A new halfspace-relaxation projection method for the split feasibility problem

Biao Qu a,b, *, Naihua Xiu a

a Department of Applied Mathematics, Beijing Jiaotong University, Beijing 100044, People’s Republic of China
b Institute of Operations Research, Qufu Normal University, Shandong 276826, People’s Republic of China

Received 29 May 2006; accepted 3 March 2007
Available online 24 March 2007
Submitted by Y. Censor

Abstract

Let C and Q be nonempty closed convex sets in ℜn and ℜm, respectively, and A an m × n real matrix. The problem, to find \( x \in C \) with \( Ax \in Q \) if such \( x \) exist, is called the split feasibility problem (SFP). This problem is important in intensity-modulated radiation therapy, signal processing, image reconstruction and so on. In this paper, based on a new reformulation for the SFP, we propose a new halfspace-relaxation projection method for the SFP. The method is implemented very easily and is proven to be fully convergent to the solution for the case where the solution set of the SFP is nonempty. Preliminary computational experience is also reported.

Keywords: The split feasibility problem; Halfspace-relaxation projection method; Fully convergent

1. Introduction

Intensity-modulated radiation therapy (IMRT) is an important recent advance in radiation therapy. In IMRT, beams of penetrating radiation are directed at the tumor lesion from...
external sources. A multileaf collimator is used to split each beam into many beamlets with individually controllable intensities. There are two principal aspects of radiation teletherapy that call for computational modeling. The first is the calculation of the radiation dose absorbed in the irradiated tissue based on a given distribution of beamlet intensities. The second aspect is to find a distribution of radiation intensities deliverable by all beamlets, which is the inverse problem of the first. To handle dose constraints and radiation source constraints in IMRT, Censor et al. [6] recently proposed and studied a unified mathematical model, the multiset split feasibility problem that can be regarded as an extension of the split feasibility problem.

The split feasibility problem (SFP) is to find \( x \in C \) with \( Ax \in Q \) if such \( x \) exist, where \( C \) and \( Q \) are nonempty closed convex sets in \( \Re^n \) and \( \Re^m \), respectively, and \( A \) is an \( m \times n \) real matrix. This problem was first presented and analyzed by Censor and Elfving [5], and also appears in signal processing, image reconstruction and so on. Many well-known iterative algorithms for solving it were established; see the survey papers [1,4]. The focus of this paper is to propose a new method for solving the SFP, and hence for addressing the IMRT.

In [5], the authors used their multidistance idea to obtain iterative algorithms for solving the SFP. Their algorithms, as well as others obtained later (see, e.g., [2]) involve matrix inverses at each iteration. In [3], Byrne presented a projection method called the CQ algorithm for solving the SFP that does not involve matrix inverses. In the CQ algorithm, one needs to calculate the orthogonal projections onto \( C \) and \( Q \), denoted by \( P_C \) and \( P_Q \), respectively, which may be impossible or needs much work to exactly compute in some cases. In [7], Yang presented a relaxed CQ algorithm for solving the SFP, where he used two halfspaces \( C_k \) and \( Q_k \) in place of \( C \) and \( Q \), respectively, at the \( k \)th iteration and the orthogonal projections onto \( C_k \) and \( Q_k \) are easily executed. Noting that both the CQ algorithm and the relaxed CQ algorithm used a fixed stepsize related to the largest eigenvalue of the matrix \( A^T A \), which sometimes affect convergence of the algorithms, the authors in [12] presented a modification of the the relaxed CQ algorithm by adopting the Armijo-like search. The modified algorithm needs not to compute the matrix inverses and the largest eigenvalue of the matrix \( A^T A \), and makes a sufficient decrease of the objective function at each iteration.

In fact, the main idea of [7,12] is the use of the halfspace-relaxation projection technique presented by Fukushima [8], that is, constructing a halfspace \( \Omega(x) \) containing the given closed convex set \( \Omega \) and being related to the current iterative point \( x \), and replacing \( P_\Omega \) by \( P_{\Omega(x)} \). From the expressions of \( \Omega(x) \), the projection \( P_{\Omega(x)} \) can be very easily computed.

In this paper, based on a new reformulation for the SFP, we propose a new halfspace-relaxation projection method for the SFP. The new approach has the following features: (1) The objective function in the new reformulation is convexly quadratic. The corresponding gradient and Hessian matrix can be computed very easily. (2) At each iteration, it needs to only compute the projection onto a halfspace containing the given closed convex set and being related to the current iterative point, which is implemented very easily. (3) It has full convergence (i.e., the whole iterate sequence is contract and convergent) to the solution for the case where the solution set of the SFP is nonempty. In addition, in Section 3 we show some interesting results on the relaxed projection for the concerned problem.

The rest of this paper is organized as follows. Section 2 gives a new reformulation for the SFP and discusses its properties. Section 3 presents some results on the relaxed projection. Section 4 gives a new halfspace-relaxation projection method for the SFP and shows its convergence. Finally, the preliminary computational experience is reported in Section 5.
2. A new reformulation for the SFP

In this section, we give a new reformulation for the SFP and discuss its properties. It is easy to see that the SFP is equivalent to the following problem:

Find \((x, y) \in \mathbb{R}^n \times \mathbb{R}^m\) with \(y = Ax, y \in Q\) and \(x \in C\) if such \((x, y)\) exist.

Let us define the function \(f : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}\) by

\[
f(x, y) = \frac{1}{2} \|y - Ax\|^2 \tag{1}
\]

and consider the problem

\[
\minimize f(x, y) \text{ subject to } (x, y) \in C \times Q. \tag{2}
\]

The next theorem which is straightforward establishes the relation between the SFP and the optimization problem (2).

**Theorem 2.1.** Let the function \(f : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}\) be defined by (1). If \((x, y) \in C \times Q\) and \(f(x, y) = 0\), then \(x\) solves the SFP.

The function \(f\) defined by (1) has the following properties.

**Proposition 2.1.** Let the function \(f : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}\) be defined by (1). Then it is convexly quadratic. Its gradient and Hessian matrix are given by

\[
\nabla f(x, y) = \left(\begin{array}{c}
-A^T \\
I
\end{array}\right) \left(\begin{array}{c}
-x \\
y
\end{array}\right) = \left(\begin{array}{cc}
A^T & -A^T \\
-A & I
\end{array}\right) \left(\begin{array}{c}
x \\
y
\end{array}\right)
\]

and

\[
\nabla^2 f(x, y) = \left(\begin{array}{c}
-A^T \\
I
\end{array}\right) \left(\begin{array}{c}
-x \\
y
\end{array}\right) = \left(\begin{array}{cc}
A^T & -A^T \\
-A & I
\end{array}\right),
\]

respectively.

**Proof.** We can get the gradient and the Hessian matrix of \(f\) by direct calculation. From the fact that \(\nabla^2 f(x, y)\) is symmetric and positive semidefinite, we know that \(f\) is convex. \(\square\)

Following the necessary and sufficient condition of the global minimum of a convex differentiable function, we get the following proposition.

**Proposition 2.2.** Let the function \(f : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}\) be defined by (1). Then the following two statements are equivalent:

(i) \(f(x^*, y^*) = 0\).

(ii) \(\nabla f(x^*, y^*) = 0\).

Let \(z = \left(\begin{array}{c}
x \\
y
\end{array}\right)\) and \(\Omega = C \times Q\). Then problem (2) becomes

\[
\min f(z) = \frac{1}{2} \|(-A, I)z\|^2 \text{ subject to } z \in \Omega. \tag{2'}
\]
To establish the halfspace-relaxation projection method for the SFP, in this paper, we always assume that following conditions are satisfied:

(H1) The solution set of the SFP is nonempty.

(H2) The set \(\Omega\) is given by
\[
\Omega = \{z \in \mathbb{R}^{n+m} | c(z) \leq 0\},
\]
where \(c : \mathbb{R}^{n+m} \to \mathbb{R}\) is a convex (not necessarily differentiable) function. For any \(z \in \mathbb{R}^{n+m}\), at least one subgradient \(\xi \in \partial c(z)\) can be calculated, where \(\partial c(z)\) is a subgradient of \(c(z)\) at \(z\) and is defined as follows:
\[
\partial c(z) = \{\xi \in \mathbb{R}^{n+m} | c(u) \geq c(z) + \langle \xi, u - z \rangle \text{ for all } u \in \mathbb{R}^{n+m}\}.
\]

**Remark 2.1.** (i) From (H1) and Proposition 2.2, we know that the solution set \(\Omega^*\) of the problem (2') is nonempty and can be expressed as
\[
\Omega^* = \{z \in \Omega | f(z) = 0\}
\]
or
\[
\Omega^* = \{z \in \Omega | \nabla f(z) = 0\}.
\]

(ii) We note that in (H2), the differentiability of \(c\) is not assumed. The representation of \(\Omega\) in (H2) is therefore general enough, because any system of inequalities \(c_j(z) \leq 0, j \in J\), where \(c_j\) are convex and \(J\) is an arbitrary index set, can be reformulated as the single inequality \(c(z) \leq 0\) with \(c(z) = \sup\{c_j(z) | j \in J\}\). Moreover, since \(c\) is assumed to be finite-valued on \(\mathbb{R}^{n+m}\), it is subdifferentiable everywhere and its subdifferentials are uniformly bounded on any bounded subset of \(\mathbb{R}^{n+m}\); see, e.g., [13] for details.

In the rest of this paper, we focus on problem (2') with (H1) and (H2).

### 3. Some results on relaxed projection

In this section, we mainly introduce and characterize halfspace-relaxation projection. First, we review some definitions and basic results which will be used in this paper.

For a given nonempty closed convex set \(X\) in \(\mathbb{R}^N\), the metric projection from \(\mathbb{R}^N\) onto \(X\) is defined by
\[
PX(s) = \arg\min\{\|s - t\| | t \in X\}, \quad s \in \mathbb{R}^N.
\]

It has the following well-known properties.

**Lemma 3.1** [16]. Let \(X\) be a nonempty closed convex subset in \(\mathbb{R}^N\), then for any \(s, t \in \mathbb{R}^N\) and \(q \in X\),
\[
\begin{align*}
(i) \quad \langle PX(s) - s, q - PX(s) \rangle &\geq 0; \\
(ii) \quad \|PX(s) - PX(t)\|^2 &\leq \langle PX(s) - PX(t), s - t \rangle; \\
(iii) \quad \|PX(s) - q\|^2 &\leq \|s - q\|^2 - \|PX(s) - s\|^2.
\end{align*}
\]

**Remark 3.1.** From [16], the result (i) in Lemma 3.1 provides not only a necessary but also a sufficient condition for a vector \(w\) to be the projection of the vector \(s\); i.e., \(w = PX(s)\) if and only if \(w \in X\) and \(\langle w - s, q - w \rangle \geq 0\) hold for any \(q \in X\). Additionally, from part (ii) of Lemma 3.1, we know that \(PX\) is a monotone and nonexpansive (i.e., \(\|PX(s) - PX(t)\| \leq \|s - t\|\)) operator.
Let $F$ be a mapping from $\mathbb{R}^N$ into $\mathbb{R}^N$. For any $s \in \mathbb{R}^N$ and $\alpha > 0$, define

$$s(\alpha) = P_X(s - \alpha F(s)),$$

$$e(s, \alpha) = s - s(\alpha).$$

**Lemma 3.2.** [15, 9]

(a) $\|s - s(\alpha)\|$ is nondecreasing with respect to $\alpha > 0$;
(b) $\frac{|s - s(\alpha)|}{\alpha}$ is nonincreasing with respect to $\alpha > 0$.

From Lemma 3.2, we immediately conclude a useful lemma.

**Lemma 3.3.** Let $F$ be a continuous mapping from $\mathbb{R}^N$ into $\mathbb{R}^N$. For any $s \in \mathbb{R}^N$ and $\alpha > 0$, we have

$$\min\{1, \alpha\}\|e(s, 1)\| \leq \|e(s, \alpha)\| \leq \max\{1, \alpha\}\|e(s, 1)\|.$$

Next, we will give some results on relaxed projection which will be used in the subsequent of this paper.

For any $z \in \mathbb{R}^{n+m}$, define the halfspace

$$\Omega(z) = \{u \in \mathbb{R}^{n+m} | c(z) + \langle \xi(z), u - z \rangle \leq 0\}, \quad (3),$$

where $\xi(z)$ is an element in $\partial c(z)$. Then, by the definition of subgradient, it is clear that $\Omega \subseteq \Omega(z)$ for $z \in \mathbb{R}^{n+m}$.

We call the orthogonal projection onto $\Omega(z)$, denoted by $P_{\Omega(z)}(\cdot)$, the halfspace relaxation projection associated with $\Omega$ at $z$. From the structure feature of $\Omega(z)$, $P_{\Omega(z)}(\cdot)$ can be written as a simple form.

**Proposition 3.1** [11]. For any $v \in \mathbb{R}^{n+m}$,

$$P_{\Omega(z)}(v) = \begin{cases} 
  v - \frac{c(z) + \langle \xi(z), v - z \rangle}{\|\xi(z)\|^2} \xi(z), & \text{if } c(z) + \langle \xi(z), v - z \rangle > 0; \\
  v, & \text{if } c(z) + \langle \xi(z), v - z \rangle \leq 0.
\end{cases}$$

For any $z \in \mathbb{R}^{n+m}$ and $\alpha > 0$, define

$$z_{\Omega(z)}(\alpha) = P_{\Omega(z)}(z - \alpha \nabla f(z)),$$

$$e_{\Omega(z)}(z, \alpha) = z - z_{\Omega(z)}(\alpha),$$

where $\Omega(z)$ is defined as (3). Then from Proposition 3.1, we can get the explicit expressions of $z_{\Omega(z)}(\alpha)$ and $e_{\Omega(z)}(z, \alpha)$, that is,

$$z_{\Omega(z)}(\alpha) = z - \alpha \nabla f(z) - \max \left\{ 0, \frac{c(z) - \alpha \langle \xi(z), \nabla f(z) \rangle}{\|\xi(z)\|^2} \right\} \xi(z),$$

$$e_{\Omega(z)}(z, \alpha) = \alpha \nabla f(z) + \max \left\{ 0, \frac{c(z) - \alpha \langle \xi(z), \nabla f(z) \rangle}{\|\xi(z)\|^2} \right\} \xi(z).$$

Obviously, they can be easily computed.
The following is an interesting result on the relaxed projection.

**Proposition 3.2.** $z$ is a solution of the problem (2) if and only if $e_{\Omega(z)}(z, \alpha) = 0$ holds for any $\alpha > 0$.

**Proof.** “$\Leftarrow$” Suppose that $e_{\Omega(z)}(z, \alpha) = 0$ holds for any $\alpha > 0$, that is, $z = P_{\Omega(z)}(z - \alpha \nabla f(z))$. From Lemma 3.1, we have

$$\alpha \langle \nabla f(z), u - z \rangle \geq 0 \quad \forall u \in \Omega(z).$$

From the above inequality and the fact that $\alpha > 0$ and $\Omega \subseteq \Omega(z)$, we get

$$\langle \nabla f(z), u - z \rangle \geq 0 \quad \forall u \in \Omega.$$

On the other hand, it follows from $z = P_{\Omega(z)}(z - \alpha \nabla f(z))$ that $z \in \Omega(z)$. Taking $u = z$ in the definition of $\Omega(z)$, we obtain that $c(z) \leq 0$, that is, $z \in \Omega$. So, $z$ is a stationary point of the problem (2), i.e., a solution of the concerned problem.

“$\Rightarrow$” Suppose that $z$ is a solution of the problem (2). Then by Proposition 2.2, $z \in \Omega$ and $\nabla f(z) = 0$. Thus $e_{\Omega(z)}(z, \alpha) = 0$ holds for any $\alpha > 0$. □

From Lemma 3.2, we obtain the following proposition.

**Proposition 3.3.** $\frac{\|z - z_{\Omega(z)}(\alpha)\|}{\alpha}$ is nonincreasing with respect to $\alpha > 0$.

**Proposition 3.4.** Let $z \in \mathbb{R}^{n+m}$ be arbitrary. Then for $\mu \in (0, 1)$ and all $\alpha > 0$ sufficiently small (where, $\alpha$ depends on $\mu$), we have

$$\|\nabla f(z) - \nabla f(z_{\Omega(z)}(\alpha))\| \leq \mu \frac{\|z - z_{\Omega(z)}(\alpha)\|}{\alpha}. \quad (4)$$

Furthermore,

$$\langle z - z_{\Omega(z)}(\alpha), \nabla f(z) - \nabla f(z_{\Omega(z)}(\alpha)) \rangle \leq \mu \frac{\|z - z_{\Omega(z)}(\alpha)\|^2}{\alpha}. \quad (5)$$

**Proof.** First we know that $z_{\Omega(z)}(\alpha) = P_{\Omega(z)}(z - \alpha \nabla f(z)) \to P_{\Omega(z)}(z)$ as $\alpha \to 0$.

If there exists an $\alpha_0 > 0$ such that $z = z_{\Omega(z)}(\alpha_0)$, then from Proposition 3.2, we obtain that $z = z_{\Omega(z)}(\alpha)$ holds for all $\alpha > 0$. In this case, this proposition is proved.

If $z \neq z_{\Omega(z)}(\alpha)$ holds for all $\alpha > 0$, we shall prove that (4) holds for all $\alpha > 0$ sufficiently small. Two cases are to be considered.

**Case 1:** $z \notin \Omega(z)$. Then the left-hand of (4) would tend to a positive number while the right-hand of (4) would tend to $+\infty$ as $\alpha \to 0$, implying that this is true.

**Case 2:** $z \in \Omega(z)$. In this case, it is obvious that $z = P_{\Omega(z)}(z)$. Since $\nabla f$ is continuous and $z_{\Omega(z)}(\alpha) \to P_{\Omega(z)}(z) = z$ as $\alpha \to 0$, the left-hand side of (4) would tend to zero as $\alpha \to 0$, while the right-hand side of (4) will be not smaller than the positive number $\|z - z_{\Omega(z)}(1)\|$ as $\alpha \to 0$ by Proposition 3.3. So this is also true.

By using (4) and the Cauchy–Schwartz inequality, we immediately obtain (5). □
4. A new algorithm and its convergence

We now formally state our algorithm.

**Algorithm 1.** Given constants \( \gamma \in (0, \infty) \), \( l \in (0, 1) \), \( \theta \in (0, 2) \), \( \rho \in (0, 1) \). Let \( z^0 \in \mathbb{R}^{n+m} \) be arbitrary. For \( k = 0, 1, \ldots, \) let

\[
\bar{z}^k = P_{\Omega_k}(z^k - \alpha_k \nabla f(z^k)),
\]

where

\[
\Omega_k = \Omega(z^k) = \{ z \in \mathbb{R}^{n+m} | c(z^k) + \langle \xi^k, z - z^k \rangle \leq 0 \},
\]

\( \xi^k \) is an element in \( \partial c(z^k) \), and \( \alpha_k = \gamma l^m_k \) and \( m_k \) is the smallest nonnegative integer \( m \) such that

\[
\alpha_k \langle z^k - \bar{z}^k, \nabla f(z^k) - \nabla f(\bar{z}^k) \rangle \leq (1 - \rho) \| z^k - \bar{z}^k \|^2.
\]

(6)

If \( \| z^k - \bar{z}^k \| = 0 \), stop. Otherwise, set

\[
z^{k+1} = z^k - \gamma_k [z^k - \bar{z}^k - \alpha_k (\nabla f(z^k) - \nabla f(\bar{z}^k))],
\]

(7)

where \( \gamma_k \) is given by

\[
\gamma_k = \frac{\theta \rho \| z^k - \bar{z}^k \|^2}{\| z^k - \bar{z}^k - \alpha_k (\nabla f(z^k) - \nabla f(\bar{z}^k)) \|^2}.
\]

(8)

We claim that Algorithm 1 is different from the algorithm of [12]. The main advantage of Algorithm 1 is that \( \nabla f(z) \) is linear and the iterative processing is very simple. Furthermore, Algorithm 1 combines the extragradient-like algorithm in [10,14] and the Fukushima’s halfspace-relaxation technique, and is well defined. Firstly, if \( \| z^k - \bar{z}^k \| \neq 0 \), then \( \| z^k - \bar{z}^k - \alpha_k (\nabla f(z^k) - \nabla f(\bar{z}^k)) \| \neq 0 \), that is, the \( \gamma_k \) in this algorithm is well defined. In fact it can be proved easily by contradiction with (6). Secondly, by (5) in Proposition 3.4, we know that, for all \( k = 0, 1, \ldots \) holds for all \( \alpha_k \) sufficiently small, that is, \( \alpha_k \) is well defined. Thirdly, if \( \| z^k - \bar{z}^k \| = 0 \), then from Proposition 3.2, we get that \( z^k \in \Omega^* \).

Now, we establish the full convergence of Algorithm 1.

**Theorem 4.1.** Let \( \{ z^k \} \) be a sequence generated by Algorithm 1. Then \( \{ z^k \} \) converges to a point \( \bar{z} = \left( \bar{x}, \bar{y} \right) \), which belongs to \( \Omega^* \) and \( \bar{x} \) is a solution of the SFP.

**Proof.** Let \( z^* \) be any element of \( \Omega^* \). For each \( k \in \{ 0, 1, \ldots \} \), we have from (7) that

\[
\| z^{k+1} - z^* \|^2 = \| z^k - z^* - \gamma_k [z^k - \bar{z}^k - \alpha_k (\nabla f(z^k) - \nabla f(\bar{z}^k))] \|^2
\]

\[
= \| z^k - z^* \|^2 - 2 \gamma_k \langle z^k - z^*, z^k - \bar{z}^k - \alpha_k (\nabla f(z^k) - \nabla f(\bar{z}^k)) \rangle
\]

\[
+ \gamma_k^2 \| z^k - \bar{z}^k - \alpha_k (\nabla f(z^k) - \nabla f(\bar{z}^k)) \|^2.
\]

(9)

We bound below the next-to-last term in (9). By the part (i) of Lemma 3.1, we have

\[
\langle z^k - \bar{z}^k + \alpha_k \nabla f(z^k), z^* - \bar{z}^k \rangle \geq 0.
\]

(10)
Using (10) and (6), the monotonicity of $\nabla f(\cdot)$ and the fact that $\nabla f(z^*) = 0$, we have

\[
0 \leq \langle z^* - z^k, \alpha_k \nabla f(z^k) + z^k - z^k \rangle + \alpha_k \langle z^k - z^*, \nabla f(z^*) \rangle \\
= \alpha_k \langle z^* - z^k, \nabla f(z^k) - \nabla f(z^*) \rangle \\
+ \langle z^* - z^k, \alpha_k \nabla f(z^k) - \alpha_k \nabla f(z^k) + z^k - z^k \rangle \\
= \alpha_k \langle z^* - z^k, \nabla f(z^k) - \nabla f(z^*) \rangle \\
+ \langle z^* - z^k + z^k - z^k, \alpha_k \nabla f(z^k) - \alpha_k \nabla f(z^k) + z^k - z^k \rangle \\
= \alpha_k \langle z^* - z^k, \nabla f(z^k) - \nabla f(z^*) \rangle + \langle z^* - z^k, \alpha_k \nabla f(z^k) - \alpha_k \nabla f(z^k) + z^k - z^k \rangle \\
+ \alpha_k \langle z^k - z^k, \nabla f(z^k) - \nabla f(z^k) \rangle - \|z^k - z^k\|^2 \\
\leq \langle z^* - z^k, \alpha_k \nabla f(z^k) - \alpha_k \nabla f(z^k) + z^k - z^k \rangle \\
+ (1 - \theta) \|z^k - z^k\|^2 - \|z^k - z^k\|^2 \\
= \langle z^* - z^k, \alpha_k \nabla f(z^k) - \alpha_k \nabla f(z^k) + z^k - z^k \rangle - \rho \|z^k - z^k\|^2,
\]

that is,

\[
\langle z^* - z^k, \alpha_k \nabla f(z^k) - \alpha_k \nabla f(z^k) + z^k - z^k \rangle \geq \rho \|z^k - z^k\|^2.
\]

Using (11) in (9) yields

\[
\|z^{k+1} - z^*\|^2 \\
\leq \|z^k - z^*\|^2 - 2\gamma_k \rho \|z^k - z^k\|^2 + \gamma_k^2 \|z^k - z^k - \alpha_k (\nabla f(z^k) - \nabla f(z^k))\|^2 \\
= \|z^k - z^*\|^2 - 2\theta \rho^2 \|z^k - z^k - \alpha_k (\nabla f(z^k) - \nabla f(z^k))\|^2 \\
+ \theta^2 \rho^2 \|z^k - z^k - \alpha_k (\nabla f(z^k) - \nabla f(z^k))\|^2 \\
= \|z^k - z^*\|^2 - \theta(2 - \theta) \rho^2 \|z^k - z^k - \alpha_k (\nabla f(z^k) - \nabla f(z^k))\|^2,
\]

which implies that the sequence $\{\|z^k - z^*\|\}$ is monotonically decreasing and hence $\{z^k\}$ is bounded. Consequently we get from (12)

\[
\lim_{k \to \infty} \frac{\|z^k - z^k\|^2}{\|z^k - z^k - \alpha_k (\nabla f(z^k) - \nabla f(z^k))\|} = 0.
\]

Moreover, it is easy to show that that $\{\bar{z}^k\}$ is bounded. In fact,

\[
\|z^k\| = \|P_{\Omega_k}(z^k - \alpha_k \nabla f(z^k))\| \\
= \|P_{\Omega_k}(z^k - \alpha_k \nabla f(z^k)) + z^* - P_{\Omega_k}(z^*)\| \\
\leq \|z^*\| + \|z^k - z^* - \alpha_k \nabla f(z^k)\| \\
\leq \|z^*\| + \|z^k - z^*\| + \alpha_k \|\nabla f(z^k)\|,
\]

which, together with the boundedness of $\{z^k\}$, deduces the desired result. So the sequence $\{\|z^k - \bar{z}^k - \alpha_k (\nabla f(z^k) - \nabla f(z^k))\|\}$ is also bounded. Thus, from (13), we have
\[ \lim_{k \to \infty} \| z^k - \bar{z}^k \| = 0. \] (14)

Assume that \( \bar{z} \) is an accumulation point of \( \{ z^k \} \). Then there exists a subsequence \( \{ z^k \}_{k \in S} \), where 
\( S \subseteq \{ 0, 1, \ldots \} \), such that
\[ \lim_{k \in S, k \to \infty} z^k = \bar{z}. \]

We are ready to show that \( \bar{z} \) is a solution of problem (2).

First we show that \( \bar{z} \in Q(f^k) \).

Since \( \bar{z}^k \in Q(f^k) \), then by the definition of \( Q(f^k) \), we have
\[ c(z^k) + \langle \xi^k, \bar{z}^k - z^k \rangle \leq 0 \quad \forall k = 1, 2, \ldots \]

Passing onto the limit in this inequality and taking into account (14) and part (ii) of Remark 2.1, we obtain that
\[ c(\bar{z}) \leq 0. \]

Hence we conclude \( \bar{z} \in Q \).

Next we need to show \( \langle \nabla f(\bar{z}), z - \bar{z} \rangle \geq 0 \forall z \in Q \). To do so, we first prove that there exists at
least a subsequence \( \{ \| e_k(z^k, 1) \| \} \) \( \in K \) (where \( K \subseteq S \)) such that
\[ \lim_{k \in K, k \to \infty} \| e_k(z^k, 1) \| = 0, \] (15)
where \( e_k(z^k, \alpha) = e_Q(z^k)(z^k, \alpha) = z^k - P_{Q_k}[z^k - \alpha \nabla f(z^k)] \).

Two cases are to be considered.

\textit{Case 1:} \( \inf_{k \in S} \{ \alpha_k \} = \alpha_{\min} > 0 \). Then from Lemma 3.3, we have
\[ \| e_k(z^k, 1) \| \leq \frac{\| z^k - z^k \|}{\min\{1, \alpha_k\}}, \]
which, together with (14), implies that
\[ \lim_{k \in S, k \to \infty} \| e_k(z^k, 1) \| \leq \lim_{k \in S, k \to \infty} \frac{\| z^k - z^k \|}{\min\{1, \alpha_k\}} \leq \lim_{k \in S, k \to \infty} \frac{\| z^k - \bar{z}^k \|}{\min\{1, \alpha_{\min}\}} = 0. \]

\textit{Case 2:} \( \inf_{k \in S} \{ \alpha_k \} = \alpha_{\min} = 0 \). Since \( \alpha_{\min} = 0 \), there must exist a subsequence \( \{ \alpha_k \}_{k \in K} \), where \( K \subseteq S \), such that \( \lim_{k \in K, k \to \infty} \alpha_k = 0 \). Thus, for all \( \alpha_k \) sufficiently small, \( \frac{\alpha_k}{T} \) must violate the search rule (6), that is
\[ \frac{\alpha_k}{T} \left( z^k - z^k \left( \frac{\alpha_k}{T} \right), \nabla f(z^k) - \nabla f \left( z^k \left( \frac{\alpha_k}{T} \right) \right) \right) \geq (1 - \rho) \left\| z^k - z^k \left( \frac{\alpha_k}{T} \right) \right\|^2, \]
where \( z^k \left( \frac{\alpha_k}{T} \right) = P_{Q_k} \left[ z^k - \frac{\alpha_k}{T} \nabla f(z^k) \right] \). Using Cauchy–Schwarz inequality and the above inequality, we have
\[ \left\| \nabla f(z^k) - \nabla f \left( z^k \left( \frac{\alpha_k}{T} \right) \right) \right\| > (1 - \rho) \frac{\| z^k - z^k \left( \frac{\alpha_k}{T} \right) \|}{\alpha_k \frac{\alpha_k}{T}}. \]

Thus, from Lemma 3.3 again, we get
\[ (1 - \rho) \| e_k(z^k, 1) \| \leq (1 - \rho) \frac{\| z^k - z^k \left( \frac{\alpha_k}{T} \right) \|}{\alpha_k \frac{\alpha_k}{T}} < \left\| \nabla f(z^k) - \nabla f \left( z^k \left( \frac{\alpha_k}{T} \right) \right) \right\|, \]
that is,
\[ \| e_k(z^k, 1) \| \leq \frac{1}{1 - \rho} \left\| \nabla f(z^k) - \nabla f \left( z^k \left( \frac{\alpha_k}{T} \right) \right) \right\|. \]
Moreover, we can get that
\[
\|z^k - z^k\left(\frac{\alpha_k}{l}\right)\| = \|z^k - \bar{z}^k + z^k - P_{\Omega_k}\left[z^k - \frac{\alpha_k}{l}\nabla f(z^k)\right]\|
\leq \|z^k - \bar{z}^k\| + \left\|\frac{\alpha_k}{l}\nabla f(z^k) - \alpha_k\nabla f(z^k)\right\|
= \|z^k - \bar{z}^k\| + \left(\frac{1}{l} - 1\right)\alpha_k\|\nabla f(z^k)\|
\rightarrow 0 \quad (k \in K, k \rightarrow \infty),
\]
which reduces that
\[
\lim_{k \in K, k \rightarrow \infty} \left\|\nabla f(z^k) - \nabla f\left(z^k\left(\frac{\alpha_k}{l}\right)\right)\right\| = 0.
\]
Thus, we have
\[
\lim_{k \in K, k \rightarrow \infty} \|e_k(z^k, 1)\| = 0.
\]
Now, we continue to prove the main result.

From part (i) of Lemma 3.1 and the fact that \(z^k - e_k(z^k, 1) = P_{\Omega_k}(z^k - \nabla f(z^k))\), we have, for \(\forall z \in \Omega \subseteq \Omega_k\)
\[
\langle \nabla f(z^k) - e_k(z^k, 1), z - z^k + e_k(z^k, 1)\rangle \geq 0,
\]
that is
\[
\langle \nabla f(z^k), z - z^k\rangle + \langle \nabla f(z^k), e_k(z^k, 1)\rangle - \langle e_k(z^k, 1), z - z^k\rangle - \|e_k(z^k, 1)\|^2 \geq 0.
\]
Letting \(k \rightarrow \infty \quad (k \in K)\), taking into account (15), we deduce that
\[
\langle \nabla f(\tilde{z}), z - \tilde{z}\rangle \geq 0 \quad \forall z \in \Omega.
\]
From the arbitrariness of \(z\), we can conclude that \(\tilde{z}\) is a solution of problem (2).

Thus we may use \(\tilde{z}\) in place of \(z^*\) in (12), and obtain that \(\{\|z^k - \tilde{z}\|\}\) is convergent. Because there is a subsequence \(\{\|z^k - \tilde{z}\|\}_{k \in S}\) converging to 0, then \(z^k \rightarrow \tilde{z}\) as \(k \rightarrow \infty\). From Theorem 2.1, \(\tilde{x}\) is a solution of the SFP. This completes the proof. \(\square\)

5. Numerical results

To give some insight into the behavior of the algorithm presented in this paper, we implemented it in MATLAB to solve the following three examples. We use \(\|z^k - \tilde{z}\| < \epsilon\) as the stopping criteria.

Throughout the computational experiments, the parameters used in Algorithm 1 were set as \(\epsilon = 10^{-10}, \gamma = 1, l = 0.5, \theta = 1, \rho = 0.5\). In the results reported below, all CPU times reported are in seconds. The approximate solution is referred to the last iteration.

**Example 1 (A convex feasibility problem).** Let \(C = \{x \in \mathbb{R}^3 | x_2^2 + x_3^2 - 4 \leq 0\}, \ Q = \{x \in \mathbb{R}^3 | x_3 - 1 - x_1^2 \leq 0\}.\) Find some point \(x\) in \(C \cap Q\).

Obviously this example can be regarded as an SFP for \(A = I\). The test results for Algorithm 1 being applied to Example 1 are listed in Table 1 using different starting points.
Table 1
Results for Example 1

<table>
<thead>
<tr>
<th>Starting points</th>
<th>Number of iterations</th>
<th>CPU (s)</th>
<th>Approximate solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(1, 2, 3, 0, 0, 0)^T$</td>
<td>43</td>
<td>0.0500</td>
<td>$(0.3213, 0.2815, 0.1425)^T$</td>
</tr>
<tr>
<td>$(1, 1, 1, 1, 1)^T$</td>
<td>67</td>
<td>0.0910</td>
<td>$(0.8577, 0.8577, 1.3097)^T$</td>
</tr>
<tr>
<td>$(1, 2, 3, 4, 5, 6)^T$</td>
<td>85</td>
<td>0.1210</td>
<td>$(1.1548, 0.8518, 1.8095)^T$</td>
</tr>
</tbody>
</table>

Table 2
Results for Example 2

<table>
<thead>
<tr>
<th>Starting points</th>
<th>Number of iterations</th>
<th>CPU (s)</th>
<th>Approximate solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(1, 2, 3, 0, 0, 0)^T$</td>
<td>1890</td>
<td>2.7740</td>
<td>$(-0.1203, 0.0285, 0.0582)^T$</td>
</tr>
<tr>
<td>$(1, 1, 1, 1, 1)^T$</td>
<td>2978</td>
<td>4.2860</td>
<td>$(0.8603, -0.1658, -0.5073)^T$</td>
</tr>
<tr>
<td>$(1, 2, 3, 4, 5, 6)^T$</td>
<td>3317</td>
<td>4.8570</td>
<td>$(3.6522, -0.1526, -2.3719)^T$</td>
</tr>
</tbody>
</table>

**Example 2** (A split feasibility problem). Let $A = \begin{pmatrix} 2 & -1 & 3 \\ 4 & 2 & 0 \\ 2 & 5 & 2 \end{pmatrix}$, $C = \{ x \in \mathbb{R}^3 : x_1 + x_2^2 + 2x_3 \leq 0 \}$, $Q = \{ x \in \mathbb{R}^3 : x_1^2 + x_2 - x_3 \leq 0 \}$. Find $x \in C$ with $Ax \in Q$.

The test results for Algorithm 1 being applied to Example 2 are listed in Table 2 using different starting points.

Algorithm 1 can also be used to solve some nonlinear programming problems. The following is a simple example.

**Example 3** (A nonlinear programming problem)

Minimize $f(z) = \sum_{i=1}^{n} z_i^2$

Subject to $\sum_{i \neq j} z_i^2 - z_j - j \leq 0$, $j = 1, 2, \ldots, n$.

This example has a unique solution $(0, \ldots, 0)^T$.

Table 3 lists the results for Algorithm 1 being applied to Example 3 with initial point $z^0 = (1, \ldots, 1)^T$ for different number of dimensions.

Table 3
Results for Example 3

<table>
<thead>
<tr>
<th>$n$ (dimension)</th>
<th>Number of iterations</th>
<th>CPU (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>36</td>
<td>0.0160</td>
</tr>
<tr>
<td>100</td>
<td>38</td>
<td>0.2970</td>
</tr>
<tr>
<td>1000</td>
<td>40</td>
<td>15.5000</td>
</tr>
<tr>
<td>5000</td>
<td>41</td>
<td>416.7340</td>
</tr>
</tbody>
</table>
The numerical experiments tested for the three simple problems are used to demonstrate the viability of the method proposed in this paper. A remarkable characteristic of the algorithm is the computational simplicity, which makes the algorithm to be implemented easily.

6. Concluding remarks

In this paper, a new halfspace-relaxation projection method for the split feasibility problem has been presented. The main advantages of the proposed method are that each iteration consists of the projection onto a halfspace implemented very easily, and that the full convergence has been established for the case where the solution set of the SFP is nonempty. Whether the proposed method can apply to the infeasible case of the SFP or not is a topic deserving further research.

Acknowledgments

We would like to thank Professor Yair Censor for his valuable comments. We would also like to thank two anonymous referees for their helpful suggestions on the earlier version of this paper.

References