^k, \xi(\omega^k - \hat{\omega}^k) \rangle + \langle z^k - \hat{z}^k, Ax^k - y^k \rangle \\
&= \| \omega^k - \hat{\omega}^k \|_{M_k}^2 - \langle x^k - \hat{x}^k, \xi_x(\omega^k, \hat{\omega}^k) \rangle - \langle y^k - \hat{y}^k, \xi_y(\omega^k, \hat{\omega}^k) \rangle \\
&\quad + \langle z^k - \hat{z}^k, Ax^k - y^k \rangle \\
&= \| \omega^k - \hat{\omega}^k \|_{M_k}^2 - \langle x^k - \hat{x}^k, \nabla p_1(x^k) - \nabla p_1(\hat{x}^k) \rangle - \langle y^k - \hat{y}^k, \nabla p_2(y^k) - p_2(\hat{y}^k) \rangle \\
&\quad + \langle z^k - \hat{z}^k, Ax^k - y^k \rangle \\
&\geq \| \omega^k - \hat{\omega}^k \|_{M_k}^2 - \tau_k \nu \| x^k - \hat{x}^k \|^2 + \beta \| Ax^k - A\hat{x}^k \|^2 - \sigma_k \nu \| y^k - \hat{y}^k \|^2 \\
&\quad + \beta \| y^k - \hat{y}^k \|^2 + \langle z^k - \hat{z}^k, A\hat{x}^k - \hat{y}^k \rangle + \langle z^k - \hat{z}^k, Ax^k - A\hat{x}^k - y^k + \hat{y}^k \rangle \\
&\geq \tau_k(1 - \nu) \| x^k - \hat{x}^k \|^2 + \sigma_k(1 - \nu) \| y^k - \hat{y}^k \|^2 + \frac{2}{\beta} \| z^k - \hat{z}^k \|^2 \\
&\quad + \beta \| Ax^k - A\hat{x}^k \|^2 + \langle z^k - \hat{z}^k, Ax^k - A\hat{x}^k \rangle + \beta \| y^k - \hat{y}^k \|^2 \\
&\quad - \langle z^k - \hat{z}^k, y^k - \hat{y}^k \rangle,
\end{aligned}
\tag{26}
$$

where the first inequality follows from (13′) and (14′).

Note that

$$
\beta \| Ax^k - A\hat{x}^k \|^2 + \langle z^k - \hat{z}^k, Ax^k - A\hat{x}^k \rangle = \beta \left\| Ax^k - A\hat{x}^k + \frac{1}{2\beta}(z^k - \hat{z}^k) \right\|^2 - \frac{1}{4\beta} \| z^k - \hat{z}^k \|^2
$$

and

$$\beta \left\| y^k - \hat{y}^k \right\|^2 - \langle z^k - \hat{z}^k, y^k - \hat{y}^k \rangle = \beta \left\| y^k - \hat{y}^k - \frac{1}{2\beta}(z^k - \hat{z}^k) \right\|^2 - \frac{1}{4\beta} \left\| z^k - \hat{z}^k \right\|^2.$$

Substituting them into (26), we have

$$\varphi(\omega^k, \hat{\omega}^k) \geq \tau_k (1 - \nu) \left\| x^k - \hat{x}^k \right\|^2 + \sigma_k (1 - \nu) \left\| y^k - \hat{y}^k \right\|^2$$

$$+ \frac{3}{2\beta} \left\| z^k - \hat{z}^k \right\|^2 + \beta \left\| Ax^k - A\hat{x}^k + \frac{1}{2\beta}(z^k - \hat{z}^k) \right\|^2$$

$$+ \beta \left\| y^k - \hat{y}^k - \frac{1}{2\beta}(z^k - \hat{z}^k) \right\|^2$$

$$\geq \tau_k (1 - \nu) \left\| x^k - \hat{x}^k \right\|^2 + \sigma_k (1 - \nu) \left\| y^k - \hat{y}^k \right\|^2 + \frac{3}{2\beta} \left\| z^k - \hat{z}^k \right\|^2.$$

For the sequences $\{\tau_k\}$ and $\{\sigma_k\}$ are bounded from (18) and (19), let

$$m = \min \left\{ (1 - \nu)\tau_0, (1 - \nu)\sigma_0, \frac{3}{2\beta} \right\},$$

therefore,

$$\varphi(\omega^k, \hat{\omega}^k) \geq m \left\| \omega^k - \hat{\omega}^k \right\|.$$

This completes the proof. □

Next, we prove the sequence ω^k is *Fejér* monotone.

Theorem 1 *Suppose ω^k and $\hat{\omega}^k$ are generated by Algorithm 3.1, and $\omega^* = (x^*, y^*, z^*)$ is a solution of* (12). *Then there exits $C > 0$ for any $k \geq 1$ such that*

$$\left\| \omega^{k+1} - \omega^* \right\|^2 \leq \left\| \omega^k - \omega^* \right\|^2 - \frac{m^2 r(2 - r)}{C^2} \left\| \omega^k - \hat{\omega}^k \right\|^2.$$

Proof From (16), we have

$$\left\| \omega^{k+1} - \omega^* \right\|^2 = \left\| \omega^k - \gamma \alpha_k^* d(\omega^k, \hat{\omega}^k) - \omega^* \right\|^2$$

$$= \left\| \omega^k - \omega^* \right\|^2 - 2\gamma \alpha_k^* \langle \omega^k - \omega^*, d(\omega^k, \hat{\omega}^k) \rangle + (\gamma \alpha_k^*)^2 \left\| d(\omega^k, \hat{\omega}^k) \right\|^2$$

$$\leq \left\| \omega^k - \omega^* \right\|^2 - 2\gamma \alpha_k^* \varphi(\omega^k, \hat{\omega}^k) + \gamma^2 \alpha_k^* \varphi(\omega^k, \hat{\omega}^k)$$

$$= \left\| \omega^k - \omega^* \right\|^2 - \gamma \alpha_k^* (2 - \gamma) \alpha_k^* \varphi(\omega^k, \hat{\omega}^k)$$

$$\leq \left\| \omega^k - \omega^* \right\|^2 - \gamma \alpha_k^* (2 - \gamma) m \left\| \omega^k - \hat{\omega}^k \right\|^2,$$

where the inequalities follow from Lemma 4 and (17).

Because $\|\xi_x(\omega^k, \hat{\omega}^k)\| \leq L_1 \|x^k - \hat{x}^k\|$ and $\|\xi_y(\omega^k, \hat{\omega}^k)\| \leq L_2 \|y^k - \hat{y}^k\|$, and

$$\left\| M_k(\omega^k - \hat{\omega}^k) \right\| = \tau_k \left\| x^k - \hat{x}^k \right\|^2 + \sigma_k \left\| y^k - \hat{y}^k \right\|^2 + \frac{1}{\beta} \left\| z^k - \hat{z}^k \right\|^2$$

$$\leq \tau_{\max} \left\| x^k - \hat{x}^k \right\|^2 + \sigma_{\max} \left\| y^k - \hat{y}^k \right\|^2 + \frac{1}{\beta} \left\| z^k - \hat{z}^k \right\|^2,$$

we have

$$
\begin{aligned}
\big\| d\big(\omega^k, \hat{\omega}^k\big) \big\| &= \big\| M_k\big(\omega^k - \hat{\omega}^k\big) - \xi\big(\omega^k, \hat{\omega}^k\big) \big\| \\
&\leq \big\| M_k\big(\omega^k - \hat{\omega}^k\big) \big\| + \big\| \xi\big(\omega^k, \hat{\omega}^k\big) \big\| \\
&\leq (L_1 + \tau_{\max})\big\| x^k - \hat{x}^k \big\| + (L_2 + \sigma_{\max})\big\| y^k - \hat{y}^k \big\| + \frac{1}{\beta}\big\| z^k - \hat{z}^k \big\|.
\end{aligned}
$$

Let $C = \max\{(L_1 + \tau_{\max}), (L_2 + \sigma_{\max}), \frac{1}{\beta}\}$, then we can get $\|d(\omega^k, \hat{\omega}^k)\| \leq C\|\omega^k - \hat{\omega}^k\|$.

So, from Lemma 4, we can yield

$$
\alpha_k^* = \frac{\varphi(\omega^k, \hat{\omega}^k)}{\|d(\omega^k, \hat{\omega}^k)\|^2} \geq \frac{m}{C^2}.
$$

Therefore,

$$
\big\| \omega^{k+1} - \omega^* \big\|^2 \leq \big\| \omega^k - \omega^* \big\|^2 - \frac{m^2 r(2 - r)}{C^2}\big\| \omega^k - \hat{\omega}^k \big\|^2. \qquad \square
$$

Theorem 2 *The sequence ω^k generated by Algorithm* 3.1 *converges to a solution of* (8).

Proof Suppose ω^* is a solution of (8). It follows from Theorem 1 that

$$
\big\| \omega^{k+1} - \omega^* \big\|^2 \leq \cdots \leq \big\| \omega^k - \omega^* \big\|^2 \leq \big\| \omega^0 - \omega^* \big\|^2, \tag{27}
$$

which means that the sequence ω^k is bounded. Thus, it has at least a cluster point.

Furthermore, Theorem 1 also shows that

$$
\frac{m^2 r(2 - r)}{C^2}\big\| \omega^k - \hat{\omega}^k \big\|^2 \leq \big\| \omega^k - \omega^* \big\|^2 - \big\| \omega^{k+1} - \omega^* \big\|^2.
$$

Summing both sides for all k, we obtain

$$
\frac{m^2 r(2 - r)}{C^2}\sum_{k=0}^{\infty}\big\| \omega^k - \hat{\omega}^k \big\|^2 \leq \sum_{k=0}^{\infty}\big\{ \big\| \omega^k - \omega^* \big\|^2 - \big\| \omega^{k+1} - \omega^* \big\|^2 \big\} \leq \big\| \omega^0 - \hat{\omega}^* \big\|^2,
$$

which means that

$$
\lim_{x \to \infty}\big\| \omega^k - \hat{\omega}^k \big\|^2 = 0.
$$

So, $\{\omega^k\}$ and $\{\hat{\omega}^k\}$ have the same cluster points. Without loss of generality, let $\bar{\omega}$ be a cluster point of $\{\omega^k\}$ and $\{\hat{\omega}^k\}$, $\bar{\tau}$ and $\bar{\sigma}$ be the cluster points of $\{\tau_k\}$ and $\{\sigma_k\}$, respectively. Let $\{\omega^{k_j}\}$, $\{\hat{\omega}^{k_j}\}$, $\{\tau_{k_j}\}$, $\{\sigma_{k_j}\}$ be the subsequences converging to them. Then, by taking limits over the subsequences in (13), (14), (15), we have

$$
\bar{x} = P_X\left\{ \bar{x} - \frac{1}{\tau}\big[\nabla p_1(\bar{x}) + \beta A^T(A\bar{x} - \bar{y})\big] \right\},
$$

$$
\bar{y} = P_Y\left\{ \bar{y} - \frac{1}{\sigma}\big[\nabla p_2(\bar{y}) - \beta(A\bar{x} - \bar{y}))\big] \right\},
$$

$$\bar{z} = \bar{z} - \beta(A\bar{x} - \bar{y}).$$

It then follows from [15] that $\bar{\omega}$ is a solution of (12).

Because of the arbitrary ω^*, we can take $\omega^* = \bar{\omega}$ in (27) and obtain

$$\left\| \omega^{k+1} - \bar{\omega} \right\| \leq \left\| \omega^k - \bar{\omega} \right\|, \quad \forall k \geq 1.$$

Therefore, the whole sequence $\{\omega^k\}$ converges to $\bar{\omega}$. This completes the proof. □

Remark 4 Our iteration method is simpler in the form and is an improvement of the corresponding result of [10].

4 Applications

The multiple-sets split feasibility problem (MSSFP) requires to find a point closest to a family of closed convex sets in one space such that its image under a linear transformation will be closest to another family of closed convex sets in the image space. It serves as a model for real-word inverse problems where constraints are imposed on the solution in the domain of a linear operator as well as in the operator's range.

In this paper, our algorithm converges to a solution of the multiple-sets split feasibility problem (MSSFP), for any starting vector $\omega^0 = (x^0, y^0, z^0)$, whenever the MSSFP has a solution. In the inconsistent case, it finds a point which is least violating the feasibility by being 'closest' to all sets as 'measured' by a proximity function.

In the general case, computing the projection in the MSSFP is difficult to implement. So, Yang [7] solves this problem by the relaxed CQ-algorithm. Without loss of generality, take the two-sets split feasibility problem for instance. He assumes the sets C and Q are nonempty and given by

$$C = \left\{ x \in R^N \,|\, c(x) \leq 0 \right\}, \quad \text{and} \quad Q = \left\{ y \in R^N \,|\, q(y) \leq 0 \right\},$$

where $c : R^N \to R$ and $q : R^M \to R$ are convex functions, respectively. Here he uses the subgradient projections instead of the orthogonal projections. This is a huge achievement and it enables the split feasibility problem to achieve computer operation.

Lastly, we want to say that our work is related to significant real-world applications. The multiple-sets split feasibility problem was applied to the inverse problem of intensity-modulated radiation therapy (IMRT). In this field, beams of penetrating radiation are directed at the lesion (tumor) from external sources in order to eradicate the tumor without causing irreparable damage to surrounding healthy tissues; see, *e.g.*, [4].

References

1. Censor, Y, Segal, A: The split common fixed point problem for directed operators. J. Convex Anal. **16**, 587-600 (2009)
2. Censor, Y, Elfving, T: A multiprojection algorithm using Bregman projections in a product space. Numer. Algorithms **8**, 221-239 (1994)
3. Byrne, C: A unified treatment of some iterative algorithms in signal processing and image reconstruction. Inverse Probl. **20**, 103-120 (2004)
4. Censor, Y, Elfving, T, Kopf, N, Bortfeld, T: The multiple-sets split feasibility problem and its applications for inverse problems. Inverse Probl. **21**, 2071-2084 (2005)
5. Censor, Y, Bortfeld, T, Martin, B, Trofimov, A: A unified approach for inversion problems in intensity-modulated radiation therapy. Phys. Med. Biol. **51**, 2353-2365 (2006)
6. Byrne, C: Iterative oblique projection onto convex sets and the split feasibility problem. Inverse Probl. **18**, 441-453 (2002)
7. Yang, Q: The relaxed CQ algorithm solving the split feasibility problem. Inverse Probl. **20**, 1261-1266 (2004)
8. Qu, B, Xiu, N: A new half space-relaxation projection method for the split feasibility problem. Linear Algebra Appl. **428**, 1218-1229 (2008)
9. Xu, HK: A variable Krasnosel'skii-Mann algorithm and the multiple-set split feasibility problem. Inverse Probl. **22**, 2021-2034 (2006)
10. Zhang, WX, Han, DR, Yuan, XM: An efficient simultaneous method for the constrained multiple-sets split feasibility problem. Comput. Optim. Appl. (2012). doi:10.1007/s10589-011-9429-8
11. Facchinei, F, Pang, JS: Finite-Dimensional Variational Inequalities and Complementarity Problems, vol. I. Springer, Berlin (2003)
12. Facchinei, F, Pang, JS: Finite-Dimensional Variational Inequalities and Complementarity Problems, vol. II. Springer, Berlin (2003)
13. Baillon, J, Haddad, G: Quelques propriétés des opérateurs angel-bornés et n-cycliquement monotones. Isr. J. Math. **26**, 137-150 (1977)
14. Zhang, WX, Han, DR, Li, ZB: A self-adaptive projection method for solving the multiple-sets split feasibility problem. Inverse Probl. **25**, 11 (2009). doi:10.1088/0266-5611/25/11/115001
15. Bertsekas, DP, Tsitsiklis, JN: Parallel and Distributed Computation. Numerical Methods. Prentice Hall, Englewood Cliffs (1989)