Yokohama Publishers
ISSN 1349-8169 ONLINE JOURNAL

Pacific Journal of Optimization
Vol. 20, No. 1, pp. 45–57, 2024

A PROJECTION ALGORITHM FOR NON-CONVEX SPLIT FEASIBILITY PROBLEM*

Yang Yang, Biao Qu† and Ming Qiu

Abstract: In this paper, the non-convex split feasibility problem is considered. We first transform the problem into an unconstrained optimization problem whose objective function is non-convex and non-smooth. Then, combining the projection gradient algorithm, we propose a projection algorithm to solve it. Based on Kurdyka-Lojasiewicz property, we obtain the global convergence of the algorithm, that is, the sequence generated by the algorithm converges to a critical point of the non-convex split feasibility problem.

Key words: *non-convex split feasibility problem, projection gradient algorithm, Kurdyka-Lojasiewicz property, convergence*

Mathematics Subject Classification: *90C26, 90C30*

1 Introduction

Let C and Q be nonempty closed sets in R^n and R^m, respectively, and A is an $m \times n$ real matrix. The split feasibility problem (SFP) then seeks to:

$$x \in C\ ,\ Ax \in Q. \tag{1.1}$$

The convex variant of SFP was introduced by Censor and Elfving [8] in 1994 and employed in solving several real-world problems, such as signal processing and intensity modulated radiation therapy [16, 21]. Therefore, many scholars have carried out a lot of researches and put forward many algorithms to solve it [6, 7, 10–14, 18, 22]. In [6, 7], Byrne presented a projection method called CQ algorithm:

$$x_{k+1} = P_C(x_k - \alpha_k A^T(I - P_Q)Ax_k), k \geq 1,$$

where $\alpha_k \in (0, \frac{2}{L})$, L denotes the largest eigenvalue of the matrix A^TA, I is the identity operator and P_C and P_Q denote the orthogonal projections onto C and Q, respectively. Additionally, Byrne assumed that both projections are easily calculated in the CQ algorithm [6]. However, in some cases, it is impossible to exactly compute the projection. Therefore if this case appears, the inexact projection-type methods will be particularly important. For example, in [22], Yang presented a relaxed CQ algorithm, which used two halfspaces C_k and

*This research was partly supported by the National Natural Science Foundation of China (No. 12271309), and the Natural Science Foundation of Shandong Province (No. ZR2022MA038).

†Corresponding Author

© 2024 Yokohama Publishers DOI: https://doi.org/10.61208/pjo-2023-012

Q_k in place of C and Q, where C_k, Q_k depend on the k-th iterative point x^k. In comparison, it is easy to compute the projections onto C_k and Q_k. Qu and Xiu [18] presented a modified version of the relaxed CQ algorithm by adopting the Armijo-type search that need not to compute the eigenvalues of the matrix $A^T A$.

With the development of technology, applications of the SFP are not limited to convex sets. Many applications are associated with non-convex split feasibility problems, for example, an outlier detection problem [9,17]. This class of problem arises in compressed sensing, such as reconstructing electrocardiogram (ECG) signals in the presence of electromyographic (EMG) noise. Let x_0 and $x = x_0 + n$ be original signal and contaminated signal, respectively, and n is an EMG noise. Here, $b = Ax = Ax_0 + An = b_0 + An$. We use b_i to represent the i-th entry of b, that is, some b_i are recorded incorrectly. Now, we transform the problem into a non-convex split feasibility problem:

$$x \in C\ ,\ Ax \in Q,$$

where $C = \{x \in R^n : ||x||_0 \leq s\}$, $||x||_0$ is the number of nonzero entries of the vector x, $s > 0$ is a given integer regulating the sparsity level and $Q = \{y \in R^m : ||y||_0 \leq r\} + b_0$, with r being an upper estimate of the number of incorrect b_i. Obviously, C and Q are two closed sets but not convex.

Another example is Sudoku puzzle [20]. The objective is to use known information to fill a 9×9 grid so that each column, each row and each of the nine 3×3 sub-grids contain all of the digits from 1 to 9. According to the rules, Sudoku puzzle can be transformed into a non-convex feasibility problem. The first step is data initialization, which is to convert two-dimensional data information into three-dimensional data. Since only nine numbers from 1 to 9 will be placed in the original Sudoku disk, these nine digits will be used as the index information of the third dimension. We now have a $9 \times 9 \times 9$ cube. After conversion, specific digital information is no longer needed on each layer, but only (0,1) is stored. Two numbers respectively mean the absence or presence of a number in the position. The second step is to design variables. Let $x_{i,j,k}, i, j, k \in \{1, 2, \dots, 9\}$ be the digital information at the i-th row, the j-th column and the k-th layer. Finally, we construct constraints. Each digit from 1 to 9 is lifted to the set $[0, 1]^9$, namely, any digit from 1 to 9 is a permutation of unit vector $e = \{1, 0, \dots, 0\}$. This leads to four Sudoku feasibility constraints:

(i) Each row of the cube, *i.e.* $C_1(i, j, :), i, j \in \{1, 2, \dots, 9\}$, is the permutation of e;

(ii) Each column of the cube, *i.e.* $C_2(:, j, k), j, k \in \{1, 2, \dots, 9\}$, is the permutation of e;

(iii)Each pillar of the cube, *i.e.* $C_3(i, :, k), i, k \in \{1, 2, \dots, 9\}$, is the permutation of e;

(iv) Each $k \in \{1, 2, \dots, 9\}$, each of the nine sub-grids is the permutation of e, *i.e.* $C_4(3(i-1) + 1 : 3i, 3(j-1) + 1 : 3j, k), i, j \in \{1, 2, 3\}$;

(v) C_5 is the constraint of the provided numbers.

Then, the Sudoku will be transformed into the following problem: find $x \in R^{s \times s \times s}$, such that

$$x \in C_1 \cap C_2 \cap C_3 \cap C_4 \cap C_5.$$

Let $\mathcal{H} = R^{s\times s\times s} \times \cdots \times R^{s\times s\times s}, \mathcal{C} = C_1 \times \cdots \times C_5, \mathcal{Q} = \{\mathbf{x} = (x_i)_i \in \mathcal{H}, x_1 = \cdots = x_5\}$, where $\mathcal{C}$ is a non-convex set. Eventually, the non-convex feasibility problem is formulated as follows:

$$\mathbf{x} \in \mathcal{C} \cap \mathcal{Q}.$$

It is easy to see that the above problem is a special case of the non-convex split feasibility problem, where the matrix A is identity matrix.

In conclusion, the theoretical research and algorithm design of non-convex split feasibility problem have important practical significance. A large part of the optimization community has been devoted to the development of the related issues and achieved some results.

Chen [9] rewrote objective function via the model using techniques of Difference-of-Convex (DC) programming and utilized NPG_{major} to derive the CQ Algorithm in the fully non-convex setting. Using the indicator function and distance function, Gibali [15] presented four optimization models for the SFP and proposed some corresponding algorithms, such as Projected Gradient (PG) method, Alternating Minimization (AM) method and Weighted Proximal ADMM method. What is more worth mentioning is Semi Alternating Projected Gradient (SAPG) method for the model

$$\min_{x\in R^n, u\in R^m}\{\delta_C(x)+\delta_Q(u)\ :\ Ax=u\},$$

where δ_C, δ_Q are indicator functions of C, Q, respectively. It derives the CQ Algorithm in non-convex setting from the perspective of algorithm.

It is worth noting that the SAPG method uses a fixed stepsize related to the largest eigenvalue of the matrix A^TA. However, if the dimension of the matrix is very large or in some other cases, it needs too much work to compute or estimate the eigenvalue. Although Chen [5] adopted a linear search to avoid the high computational cost, in the theorem of global convergence, he assumed that the objective function is a KL function and the distance function $y \longmapsto \frac{1}{2}d^2(y, D)$ is continuously differentiable at Ax^* with locally Lipschitz gradient, where x^* is an accumulation point of $\{x^t\}$. This greatly reduces the application field of the algorithm. In this paper, we present a projection algorithm with an Armijo-type search to overcome the difficulties. We also show the convergence of the projection algorithm under Kurdyka-Lojasiewicz (KL) property.

The rest of this paper is organized as follows. The mathematical formulation of non-convex split feasibility problem and some preliminaries are given in Section 2. In Section 3, we introduce a projection algorithm with line-search and prove its convergence.

2 The Problem and Some Preliminaries

In this section, we show some preliminaries, definitions, the mathematical formulation of the SFP and KL property, uniformized KL property.

Definition 2.1 ([19]). Given a nonempty closed set $D \subseteq R^n$, the indicator function is defined by

$$\delta_D(x)=\begin{cases}0 & x\in D\\ +\infty & x\notin D\end{cases}.$$

An extended real-valued function $h : R^n \to R\cup\{\infty\}$ is said to be proper if dom $h := \{x : h(x) < \infty\} \neq \emptyset$, and closed if it is lower semicontinuous, *i.e.* $f(x) \leq \liminf_{k\to\infty} f(x^k)$, as $x^k \to x$.

In this paper, we are interested in solving the following non-convex and non-smooth minimization problem

$$\min_{x\in R^n, y\in R^m}\{\psi(x,y)=\delta_C(x)+\delta_Q(y)+\frac{1}{2}||Ax-y||^2\}. \tag{2.1}$$

It can be easy to see that a point $x^* \in C$ is a solution of problem (1.1) if and only if it is a minimizer of problem (2.1) and its optimal value is 0. Therefore, we can solve the problem (1.1) by solving the problem (2.1).

Remark 2.2. (1). When D is a nonempty and closed set, the indicator function of D is a lower semicontinuous and proper function [15]. Hence, the objective function $\psi(x, y)$ is a closed and proper function.

(2). Let $f(x, y) = \frac{1}{2}||Ax - y||^2$, then

$$\nabla_x f(x, y) = A^T(Ax - y) \ , \quad \nabla_y f(x, y) = y - Ax.$$

For a given nonempty and closed sets Ω in R^n, the projection from R^n onto Ω is defined by

$$P_\Omega(x) = \text{argmin}\{\|x - y\| : y \in \Omega\}.$$

Since the Ω is not necessarily convex, the projector $P_\Omega(x)$ may be a set-valued operator, as opposed to the convex case where projections are guaranteed to be single-valued.

Definition 2.3 ([1])**.** Let $f : R^n \to (-\infty, +\infty]$ be a proper lower semicontinuous function.

(i) For each $x \in dom\ f$, the Fréchet subdifferential of f at x, written $\widehat{\partial} f(x)$, is the set of vectors $u \in R^n$ which satisfy

$$\liminf_{y \to x, y \neq x} \frac{f(y) - f(x) - \langle u, y - x\rangle}{||y - x||} \geq 0.$$

If $x \notin dom\ f$, then $\widehat{\partial} f(x) = \emptyset$.

(ii) The limiting-subdifferential, or simply the subdifferential for short, of f at $x \in dom f$, written $\partial f(x)$, is defined as follows

$$\partial f(x) = \{u \in R^n : \exists x^k \to x, f(x^k) \to f(x), u^k \in \widehat{\partial} f(x^k) \to u\}.$$

Remark 2.4 ([1])**.** (1) Let $(x^k, u^k)_{k \in N}$ be a sequence that converges to (x, u) as $k \to \infty$. By the definition of $\partial f(x)$, if $f(x^k)$ converges to $f(x)$ as $k \to \infty$, then $(x, u) \in Graph\ \partial f := \{(x, u) \in R^n \times R^m :\ u \in \partial f(x)\}$.

(2) A necessary (but not sufficient) condition for $x \in R^n$ to be a minimizer of f is

$$0 \in \partial\ f(x). \tag{2.2}$$

A point that satisfies (2.2) is called limiting-critical or simply critical. The set of critical points of f is denoted by $crit\ f$.

Proposition 2.5 ([5])**.** *Assume that the coupling function* $H(x, y)$ *in problem* $\min\{\psi(x, y) = f(x) + H(x, y) + g(y)\}$ *is continuously differentiable. Then for all* $(x, y) \in R^n \times R^m$*, we have*

$$\partial\psi(x, y) = (\partial f(x) + \nabla_x H(x, y), \partial g(y) + \nabla_y H(x, y)) = (\partial_x \psi(x, y), \partial_y \psi(x, y)).$$

Definition 2.6 ([1])**.** Let C be a nonempty closed subset of R^n.

(i) For any $x \in C$ the Fréchet normal cone to C at x is defined by

$$\widehat{N}_C(x) = \{v \in R^n : \langle v, y - x\rangle \leq o(x - y), y \in C\}.$$

(ii) The (limiting) normal cone to C at x is denoted $N_C(x)$ and is defined by

$$v \in N_C(x) \Leftrightarrow \exists x_k \in C, x_k \to x, \exists v_k \in \widehat{N}_C(x_k), v_k \to v.$$

When $x \notin C$, we set $N_C(x) = \emptyset$.

Remark 2.7 ([1]). An elementary but important fact about normal cone and subdifferential is the following

$$\partial\delta_C = N_C.$$

Definition 2.8 ([9]). Let $\eta \in (0,+\infty)$. We denote by Φ_η the class of all concave and continuous functions $\varphi : [0,\eta) \to R_+$ which satisfy the following conditions

(i) $\varphi(0) = 0$;

(ii) φ is continuously differentiable on $(0,\eta)$ with $\varphi' > 0$;

Definition 2.9 ([3]). Let $f : R^n \to R \cup \{+\infty\}$ be proper and lower semicontinuous.

(i) The function f is said to have the Kurdyka-Lojasiewicz property at $\overline{u} \in dom\ \partial f$ if there exist $\eta \in (0,+\infty]$, a neighborhood U of $\overline{u}$ and a function $\varphi \in \Phi_\eta$ such that for all

$$u \in U \cap [f(\overline{u}) < f < f(\overline{u}) + \eta],$$

the following inequality holds

$$\varphi'(f(u) - f(\overline{u}))dist(0, \partial f(u)) \geq 1.$$

(ii) If f satisfies the KL property at each point of $dom\ \partial f$ then f is called KL function.

Lemma 2.10 ([3]). *Let Ω be a compact set and let $f : R^n \to R \cup \{+\infty\}$ be a proper and low semicontinuous function. Assume that f is constant on Ω and satisfies the KL property at each point of Ω. Then, there exist $\varepsilon > 0, \eta > 0$, and $\varphi \in \Phi_\eta$ such that for all $\overline{u}$ in Ω and all u in the following intersection:*

$$\{u \in R^n \ : \ dist(u,\Omega) < \varepsilon\} \cap \{u \in R^n \ : \ f(\overline{u}) < f(u) < f(\overline{u}) + \eta\}, \tag{2.3}$$

one has,

$$\varphi'(f(u) - f(\overline{u}))dist(0, \partial f(u)) \geq 1.$$

Definition 2.11 ([2]). (i) A subset S of R^n is a real semialgebraic if there exists a finite number of real polynomial functions $P_{ij}, Q_{ij} : R^n \to R$ such that

$$S = \bigcup_{j=1}^{p} \bigcap_{i=1}^{q} \{x \in R^n : P_{ij} = 0, Q_{ij} < 0\};$$

(ii) A function $f : R^n \to R \cup \{+\infty\}$ is called semi-algebraic if its graph $\{(x,\lambda) \in R^{n+1} : f(x) = \lambda\}$ is a semialgebraic subset of R^{n+1}.

Remark 2.12. (1) ([1,3,4]) If a function is semialgebraic, it satisfies the KL property with $\varphi(s) = cs^{1-\theta}$, for some $\theta \in [0,1) \cap \mathbb{Q}$, and some $c > 0$.

(2) ([1]) The indicator function of semialgebraic sets are semialgebraic.

3 A Projection Algorithm and Convergence Analysis

We now formally state our projection algorithm with an Armijo-like search.

Algorithm 3.1.

Step 0 Given constants $l \in (0,1)$, $\mu \in (0, \frac{1}{2})$, $\varepsilon > 0$. Let $x^0 \in C$. Set $k = 0$.

Step 1 Compute

$$y^{k+1} \in P_Q(Ax^k),$$

$$x^{k+1} \in P_C(x^k - \alpha_k A^T(Ax^k - y^{k+1})),$$

where $\alpha_k = l^{m_k}$, m_k is the smallest nonnegative integer m such that

$$||\nabla_x f(x^k, y^{k+1}) - \nabla_x f(x^{k+1}, y^{k+1})|| \le \frac{\mu||x^k - x^{k+1}||}{\alpha_k}.$$

Step 2 If $||Ax^{k+1} - y^{k+1}|| = 0$, stop. Otherwise let $k = k + 1$. Go to **Step 1**

Remark 3.1. Algorithm 3.1 can be seen as the CQ algorithm which solves the convex split feasibility problem in the non-convex setting. Also, in Algorithm 3.1, we can obtain clearly that if $||Ax^{k+1} - y^{k+1}|| = 0$, then x^{k+1} is the optimal solution of the problem (1.1).

Theorem 3.2. *For all $k = 0, 1, \ldots$, the Armijo-like search rule is well defined.*

Proof. When $\alpha_k > 0$ is sufficiently small, $\alpha_k \le \frac{\mu}{||A^T A||}$, then $||A^T A|| \le \frac{\mu}{\alpha_k}$. Therefore,

$$\begin{aligned}||\nabla_x f(x^k, y^{k+1}) - \nabla_x f(x^{k+1}, y^{k+1})|| &= ||A^T A(x^k - x^{k+1})|| \\ &\le ||A^T A||||x^k - x^{k+1}|| \\ &\le \frac{\mu||x^k - x^{k+1}||}{\alpha_k}.\end{aligned}$$

The proof is completed. □

Remark 3.3. Remark 3.1 and Theorem 3.2 illustrate that Algorithm 3.1 is practicable.

Lemma 3.4.

$$\alpha_k \ge \min\{1, \frac{\mu l}{||A^T A||}\}, \qquad k = 0, 1, \ldots.$$

Proof. If $\alpha_k = 1$, then the lemma is proved.

If $\alpha_k < 1$, from the search rule, we know that exist $\exists\ \widehat{x}^k \in P_C(x^k - \frac{\alpha_k}{l}\nabla_x f(x^k, y^{k+1}))$ such that

$$||\nabla_x f(x^k, y^{k+1}) - \nabla_x f(\widehat{x}^k, y^{k+1})|| > \frac{\mu l||x^k - \widehat{x}^k||}{\alpha_k}.$$

Then, we can get the following inequality

$$||A^T A|| > \frac{\mu l}{\alpha_k},$$

namely,

$$\alpha_k > \frac{\mu l}{||A^T A||}.$$

This completes the proof. □

Definition 3.5 ([15])**.** (Gradient-like Descent Sequence). Let $F : R^d \times R^p \to (-\infty, +\infty]$ be a proper and lower semicontinuous function which is bounded from below and $\{(u^k, v^k)\}_{k\in N}$ be a sequence generated by a certain algorithm for solving the problem

$$inf\{F(u, v) : u \in R^d, v \in R^p\}.$$

We say that $\{(u^k, v^k)\}_{k\in N}$ is a gradient-like descent sequence for minimizing F if the following three conditions hold:

(C1) (Sufficient decrease property). There exists a positive scalar ρ_1, such that $\rho_1||u^{k+1}-u^k||^2 \le F(u^k, v^k) - F(u^{k+1}, v^{k+1})$;

(C2) (A subgradient lower bound for the iterates gap). There exists a positive scalar ρ_2, such that $||w^{k+1}|| \le \rho_2||u^{k+1} - u^k||$;

(C3) Let $(\overline{u}, \overline{v})$ be a limit point of a subsequence $\{(u^k, v^k)\}_{k\in K}$, then $\limsup_{k\in K\subseteq N} F(u^k, v^k) \le F(\overline{u}, \overline{v})$.

Theorem 3.6 ([15])**.** *Let $\{(u^k, v^k)\}_{k\in N}$ be a bounded gradient-like descent sequence for minimizing F. If F is semialgebraic, then the sequence $\{u^k\}_{k\in N}$ converges to some u^*. In addition, for any limit point v^* of $\{v^k\}_{k\in N}$ we have that $(u^*, v^*) \in crit\ F$.*

Now we establish the convergence of Algorithm 3.1.

Theorem 3.7. *Let $\{(x^k, y^k)\}_{k\in N}$ be a bounded sequence generated by Algorithm 3.1. If C, Q are semialgebraic sets, then the sequence $\{x^k\}_{k\in N}$ converges globally to some point x^* such that for any limit point y^* of the sequence $\{y^k\}_{k\in N}$, the pair $(x^*, y^*) \in crit\ \psi$.*

Proof. We only prove that the sequence $\{(x^k, y^k)\}_{k\in N}$ is a gradient-like descent sequence according to Theorem 3.6 and Remark 2.12. We begin with the first condition (C1). From the definition of projection, we have that

$$x^{k+1} \in \arg\min_x \{\delta_C(x) + \alpha_k\langle x - x^k, \nabla_x f(x^k, y^{k+1})\rangle + \frac{1}{2}||x - x^k||^2\}. \tag{3.1}$$

This implies

$$\delta_C(x^{k+1}) + \alpha_k\langle x^{k+1} - x^k, \nabla_x f(x^k, y^{k+1})\rangle + \frac{1}{2}||x^{k+1} - x^k||^2 \le \delta_C(x^k).$$

Therefore, we have

$$\delta_C(x^{k+1}) + \frac{1}{2}||x^{k+1} - x^k||^2 \le \delta_C(x^k) + \alpha_k\langle x^k - x^{k+1}, \nabla_x f(x^k, y^{k+1})\rangle.$$

By subtracting $\alpha_k\langle x^k - x^{k+1}, \nabla_x f(x^{k+1}, y^{k+1})\rangle$ on both side of the above inequality, we arrive at

$$\begin{aligned}
&\delta_C(x^{k+1}) + \tfrac{1}{2}||x^{k+1} - x^k||^2 - \alpha_k\langle x^k - x^{k+1}, \nabla_x f(x^{k+1}, y^{k+1})\rangle \\
&\le \delta_C(x^k) + \alpha_k\langle x^k - x^{k+1}, \nabla_x f(x^k, y^{k+1}) - \nabla_x f(x^{k+1}, y^{k+1})\rangle \\
&\le \delta_C(x^k) + \alpha_k||x^k - x^{k+1}||||\nabla_x f(x^k, y^{k+1}) - \nabla_x f(x^{k+1}, y^{k+1})|| \\
&\le \delta_C(x^k) + \mu||x^k - x^{k+1}||^2.
\end{aligned}$$

Also, we can get when y is fixed, $f(x, y)$ is a convex function of x. From the search rule, we obtain that $\alpha_k \le 1$. Thus,

$$\begin{aligned}
&\delta_C(x^{k+1}) + (\tfrac{1}{2} - \mu)||x^{k+1} - x^k||^2 \\
&\le \delta_C(x^k) + \alpha_k\langle x^k - x^{k+1}\ ,\ \nabla_x f(x^{k+1}, y^{k+1})\rangle \\
&\le \delta_C(x^k) + f(x^k, y^{k+1}) - f(x^{k+1}, y^{k+1}),
\end{aligned}$$

namely,

$$\delta_C(x^{k+1}) + (\frac{1}{2} - \mu)||x^{k+1} - x^k||^2 + f(x^{k+1}, y^{k+1}) \le \delta_C(x^k) + f(x^k, y^{k+1}).$$

Moreover, in a similar way, we have

$$\delta_Q(y^{k+1}) + f(x^k, y^{k+1}) \leq \delta_Q(y^k) + f(x^k, y^k).$$

Combining the last two inequalities yields the desired result with $\rho_1 = \frac{1}{2} - \mu$.

For proving condition (C2) we use the optimality condition associated with the (3.1), meaning

$$(x^k - x^{k+1}) - \alpha_k \nabla_x f(x^k, y^{k+1}) \in \partial\delta_C(x^{k+1}).$$

Multiplying both sides by $\frac{1}{\alpha_k}$ and from Remark 2.7, we obtain that

$$\frac{1}{\alpha_k}(x^k - x^{k+1}) - \nabla_x f(x^k, y^{k+1}) \in \partial\delta_C(x^{k+1}).$$

By using that $\nabla_x f(x^k, y^{k+1}) = A^T(Ax^k - y^{k+1})$ and adding $A^T Ax^{k+1}$ on both side, we arrive at

$$\frac{1}{\alpha_k}(x^k - x^{k+1}) - A^T Ax^k + A^T Ax^{k+1} \in \partial\delta_C(x^{k+1}) + A^T Ax^{k+1} - A^T y^{k+1}.$$

Also, we can have the following inequality via applying the same method to y,

$$Ax^k \in \delta_Q(y^{k+1}) + y^{k+1}.$$

By subtracting both sides by Ax^{k+1}, we get that

$$A(x^k - x^{k+1}) \in \partial\delta_Q(y^{k+1}) - Ax^{k+1} + y^{k+1}.$$

Letting

$$A_x^{k+1} = \frac{1}{\alpha_k}(x^k - x^{k+1}) + A^T A(x^{k+1} - x^k), A_y^{k+1} = A(x^k - x^{k+1}),$$

from Proposition 2.5, we obtain that $(A_x^{k+1}, A_y^{k+1}) \in \partial\psi(x^{k+1}, y^{k+1})$. Therefore,

$$||A_x^{k+1}|| \leq \frac{1}{\alpha_k}||x^{k+1} - x^k|| + ||A^T A||||x^{k+1} - x^k|| \leq (\frac{||A^T A||}{\mu l} + ||A^T A||)||x^{k+1} - x^k||,$$

$$||A_y^{k+1}|| \leq ||A||||x^{k+1} - x^k||.$$

Letting $w^{k+1} = (A_x^{k+1}, A_y^{k+1})$. And combining the last two inequalities yields that

$$||w^{k+1}|| \leq ||A_x^{k+1}|| + ||A_y^{k+1}|| \leq (\frac{||A^T A||}{\mu l} + ||A^T A|| + ||A||)||x^{k+1} - x^k||.$$

Letting $\rho_2 = (\frac{1}{\mu l} + 1)||A^T A|| + ||A||$, we can obtain the desired result.

Lastly, from the structural properties of objective function and **Step 1** in Algorithm 3.1, we deduce that $\psi(x^k, y^k) = \frac{1}{2}||Ax^k - y^k||^2$ is continuous. Therefore, condition (C3) easily follows. This completes the proof. □

From Algorithm 3.1, it may be seen that the sequence $\{(x^k, y^k)\}_{k\in N}$ generated by Algorithm 3.1 only has the global convergence for the variable $\{x^k\}_{k\in N}$. Thus, we consider adding the necessary line search to y in order to obtain the global convergence of the whole sequence.

Algorithm 3.2.

Step 0 Given constants $l \in (0,1)$, $\mu \in (0, \frac{1}{2})$, $\varepsilon > 0$. Let $x^0 \in C$, $y^0 \in Q$. Set $k = 0$.
Step 1 Compute

$$x^{k+1} \in P_C(x^k - \alpha_k \nabla_x f(x^k, y^k)),$$

where $\alpha_k = l^{m_k}$, m_k is the smallest nonnegative integer m such that

$$||\nabla_x f(x^k, y^k) - \nabla_x f(x^{k+1}, y^k)|| \leq \frac{\mu||x^k - x^{k+1}||}{\alpha_k}.$$

Step 2 Compute

$$y^{k+1} \in P_Q(y^k - \beta_k \nabla_y f(x^{k+1}, y^k)),$$

where $\beta_k = l^{n_k}$, n_k is the smallest nonnegative integer n such that

$$||\nabla_y f(x^{k+1}, y^k) - \nabla_y f(x^{k+1}, y^{k+1})|| \leq \frac{\mu||y^k - y^{k+1}||}{\beta_k}.$$

Step 3 If $||Ax^{k+1} - y^{k+1}|| = 0$, stop. Otherwise let $k = k + 1$. Go to **Step 1**

By the similar proof process with Theorem 3.2 and Lemma 3.4, we can obtain that the search rule in **Step 2** is well defined and

$$\beta_k > \mu l, \qquad k = 0, 1, \ldots.$$

Theorem 3.8. *Let $\{z^k\}_{k\in N} = \{(x^k, y^k)\}_{k\in N}$ be a sequence generated by Algorithm 3.2 which is assumed to be bounded. Suppose that C, Q are semialgebraic sets. The following assertions hold:*
(i) $\sum_{k=1}^{\infty} ||z^{k+1} - z^k|| < \infty$.
(ii) *The sequence $\{z^k\}_{k\in N}$ converges to a critical point $z^* = (x^*, y^*) \in crit\ \psi$.*

Proof. By using the method that is similar to Theorem 3.7, we obtain that

$$\psi(z^k) - \psi(z^{k+1}) \geq \rho_1 ||z^{k+1} - z^k||^2, \text{with } \rho_1 = \frac{1}{2} - \mu. \tag{3.2}$$

$$||w^k|| \leq \rho_2 ||z^k - z^{k-1}||, \text{with } \rho_2 = \sqrt{2} max\{\frac{||A||^2}{\mu l} + ||A||^2, \frac{1}{\mu l} + 1 + ||A||\}. \tag{3.3}$$

(i) Since $\{z^k\}_{k\in N}$ is bounded there exists a subsequence $\{z^{k_j}\} \to \overline{z}, k_j \to \infty$, we get

$$\lim_{k_j \to \infty} \psi(z^{k_j}) = \psi(\overline{z}).$$

From (3.2), we can get that $\lim_{k\to\infty} \psi(z^k)$ exists. So

$$\lim_{k\to\infty} \psi(z^k) = \psi(\overline{z}). \tag{3.4}$$

If there exists an integer $\overline{k}$ for which $\psi(z^{\overline{k}}) = \psi(\overline{z})$, then the decreasing property would imply that $z^{\overline{k}+1} = z^{\overline{k}}$. Thus, the conclusion is obvious.

If $\psi(\overline{z^k}) \neq \psi(\overline{z})$ always holds for each $k = 0, 1, 2, \dots$, then, because $\{\psi(z^k)\}_{k\in N}$ is a nonincreasing sequence, it is clear from (3.4) that

$$\psi(\overline{z}) < \psi(z^k).$$

Again from (3.4), for any $\eta > 0$ there exists k_0 such that

$$\psi(z^k) < \psi(\overline{z}) + \eta, \quad \forall k > k_0.$$

We denote by $\omega(z^0)$ the set of all limit points. This implies

$$\lim_{k\to+\infty} dist(z^k, \omega(z^0)) = 0.$$

Therefore, for any $\varepsilon > 0$ there exists k_1 such that

$$dist(z^k, \omega(z^0)) < \varepsilon, \quad \forall k > k_1.$$

Combining the above two formulas, we find that $\{z^k\}_{k\in N}$ satisfies (2.3) for all $k > l = max\{k_0, k_1\}$.

Moreover, since $\omega(z^0)$ is nonempty and compact, and $\psi(z^k)$ is finite and constant on $\omega(z^0)$, we can apply Lemma 2.10 with $\Omega = \omega(z^0)$. Therefore, for any $k > l$, we have

$$\varphi'(\psi(z^k) - \psi(\overline{z}))dist(0\ ,\ \partial\psi(z^k)) \geq 1.$$

Using (3.3), we get that

$$\varphi'(\psi(z^k) - \psi(\overline{z})) \geq \frac{1}{\rho_2||z^k - z^{k-1}||}.$$

On the other hand, from the concavity of φ, we get that

$$\varphi(\psi(z^k) - \psi(\overline{z})) - \varphi(\psi(z^{k+1}) - \psi(\overline{z})) \geq \varphi'(\psi(z^k) - \psi(\overline{z}))(\psi(z^k) - \psi(z^{k+1})).$$

For convenience, for all $p, q \in N$ and $\overline{z}$, we define

$$\triangle_{p,q} = \varphi(\psi(z^p) - \psi(\overline{z})) - \varphi(\psi(z^q) - \psi(\overline{z})).$$

Combining the above two inequalities with (3.2) yields for any $k > l$ that

$$\triangle_{k,k+1} \geq \varphi'(\psi(z^k) - \psi(\overline{z}))(\psi(z^k) - \psi(z^{k+1})) \geq \frac{\rho_1||z^{k+1} - z^k||^2}{\rho_2||z^k - z^{k-1}||},$$

and hence

$$||z^{k+1} - z^k|| \leq \sqrt{c\triangle_{k,k+1}||z^k - z^{k-1}||}$$

with $c = \frac{\rho_2}{\rho_1}$. Using the fact that $2\sqrt{ab} \leq a + b$ for all $a, b \geq 0$, we infer

$$2||z^{k+1} - z^k|| \leq ||z^k - z^{k-1}|| + c\triangle_{k,k+1}. \tag{3.5}$$

Summing up (3.5) for $i = l + 1, \dots, k$ yields

$$\begin{aligned} 2\sum_{i=l+1}^{k}||z^{i+1} - z^i|| &\leq \sum_{i=l+1}^{k}||z^i - z^{i-1}|| + c\sum_{i=l+1}^{k}\triangle_{i,i+1} \\ &= \sum_{i=l+1}^{k}||z^{i+1} - z^i|| + ||z^{l+1} - z^l|| + c\sum_{i=l+1}^{k}\triangle_{i,i+1} \\ &= \sum_{i=l+1}^{k}||z^{i+1} - z^i|| + ||z^{l+1} - z^l|| + c\triangle_{l+1,k+1}, \end{aligned}$$

where the last equality follows from the fact that $\triangle_{p,q} + \triangle_{q,r} = \triangle_{p,r}$ for all $p, q, r \in N$. Since $\varphi \geq 0$, we thus have for all $k > l$ that

$$\begin{aligned}\sum_{i=l+1}^{k} \|z^{i+1} - z^i\| &\leq \|z^{l+1} - z^l\| + c(\varphi(\psi(z^{l+1}) - \psi(\overline{z})) - \varphi(\psi(z^{k+1}) - \psi(\overline{z}))) \\ &\leq \|z^{l+1} - z^l\| + c\varphi(\psi(z^{l+1}) - \psi(\overline{z})) \\ &\leq \|z^{l+1} - z^l\| + c\varphi(\psi(z^{l}) - \psi(\overline{z})).\end{aligned}$$

Since the right hand-side of the inequality above does not depend on k at all, it is easily shows that $\sum_{k=1}^{\infty} \|z^{k+1} - z^k\| < +\infty$.

(ii) From (i), we know that $\{z^k\}_{k\in N}$ is a Cauchy sequence. So we can get that $\{z^k\}_{k\in N} \to \overline{z} = z^*$, and $\omega^k \to 0$. We finally conclude from (3.4) and Remark 2.4 (1) that $0 \in \partial\psi(z^*)$. Hence, z^* is a critical point of φ. The proof is completed. □

References

[1] H. Attouch, J. Bolte, P. Redont and A. Soubeyran, Proximal alternating minimization and projection methods for nonconvex problems: an approach based on the Kurdyka-Lojasiewicz inequality, *Math. Oper. Res.* 35 (2010) 438–457.

[2] H. Attouch, J. Bolte and B. F. Svaiter, Convergence of descent methods for semi-algebraic and tame problems: proximal algorithms, forward-backwards plitting, and regularized Gauss-Seidel methods, *Math. Program.* 137 (2013) 91–129.

[3] J. Bolte, A. Daniilidis and A. Lewis, The Lojasiewicz inequality for nonsmooth subanalytic functions with applications to subgradient dynamical system, *SIAM J. Optim.* 17 (2006) 1205–1223.

[4] J. Bolte, A. Daniilidis, A. Lewis and M. Shiota, Clarke subgradients of stratifiable functions, *SIAM J. Optim.* 18 (2007) 556–572.

[5] J. Bolte, S. Sabach and M Teboulle, Proximal alternating linearized minimization for nonconvex and nonsmooth problems, *Math. Program.* 146 (2014) 459–494.

[6] C. Byrne, Iterative oblique projection onto convex sets and the split feasibility problem, *Inverse Problems* 18 (2002) 441–453.

[7] C. Byrne, A unified treatment of some iterative algorithms in signal processing and image reconstruction, *Inverse Problems* 20 (2004) 103–120.

[8] Y. Censor, T. Elfving, A multiprojection algorithm using Bregman projections in a product space, *Numer. Algorithms* 8 (1994) 221–239.

[9] C. Chen, T. Pong, L. Tan and L. Zeng, A difference-of-convex approach for split feasibility with applications to matrix factorizations and outlier detection, *J. Global Optim.* 78 (2020) 107–136.

[10] Y. Z. Dang, F. W. Meng, J. Sun, Convergence of an algorithm for the split common fixed-point of asymptomatic quasi-nonexpansive operators, *Pac. J. Optim.* 10 (2014) 453–460.

[11] Y. Z. Dang, F. W. Meng, J. Sun, Convergence analysis of a parallel projection algorithm for solving convex feasibility problems, *Numer. Algebra Control Optim.* 6 (2016) 505–519.

[12] Y. Z. Dang, F. W. Meng, J. Sun, An iterative algorithm for split common fixed-point problem for demicontractive mappings, in: *Optim. Methods, Theory and Appl.*, H. Xu, S. Wang, S. Wu (eds), Springer, 2015 pp 85–94.

[13] Y. Z. Dang, J. Sun, H. L. Xu, Inertial accelerated algorithms for solving split feasibility problems, *J. Ind. Manag. Optim.* 13 (2017) 1383–1394.

[14] Y. Z. Dang, B. Rodrigues, J. Sun, Strong convergence of a KM iterative algorithm for computing a split common fixed-point of quasi-nonexpansive operators, *J. Nonlinear Convex Anal.* 22 (2021) 969–978.

[15] A. Gibali, S. Sabach and S. Voldman, Non-convex split feasibility problems: models, algorithms and theory, *Open J. Math. Optim.* 1 (2020) 1–15.

[16] G. López, V. Martín-Márquez, F. Wang and H. Xu, Solving the split feasibility problem without prior knowledge of matrix norms, *Inverse Problems* 28 (2012) 374–389.

[17] L. F. Polania, R. E. Carrillo, M. Blanco-Velasco and K. E. Barner, Compressive sensing for ECG signals in the presence of electromyographic noise, in: *38th Annual Northeast Bioengineering Conference*, 2012, pp. 295–296.

[18] B. Qu, N. H. Xiu, A note on the CQ algorithm for the split feasibility problem, *Inverse Problems* 21 (2005) 1655–1665.

[19] R. T. Rockafellar and R.J-B. Wets, *Variational Analysis*, Springer-Verlag, Berlin Heidelberg, 1998.

[20] R. Tovey, J. Liang, The fun is finite: Douglas-Rachford and sudoku puzzle-finite termination and local linear convergence, (2020) doi: 10.48550/arXiv.2009.04018.

[21] J. Xu, E. C. Chi, M. Yang and K. Lange, A majorization minimization algorithm for split feasibility problems, *Comput. Optim. Appl.* 71 (2018) 795–828.

[22] Q. Yang, The relaxed CQ algorithm solving the split feasibility problem, *Inverse Problems* 20 (2004) 1261–1266.

Manuscript received 22 September 2022
revised 16 February 2023
accepted for publication 20 February 2023

Yang Yang
Institute of Operations Research
Qufu Normal University, Shandong, 276826, China
E-mail address: yang__y123@163.com

Biao Qu
Institute of Operations Research
Qufu Normal University, Shandong, 276826, China
E-mail address: qubiao001@163.com

Ming Qiu
Institute of Operations Research
Qufu Normal University, Shandong, 276826, China
E-mail address: peace_halo@163.com