A Note on the Finite Convergence of Alternating Projections

Hoa T. Bui · Ryan Loxton · Asghar Moeini

Received: date / Accepted: date

Abstract We establish sufficient conditions for finite convergence of the alternating projection method for finding the closest points in two non-intersecting sets. To the best of our knowledge, these are the first known results on finite convergence for alternating projections. In the special case of a polyhedron and closed half space, our sufficient conditions allow us to quantify how large the distance between these sets needs to be for alternating projects to converge in a single iteration.

Keywords Alternating projections · non-intersecting · proximal normal cone · intrinsic transversality · finite convergence · polyhedrons

Mathematics Subject Classification (2000) 65K10, 49J52, 58E30, 90C30

1 Introduction

Throughout this paper, $X$ is a Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and associated norm $\| \cdot \|$. The method of alternating projections for two nonempty sets $A, B \subset X$, involves iterating the following steps, starting with $x_n \in A$:

$$
\begin{align*}
y_n &\in P_B(x_n), \\
x_{n+1} &\in P_A(y_n).
\end{align*}
$$

Here, $P_B(x_n)$ is the set of all projections of $x_n$ onto $B$ and $P_A(y_n)$ is the set of all projections of $y_n$ onto $A$. The alternating projection method is mainly used for solving feasibility problems, i.e. finding a point in the intersection of a collection of sets. The study of the convergence of this method in the consistent case (i.e. $A \cap B \neq \emptyset$) has a long history that can be traced back to von Neumann (see [11, 15, 16, 18] for historical comments). In particular, for convex settings, Bregman [4] showed that the method always converges, and a linear convergence rate was established by Gubin et al. [14] and Bauschke & Borwein [1]. For general nonconvex settings, Drusvyatskiy et al. [11] proved that linear convergence is assured with a regularity-type property, called intrinsic transversality. In [18], Noll and Rondepierre studied a general setting that allows for nonlinear convergence under more subtle nonlinear regularity assumptions.

For the inconsistent case, when $A \cap B = \emptyset$, the method does not converge to a single point, but under certain conditions it will converge to a pair of points of nearest distance. For example, Cheney and Goldstein showed in [8] that in Euclidean spaces when the two sets are closed and convex, and one of the sets is compact, the method converges and attains the distance between the two sets. Consequently, if two sets are polytopes, then the method converges to a pair of points of nearest distance. Because of its convergence properties, the alternating projection method for inconsistent cases has been widely applied [9, 10, 14]; see also [5, 6] for a review. In this paper, we are interested in the finite convergence of the method for inconsistent cases.
A key step in our approach is to extend the concept of intrinsic transversality, first defined for consistent cases in [11], to the more general setting when the intersection can be empty or nonempty (Conditions 1, 2, 3 in Section 3). We show that under these conditions, finite convergence of alternating projections is guaranteed. Additionally, the number of iterations depends on the distance between the two sets, the starting point and a constant, which is the maximum angle between vectors of type $(a - b)$, where $a \in A$ and $b \in B$, and the proximal normal cones $N^\text{prox}_A(a)$ and $N^\text{prox}_B(b)$. Using these conditions, we then prove the finite convergence of alternating projections and calculate the number of steps for a case when the two sets are a polyhedron and a closed half space.

Projection methods are often used as robust methods for solving large-scale linear programming problems and systems of linear equations (see [2,3,7,12,13,19,20]). A linear programming problem can indeed be solved by finding the closest points of the following sets

(1) the problem feasible’s region; and

(2) the closed half space containing all vectors whose objective function value does not exceed a specified lower bound for the cost function.

These two sets are polyhedrons and hence our results for finite convergence of the alternating projection method are applicable. We can also determine the minimum distance (or lower bound), relative to the starting point, needed to ensure the convergence after one iteration. This idea is valid not just for linear programming problems, but also some other optimization problems.

The paper is organized as follows. In Section 2, we recall some essential results that will be used in subsequent sections, and we provide a new proof for one of these results. Section 3 contains our main result on the finite convergence of the alternating projection method. Section 4 explores its finite convergence for some special cases and then discusses the application of this technique to linear programming.

2 Preliminaries and auxiliary results

Let $\mathbb{B}$ denote the open unit ball and furthermore let $B_\delta(x)$ and $\overline{B}_\delta(x)$ denote, respectively, the open and closed balls with center $x$ and radius $\delta > 0$. We use $\mathbb{R}$ and $\mathbb{R}_+$ to denote the real line (with the usual norm) and the set of all non-negative real numbers. The boundary and interior of a set $A$ are denoted as $\text{bd} A$ and $\text{int} A$, respectively. The distance from a point $x$ to a set $A$ is defined by $d(x, A) := \inf_{u \in A} \|u - x\|$, and we use the convention $d(x, \emptyset) = +\infty$. The set of all projections of $x$ onto $A$ is

$$ P_A(x) := \{a \in A : d(x, a) = d(x, A)\}. $$

If $A$ is a closed subset of a finite dimensional space, then $P_A(x) \neq \emptyset$. Additionally, if $A$ is a closed convex set of an Euclidean space, then $P_A(x)$ is a singleton. As defined in [17], the proximal normal cone to $A$ at $a \in A$ is defined as follows:

$$ N^\text{prox}_A(a) := \text{cone}(P_A^{-1}(a) - a) = \{\lambda(x - a) : \lambda \geq 0, a \in P_A(x)\}. $$

For convenience, we will use the notation $N_A(a)$ instead of $N^\text{prox}_A(a)$ throughout. Observe that if $a \in P_A(x)$, then $x - a \in N_A(a)$. A functional counterpart of the proximal normal cone is the proximal subdifferential of a proper lower semicontinuous function, denote $\partial_P f$, or $\partial f$ for simplicity. Note that the
fuzzy sum rule holds for $\partial f$. Note also that $\partial I_A(a) = N_A(a)$ for any $a \in A$, here $I_A$ is the indicator function of the set $A$. If we replace proximal normal cones by Fréchet normal cones, the analysis throughout the paper still holds.

The following result has been proved in [11] and here we give an alternative proof.

Definition 2 (Intrinsic transversality) [11, Definition 3.1] Given two closed sets $A, B$ of a Hilbert space $X$, $\bar{x} \in A \cap B$, we say that $(A, B)$ is intrinsically transversal at $\bar{x}$ with degree $\alpha \in (0, 1)$ if there is $\rho > 0$ such that for all $x \in (A \setminus B) \cap B_\rho(\bar{x})$, $y \in (B \setminus A) \cap B_\rho(\bar{x})$, we have

$$\max \left\{ d\left( \frac{x-y}{\|x-y\|}, N_B(y) \right), d\left( \frac{y-x}{\|x-y\|}, N_A(x) \right) \right\} \geq \alpha. \quad (1)$$

The next theorem establishes the linear convergence of the alternating projection method under intrinsic transversality assumption.

Theorem 2 (Linear convergence) [11, Theorem 6.1] If two closed sets $A, B$ of an Hilbert space $X$, a closed set $A$, and points $a \in A, b \notin A$ with $\rho := \|a-b\|$ and $\alpha > 0$. If there is $\delta > 0$ such that

$$\inf \left\{ d\left( \frac{b-x}{\|b-x\|}, N_A(x) \right) : x \in B_\rho(b) \cap B_\delta(a) \cap A \right\} \geq \alpha, \quad (2)$$

then $d(b, A) \leq \|a-b\| - \alpha \delta$.

Proof Consider the continuous function $f(x) = \|x-b\|$ and suppose to the contrary that $d(b, A) > \|a-b\| - \alpha \delta$. Take $\alpha' \in (0, \alpha)$ such that $d(b, A) > \|a-b\| - \alpha' \delta$. This is equivalent to $\inf f(x) > f(a) - \alpha' \delta$. By Ekeland Variational Principle, there is a point $x_0 \in A \cap B_\delta(a)$ such that

$$f(x_0) < f(a) \quad (3)$$

$$f(x) \leq f(x_0) + \alpha' \|x-x_0\|, \quad \forall x \in A. \quad (4)$$

Due to (3), $\|x_0 - b\| < \|a-b\| = \rho$, or $x_0 \in A \cap B_\rho(b)$. By (4), we have $x_0$ is the global minimizer of the sum $f(x) + \alpha' \|x-x_0\| + I_A(x)$. Thus,

$$0 \in \partial \left( f(x_0) + I_A(x_0) + \alpha' \|x-x_0\| \right).$$

Take $\epsilon > 0$ such that $\epsilon < \min \left\{ \alpha - \alpha', \rho - \|x_0 - b\|, \delta - \|x_0 - x\| \right\}$. There is $\bar{x} \in A \cap B_\epsilon(x_0)$ such that $\partial f(\bar{x}) + I_A(\bar{x}) \cap (a\mathbb{B}) \neq \emptyset$. On the other hand, $f(\bar{x}) = \|\bar{x} - b\| > \|x_0 - b\| - \|\bar{x} - x_0\| > 0$, then the function $f$ is differentiable at $\bar{x}$ and $\nabla f(\bar{x}) = \frac{b-\bar{x}}{\|b-\bar{x}\|}$. Thus, $\partial f(\bar{x}) + I_A(\bar{x}) = \frac{b-\bar{x}}{\|b-\bar{x}\|} + N_A(\bar{x})$. Hence,

$$\left( \frac{b-\bar{x}}{\|b-\bar{x}\|} + N_A(\bar{x}) \right) \cap (a\mathbb{B}) \neq \emptyset, \quad \text{or} \quad d\left( \frac{b-\bar{x}}{\|b-\bar{x}\|}, N_A(\bar{x}) \right) < \alpha. \quad (5)$$

Since $\bar{x} \in A \cap B_\rho(b)$ and $\|\bar{x} - a\| \leq \|\bar{x} - x_0\| + \|\bar{x} - a\| < \delta$, the previous inequality contradicts (2). \qed

The next proposition provides complementary results for the convergence of the alternating projections.

Proposition 1 For two closed sets $A, B$ of a Hilbert space $X$ with $d(A, B) > 0$ and $a \in A, b \in B, \text{ if } \|a-b\| = d(A, B)$, then

$$(b-a) \in N_A(a), \quad \text{and } \quad (a-b) \in N_B(b); \quad (5)$$

furthermore, if $A, B$ are convex, then condition (5) is sufficient.

Proof If $\|a-b\| = d > 0$ with $a \in A, b \in B$, then $a \in P_A(b)$ and $b \in P_B(a)$. By the definition of the proximal normal cones, $b-a \in N_A(a)$ and $a-b \in N_B(b)$.

Consider the distance function restricted to the sets $A$ and $B$, $f : X \times X \to \mathbb{R}_{+}, f(x, y) = \|x-y\| + 1_{A \times B}(x, y)$, here $1_{A \times B}$ is the indicator function of the set $A \times B$. The product space $X \times X$ is equipped with the usual sum norm. When two sets $A, B$ are convex, the function $f$ is convex. We have $(a, b)$ is the global minimizer of $f$, or equivalently a pair of shortest distance between $A$ and $B$ if and only if

$$0 \in \partial f(a, b) = \partial \|a-b\| + N_{A \times B}(a, b) = \partial \|a-b\| + N_A(a) \times N_B(b).$$

Since $A \cap B = \emptyset$, then $a-b \neq 0$ and $\partial \|a-b\| = \left\{ \left( \frac{a-b}{\|a-b\|}, \frac{b-a}{\|b-a\|} \right) \right\}$. The inclusion $0 \in \partial f(a, b)$ is equivalent to (5). \qed

The next theorem establishes the linear convergence of the alternating projection method under intrinsic transversality assumption.

Theorem 3 (Linear convergence) [11, Theorem 6.1] If two closed sets $A, B$ of an Euclidean space $X$ are intrinsically transversal at a point $\bar{x} \in A \cap B$, with degree $\alpha > 0$, then, for any constant $c$ in the interval $(0, \alpha)$ the method of alternating projections, initiated sufficiently near $\bar{x}$, converges to a point in the intersection $A \cap B$ with linear rate $1 - c^2$. \footnote{This theorem is a special case of the more general result in [11, Theorem 6.1].}
3 Convergence results

We extend the definition of the intrinsic transversality in Definition 1 to more general frameworks without the assumption $A \cap B = \emptyset$, relaxing $\bar{x}$ and its local neighborhood $B_\rho(\bar{x})$. Condition 1 is applied for the whole sets $A, B$. In Condition 2, we consider local neighbourhoods around the given points.

**Condition 1** Given two closed sets $A, B$ of a Hilbert space $X$ and a constant $\alpha \in (0,1)$, for all $x \in A \setminus B, y \in B$ such that $d(y, A) > d(A, B)$ (or $d(x, B) > d(A, B)$), inequality (1) holds.

**Condition 2** Given two closed sets $A, B$ of a Hilbert space $X$, $a \in A, b \in B$ such that $\|a - b\| = d(A, B) \geq 0$ and a constant $\alpha \in (0,1)$, then there is a constant $\rho > 0$ such that for all $x \in (A \setminus B) \cap B_\rho(a)$ and $y \in B$ such that $d(y, A) > d(A, B)$ (or $d(x, B) > d(A, B)$), inequality (1) holds.

**Remark 1** If $A \cap B \neq \emptyset$, then $d(A, B) = 0$ and Condition 2 reduces to Definition 1 and Condition 1 reduces to the following condition.

**Condition 1’** Given two closed sets $A, B$ of a Hilbert space $X, A \cap B \neq \emptyset$, and a constant $\alpha \in (0,1)$, for all $x \in A \setminus B, y \in B \setminus A$, the pair $\{A, B\}$ is intrinsically transversal at $x$ with $\rho = +\infty$.

We will show later in this section that under Conditions 1 or 2, when $A \cap B = \emptyset$, the method of alternating projection converges after finite number of steps. First, we establish the key result for the convergence theorems.

**Lemma 1** Suppose $A, B$ are closed subsets of a Hilbert space $X$, $x \in A \setminus B$ and $y \in P_B(x)$ satisfying $d(y, A) > d(A, B)$, and $\alpha \in (0,1), \delta = \alpha \|x - y\|$. If inequality (1) holds with any vector $z \in A \cap B_\delta(x) \setminus B$ in place of $x$, then

$$d(y, A) \leq (1 - \alpha^2) \|x - y\|. \quad (6)$$

Consequently, if Condition 1 holds, inequality holds for any $x \in A \setminus B$ and $y \in P_B(x)$ such that $d(y, A) > d(A, B)$. Let $\rho := \|x - y\| > 0$ and $\delta := \alpha \|x - y\| > 0$. Take $z \in A \cap B_\delta(x) \cap B_\rho(y)$. Then,

$$d(z, B) \geq d(x, B) - \|x - z\| = \|x - y\| - \|x - z\| \geq \|x - y\| - \delta = (1 - \alpha) \|x - y\| > 0.$$  

This implies $z \in A \setminus B$. Hence, inequality (1) holds with $z$ in place of $x$. On the other hand,

$$d \left( \frac{z - y}{\|z - y\|}, N_B(y) \right) \leq d \left( \frac{z - y}{\|z - y\|}, R_+ (x - y) \right)$$

$$= d \left( \frac{z - y}{\|z - y\|}, R_+ (z - y) \right)$$

$$\leq \frac{\|x - y\|}{\|z - y\|} - \frac{\|z - y\|}{\|x - y\|} \leq \frac{\|z - x\|}{\|x - y\|} < \alpha \frac{\|x - y\|}{\|x - y\|} = \alpha.$$  

The estimations above together with inequality (1) yield $d \left( \frac{y - z}{\|y - z\|}, N_A(z) \right) \geq \alpha$. Hence,

$$\inf \left\{ d \left( \frac{y - z}{\|y - z\|}, N_A(z) \right) : z \in A \cap B_\delta(x) \cap B_\rho(y) \right\} \geq \alpha.$$  

Then, applying Theorem 1, we obtain $d(y, A) \leq \|x - y\| - \alpha \delta = \|x - y\| (1 - \alpha^2)$. \Box

Lemma 1 studies the main building block of the method, which involves two successive projections:

$$x_{2n-1} \in A, \quad x_{2n} \in P_B(x_{2n-1}), \quad \text{and} \quad x_{2n+1} \in P_A(x_{2n}). \quad (7)$$

Under Condition 1, thanks to Lemma 1 we have

$$\|x_{2n+1} - x_{2n}\| = d(x_{2n}, A) \leq (1 - \alpha^2) \|x_{2n} - x_{2n-1}\|.$$  

This idea plays the core role in the following theorem.

**Theorem 3** Suppose $A, B$ are closed subsets of a Hilbert space $X$ and Condition (1) holds for some $\alpha \in (0,1)$. Consider a sequence of alternating projections $x_{2n} \in A$ and $x_{2n+1} \in B$ $(n \geq 0)$. 

(i) If \( d(A, B) > 0 \), then the alternating projections \((z_n)\) converge finitely to a pair of points of nearest distance. Furthermore, the number of steps is \(2N\) with
\[
N := \left\lceil \log_{1-\alpha^2} \left( \frac{d(A, B)}{d(x_0, B)} \right) \right\rceil. \tag{8}
\]
(ii) If \( d(A, B) = 0 \), then the alternating projections \((x_n)\) converge linearly to a point in the intersection \(A \cap B\) with the rate \((1 - \alpha^2)^n\), i.e.
\[
\|x_{n+1} - x_n\| \leq (1 - \alpha^2)^n \|x_1 - x_0\|. \tag{9}
\]

Proof: Assume that Condition 1 holds with \(\alpha \in (0, 1)\).

(i) Let \( d(A, B) > 0 \). Choose an initial point \(x_0 \in A\). Suppose the algorithm does not converge after \(2n_0 - 1\) steps with \(n_0 \geq 1\). Since \(x_{2n+1} \in P_B(x_{2n})\) and \(x_{2n+2} \in P_A(x_{2n+1})\) for \(n \in [1, 2n_0]\), by Lemma 1, we have
\[
\|x_2 - x_1\| \leq (1 - \alpha^2)\|x_1 - x_0\|;
\]
\[
\|x_4 - x_3\| \leq (1 - \alpha^2)\|x_3 - x_2\| \leq (1 - \alpha^2)^2\|x_1 - x_0\|;
\]
\[
\vdots
\]
\[
\|x_{2n} - x_{2n-1}\| \leq (1 - \alpha^2)^n\|x_1 - x_0\|.
\]

On the other hand, we also have
\[
d(A, B) < d(x_{n_0}, A) = \|x_{2n} - x_{2n-1}\| \leq (1 - \alpha^2)^n\|x_0 - x_1\| = (1 - \alpha^2)^n d(x_0, B). \tag{10}
\]
This implies \(n < \log_{1-\alpha^2} \left( \frac{d(A, B)}{d(x_0, B)} \right)\). Hence \(n < \left\lceil \log_{1-\alpha^2} \left( \frac{d(A, B)}{d(x_0, B)} \right) \right\rceil = N\). Therefore, the constant \(N\) defined in (8) is an upper bound of all \(n_0\) and all the equalities in (10) attain with \(d(A, B) = d(x_{2n} - x_{2n-1}, A) = \|x_{2n} - x_{2n-1}\|\) when \(n = N\). Thus, the alternating projections sequence converges to a pair of points of nearest distance.

(ii) Let \( d(A, B) = 0 \). Observe that the condition \(d(y, A) > d(A, B)\) is equivalent to \(y \in B \setminus A\). Therefore, we can apply Lemma 1 for both iterations \(x_{2n} \in A, x_{2n+1} \in P_B(x_{2n})\) and \(x_{2n+1} \in B, x_{2n+2} \in P_A(x_{2n+1})\) to obtain
\[
\|x_1 - x_0\| \geq \frac{1}{1 - \alpha^2} \|x_2 - x_1\| \geq \frac{1}{(1 - \alpha^2)^2} \|x_3 - x_2\| \geq \ldots \geq \frac{1}{(1 - \alpha^2)^n} \|x_{2n+1} - x_{2n}\|
\]
which yields (9). \(\square\)

The following examples demonstrate the application of Condition 1 in Theorem 3.

**Example 1** Consider the space \(X = \mathbb{R}^2\) equipped with the Euclidean norm.

(i) Given two closed sets \(A := \{(u, v) : v \leq 0\}, B := \{(u, v) : v \geq |u|\}\) and \(x_0 \in A\), see Figure 3, we show that Condition 1 holds for this setting.

Take \(x \in A \setminus B\) and \(y \in B \setminus A\). If \(y \in \text{int} B\) or \(x \in \text{int} A\), the proximal normal cones at these points are trivial i.e. \(N_B(y) = \{0\}\) or \(N_A(x) = \{0\}\), thus \(d \left( \frac{x - y}{\|x - y\|}, N_B(y) \right) = 1\) or \(d \left( \frac{y - x}{\|x - y\|}, N_A(x) \right) = 1\),

![Diagram](attachment:image.png)
respectively. Hence, it is sufficient to consider \( x \in \text{bd} \ A \) and \( y \in \text{bd} \ B \). Take \( x = (x_1, |x_1|) \) and \( y = (x_2, 0) \) with \( x_1, x_2 \in \mathbb{R}, x_1 \neq 0 \). Observe that

\[
N_B(y) = \begin{cases} \mathbb{R}_+(1, -1) & x_1 > 0 \\ \mathbb{R}_+(-1, -1) & x_1 < 0 \end{cases}; \quad N_A(x) = \mathbb{R}_+(0, 1).
\]

We have \( \frac{x - y}{\|x - y\|} = \left( \frac{x_1 - x_2}{\sqrt{(x_1 - x_2)^2 + x_1^2}}, \frac{|x_1|}{\sqrt{(x_1 - x_2)^2 + x_1^2}} \right) \) and

\[
d\left( \frac{y - x}{\|y - x\|}, N_A(x) \right) = \sqrt{\frac{(x_1 - x_2)^2}{(x_1 - x_2)^2 + x_1^2}},
\]

\[
d\left( \frac{x - y}{\|x - y\|}, N_B(y) \right) = \sqrt{\frac{x_2^2}{2(x_1 - x_2)^2 + 2x_1^2}}.
\]

The two equalities above imply

\[
d^2 \left( \frac{y - x}{\|y - x\|}, N_A(x) \right) + d^2 \left( \frac{x - y}{\|x - y\|}, N_B(y) \right) = \frac{2(x_1 - x_2)^2 + x_2^2}{2(x_1 - x_2)^2 + 2x_1^2} \geq \frac{(x_1 - x_2)^2 + [(x_2 - x_1)^2 + x_2^2]}{2(x_1 - x_2)^2 + 2x_1^2} \geq \frac{1}{4}.
\]

Hence,

\[
\max \left\{ d\left( \frac{x - y}{\|x - y\|}, N_B(y) \right), d\left( \frac{y - x}{\|y - x\|}, N_A(x) \right) \right\} \geq \frac{1}{2\sqrt{2}}.
\]

The method of alternating projections, initiated at \( x_0 \), converges linearly to \( \bar{x} \) with the rate 7/8. \( \Box \)

(ii) By moving \( B \) with a shift \( b \), where \( b = (0, k), k > 0 \), the setting \( \{A, (B+b)\} \) has empty intersection with \( d(A, B+b) = \|b\| = k \), see Figure 3. The vectors \( b \) and \( a = \bar{x} = (0, 0) \) are nearest points of the sets \( A, (B+b) \). Take \( x \in \text{bd} \ A \) and \( y \in \text{bd} \ B \) with \( x = (x_1, |x_1| + k), y = (x_2, 0), x_1, x_2 \in \mathbb{R}, x_1 \neq 0 \). We apply the analysis in (i) for the pair \( x' = (x_1 + k, |x_1| + k) \) and \( y' = (x_2 + k, 0) \) if \( x_1 > 0 \) or \( x' = (x_1 - k, |x_1| + k) \) and \( y' = (x_2 - k, 0) \) if \( x_1 < 0 \) to derive

\[
\max \left\{ d\left( \frac{x' - y'}{\|x' - y'\|}, N_B(y') \right), d\left( \frac{y' - x'}{\|y' - x'\|}, N_A(x') \right) \right\} \geq \frac{1}{2\sqrt{2}}.
\]

Hence, the alternating projections converge after \( \left\lceil \log_{7/8} \left( \frac{k}{d(x_0, B)} \right) \right\rceil \) steps. \( \Box \)

Remark 2 Theorem 3 explains the finite convergence in Example 1 (ii). Additionally, the estimation on the number of steps (8) shows that the larger the distance between the two sets relative to the starting point, the fewer steps of the method. By increasing \( k \) (in \( b = (0, k) \)), the distance between two sets \( A + b \) increases accordingly. With the same starting point \( x_0 = (u_0, 0) \in A \), if \( k \geq d(x_0, A) \), the algorithm will reach one of the nearest points after one projection, i.e. \( b \in P_{A+b}(x_0) \).

The next example shows that if the setting does not satisfy Condition 1, the linear convergence and finite convergence as in previous example does not hold.

Example 2 (i) Consider two closed sets \( A := \{(x, y) : y \leq 0 \}, B := \{(x, y) : y \geq x^2 \} \) and \( x_0 \in \text{bd} \ A = \{(x, y) : y = 0 \} \), see Figure 3. Condition 1 does not hold and the alternating projection algorithm, initiated at \( x_0 \), does not converge linearly.

(ii) Shifting \( B \) by a translation \( b = (0, k), k > 0 \), the pair \( \{A, B+b\} \) has two nearest points \( a = \bar{x} = (0, 0) \in A \) and \( b \in B \), see Figure 3. Unlike Example 1, the alternating projections, started at \( x_0 \), does not converge to \( a, b \) after finite number of steps.

Indeed, if the algorithm reaches \( b \) after \( n \) steps, this implies there is an \( x \in (a, x_0) \) such that \( b \in P_{B+b}(x) \), or \( b - x \in N_{B+b}(b) \), which contradicts the facts that \( x \neq a \) and \( N_{B+b}(b) = \mathbb{R}_+(a - b) \). \( \Box \)
**Theorem 4** Suppose $A, B$ are closed subsets of a Hilbert space $X$, $a \in A, b \in B$ such that $\|a - b\| = d(A, B)$ and Condition 2 holds for some $\alpha \in (0, 1)$ and $\rho > 0$. Consider a sequence of alternating projections $(x_n)$ with $x_{2n} \in A$ and $x_{2n+1} \in B$ ($n \geq 0$), initiated sufficiently close to $a$ (or $b$).

(i) If $d(A, B) > 0$, then $(x_n)$ converges finitely to a pair of points of nearest distance.

(ii) If $d(A, B) = 0$, then $(x_n)$ converges linearly to a point in the intersection with a rate $1 - \alpha^2$.

**Proof** Suppose that Condition 2 holds with $\alpha \in (0, 1)$ and $\rho > 0$.

(i) Suppose $d(A, B) > 0$. Take $\kappa := \min\{\alpha^2 d, \rho\}$ and $x_0 \in A \cap B_\kappa(a)$, $x_1 \in P_B(x_0)$. If $d(x_1, A) > d$, then by Lemma 1, we have

$$d(x_1, A) \leq \|x_1 - x_0\| - \alpha^2 \|x_1 - x_0\| \leq \|x_0 - b\| - \alpha^2 d$$

which is a contradiction. This implies $d(x_1, A) = d(A, B)$. Then, $(x_n)$ converges after 2 steps.

(ii) Note that with $d(A, B) = 0$, we have $a \equiv b \equiv \bar{x} \in A \cap B$. If $x \in A \cap B_\kappa(\bar{x})$ and $y \in P_B(x)$, then $\|y - \bar{x}\| \leq \|y - x\| \leq \|x - \bar{x}\| \leq \rho$, thus $y \in B_\rho(\bar{x})$. We can apply Lemma 1 for both iterations $x_{2n} \in A$, $x_{2n+1} \in P_B(x_{2n})$ and $x_{2n+2} \in P_A(x_{2n+1})$ to obtain

$$\|x_1 - x_0\| \geq \frac{1}{1 - \alpha^2} \|x_2 - x_1\| \geq \frac{1}{1 - \alpha^2} \|x_3 - x_2\| \geq \ldots \geq \frac{1}{(1 - \alpha^2)^{2n}} \|x_{2n+1} - x_{2n}\|.$$

Thus, the alternating projections converge linearly with the rate $1 - \alpha^2$. \qed

**Remark 3** (i) Since Condition 2 is weaker than intrinsic transversality property, Theorem 4 is rich enough to cover Theorem 2.

(ii) When two sets $A \cap B = \emptyset$ and $a \in A, b \in B$ with $\|a - b\| = d(A, B)$, it is hard to localize the setting $A, B$ at $a$ and $b$ in the sense that all the projections of any points sufficiently near $a$ or $b$ will remain stay near $a, b$. In the global version (Condition 1), we do not encounter this obstacle. This explains why we cannot estimate an upper bound for the number of steps for the method in Theorem 4.

**Condition 3** Given two closed subsets $A, B$ of a Hilbert space $X$ and constants $\alpha \in (0, 1), \beta \in [0, 1]$, then for all $x \in A \setminus B$ and $y \in P_B(x)$ such that $d(y, A) > d(A, B)$, then

$$\inf_{z \in A \cap B_\delta(y) \cap B_\beta(x)} \left( \frac{y - z}{\|y - z\|^2}, N_A(z) \right) \geq \alpha,$$

with $\delta = \alpha(\|x - y\| - \beta d(A, B)) > 0$ and $\rho = \|x - y\|$.

In Condition 3, instead of considering all vectors $y \in B \setminus A$, we only consider the projections of the vector $x$ onto $A$. This condition also gives us flexibility in choosing the neighborhoods of $x$ in $B$. The next theorem presents the convergence results of the method under this condition.

**Theorem 5** Suppose $A, B$ are closed subsets of a Hilbert space $X$ with $d(A, B) > 0$ and Condition 3 holds for some $\alpha \in (0, 1)$ and $\beta \in [0, 1]$. Consider a sequence of alternating projections $(x_n)$ with $x_{2n} \in A$ and $x_{2n+1} \in B$ ($n \geq 0$).
(i) If $\beta < 1$, then $(x_n)$ converges finitely to a pair of points of nearest distance and the number of step is $2N$ with

$$\left\lceil \log_{(1-\alpha^2)} \left( \frac{d(1-\beta)}{\|x_1 - x_0\| - \beta d} \right) \right\rceil.$$  \hspace{1cm}  (12)

(ii) If $\beta = 1$, then the sequence $(\|x_{2n} - x_{2n+1}\| - d)$ converges linearly to 0 with the rate $(1 - \alpha^2)$.

Proof Assume that Condition 3 holds for $\alpha \in (0, 1)$ and $\beta \in [0, 1]$. Denote $d(A, B) = d$. For any $n \geq 1$, and $x_{2n} \in A, x_{2n+1} \in PB(x_{2n}), x_{2n+2} \in PA(x_{2n})$, set $\rho := \|x_{2n} - x_{2n-1}\|$ and $\delta := \alpha \|x_{2n+1} - x_{2n}\|$. By Condition 3, for any $z \in A \cap B_P(x_{2n+1}) \cap B_{\delta}(x_{2n+2})$, inequality (11) holds. By Theorem 1, we have

$$\|x_{2n+2} - x_{2n+1}\| \leq \|x_{2n+1} - x_{2n}\| - \alpha^2(\|x_{2n+1} - x_{2n}\| - \beta d) = (1 - \alpha^2)\|x_{2n+1} - x_{2n}\| + \alpha^2\beta d.$$  \hspace{1cm}  (13)

Suppose the algorithm does not converge after $2n_0 - 1$ steps with $n_0 \geq 1$. Since $x_{2n+1} \in PB(x_{2n})$ and $x_{2n+2} \in PA(x_{2n+1})$ for $n \in [1, 2n_0]$, by Lemma 1, we have

$$\frac{\|x_{2n+1} - x_{2n}\|}{d} - 1 \leq (1 - \alpha^2) \frac{\|x_{2n} - x_{2n-1}\|}{d} - \alpha^2(1 - \beta).$$

The following estimations hold for any $n \leq n_0$:

$$\|x_2 - x_1\| \leq (1 - \alpha^2)\|x_1 - x_0\| + \alpha^2\beta d$$

$$\|x_4 - x_3\| \leq (1 - \alpha^2)\|x_3 - x_2\| + \alpha^2\beta d$$

$$\leq (1 - \alpha^2)^2\|x_1 - x_0\| + \alpha^2\beta d(1 + (1 - \alpha^2))$$

$$\vdots$$

$$\|x_{2n} - x_{2n-1}\| \leq (1 - \alpha^2)^n\|x_1 - x_0\| + \alpha^2\beta d \frac{\sum_{k=1}^{n-1} (1 - \alpha^2)^k}{\alpha^2}$$

$$= (1 - \alpha^2)^n\left(\frac{\|x_1 - x_0\|}{d} - 1\right) + \alpha^2\beta d \frac{1 - (1 - \alpha^2)^n}{\alpha^2}$$

$$= (1 - \alpha^2)^n(\|x_1 - x_0\| - \beta d) + \beta d.$$

From here we consider two cases.

(i) If $\beta < 1$, from the above inequalities, we have

$$n \leq \log_{(1-\alpha^2)} \left( \frac{d(1-\beta)}{\|x_1 - x_0\| - \beta d} \right).$$

Hence, $n_0 \leq \left\lceil \log_{(1-\alpha^2)} \left( \frac{d(1-\beta)}{\|x_1 - x_0\| - \beta d} \right) \right\rceil = N$ and the sequence $(x_n)$ converges after $N$ steps.

(ii) If $\beta = 1$, we have

$$\|x_{2n} - x_{2n-1}\| - d \leq (1 - \alpha^2)^n(\|x_1 - x_0\| - d).$$

The sequence $(\|x_{2n} - x_{2n-1}\| - d)$ converges linearly to 0 with a linear rate $(1 - \alpha^2)$. \hspace{1cm} \Box

Remark 4 Theorem 5 states that if the neighborhood associated with the vector $x$ in Condition 3 is small, the method will not converge finitely. Observe that in (i), the radius $\delta$ is always bigger than the constant $\alpha(1 - \beta)d(A, B)$. On the other hand, in (ii), the radius $\delta$ tends to 0 when $\|x - y\| \to d(A, B)$. When $d(A, B) = 0$, Theorem 5 can be formulated as in Theorem 3.
4 Alternating projections for two polyhedrons

In this section, we apply the results in Section 3 for some special cases in Euclidean spaces. First, we consider the case when the two sets are a polyhedron and a closed half space.

**Proposition 2** Suppose $A, B$ are closed subsets of an Euclidean space $X = \mathbb{R}^n$ ($n \geq 1$) with $A \cap B \neq \emptyset$. Suppose further that $A$ is a closed half space and $B$ is a polyhedron

$$A := \{x : \langle c, x \rangle \leq M\}, \quad \text{and} \quad B := \{x : Ax \leq b\},$$

here $A = \begin{bmatrix} a_1 & \cdots & a_m \end{bmatrix}$ is a matrix $m \times n$ with $a_i \in \mathbb{R}^n$ ($i = 1, \ldots, m$), $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n \setminus \{0\}$ and $M \in \mathbb{R}$. Then, the pair $\{A, B\}$ satisfies Condition 3 with $\beta = 0$ and

$$\alpha := \frac{1}{2} \times \min_{0 > (a_i, c) > -\|a_i\|c} d \left( \frac{a_i}{\|a_i\|}, \mathbb{R}_+(-c) \right), \quad (14)$$

with the convention that the minimum over an empty set is 1.

**Proof** Take $x \in A$ and $y \in P_B(x)$ such that $d(y, A) > d(A, B)$. Due to Proposition 1 (i), $y - x \notin N_A(x) = \mathbb{R}_+(-c)$ and $c \notin N_B(y)$, hence $(y - x, c) < \|y - z\| \cdot \|c\|$. On the other hand, $(c, x) \leq M < (c, y)$, therefore, $(c, x - y) < 0$ and due to (14) together with $c \notin N_B(y)$ and $x - y \in N_B(y)$, we have

$$d \left( \frac{c}{\|c\|}, \mathbb{R}_+(y - x) \right) = d \left( \frac{x - y}{\|x - y\|}, \mathbb{R}_+(-c) \right) \geq 2\alpha.$$

Set $\delta := \alpha \|x - y\|$. Take $z \in A \cap B_\delta(x) \cap B_\delta(y)$. Observe that either $z \in \text{int } A$ and $N_A(z) = \{0\}$ or $z \in \text{bd } A$ and $N_A(z) = \mathbb{R}_+(c)$. For both cases, we have the following estimations

$$d \left( \frac{y - z}{\|y - z\|}, N_A(z) \right) \geq d \left( \frac{y - z}{\|y - z\|}, \mathbb{R}_+(-c) \right) = d \left( \frac{c}{\|c\|}, \mathbb{R}_+(y - z) \right) = \min_{t \geq 0} \frac{c}{\|c\|} - t(y - z) \geq \min_{t \geq 0} \left( \frac{c}{\|c\|} - t(y - x) \right) - t \left\|x - z\right\| \geq \min_{t \geq 0} \left( \frac{c}{\|c\|} - t(y - x) \right) - \|x - y\| = d \left( \frac{c}{\|c\|}, \mathbb{R}_+(y - x) \right) - \alpha \geq 2\alpha - \alpha = \alpha.$$

Hence, inequality (11) holds. The pair $\{A, B\}$ satisfies Condition 3 with $\beta = 0$ and $\alpha$ defined as in (14). □

**Theorem 6** Let $X$ be an Euclidean space, $A$ a closed subspace and $B$ a polyhedron defined as in Proposition 2 with $A \cap B = \emptyset$. Then, the method of alternating projections, initiated at $x_0 \in A$, converges after $2N$ steps with

$$N := \left\lceil \log_{1-\alpha^2} \left( \frac{d(A, B)}{d(x_0, B)} \right) \right\rceil,$$

where the constant $\alpha$ is given in (14). Furthermore, if $d(x_0, B) \leq \frac{d(A, B)}{1 - \alpha^2}$, then the alternating projections converge after 2 steps.

**Proof** The constant $\alpha$ given in (14) is positive due to the fact that the set $S := \{a_i : \langle a_i, c \rangle > -\|a_i\|c\}, i = 1, \ldots, m$ is finite and $d \left( \frac{a_i}{\|a_i\|}, \mathbb{R}_+(-c) \right) > 0$ for all $a_i \in S$. Thanks to Proposition 2 and Theorem 5, we conclude that the method of alternating projections converge after $2N$ steps with

$$N = \left\lceil \log_{1-\alpha^2} \left( \frac{d(A, B)}{\|x_1 - x_0\|} \right) \right\rceil = \left\lceil \log_{1-\alpha^2} \left( \frac{d(A, B)}{d(x_0, B)} \right) \right\rceil.$$

When $d(x_0, B) \leq \frac{d(A, B)}{1 - \alpha^2}$, we have $N = 1$ and the method converges after 2 steps. □
Remark 5 For two closed sets $A \cap B \neq \emptyset$ and vectors $a \in A$ and $b \in B$ with $\|a - b\| = d(A, B)$, the sets $A$ and $B$ are separated by
\[
\inf_{y \in B} \langle b - a, y \rangle \geq \sup_{y \in A} \langle b - a, y \rangle.
\]
If $B$ is a polyhedron, we can replace $A$ by the closed half space
\[C := \left\{ x : \langle b - a, x \rangle \leq \sup_{y \in A} \langle b - a, y \rangle \right\}.
\]
We have $A \subset C$ and $C \cap B = \emptyset$ with $d(C, B) = \inf_{y \in B} \langle b - a, y \rangle - \sup_{y \in A} \langle b - a, y \rangle = d(A, B)$. If the starting point $x_0 \in A \subset C$ is close enough to $B$, i.e.
\[d(x_0, B) \leq \frac{d(A, B)}{1 - \alpha^2},
\]
with $\alpha$ defined in (14), then by Theorem 6, the projections of $x_0$ onto $B$ belong to the set $\{ b \in B : d(b, C) = d(C, B) = d(A, B) \}$, the set of nearest points between $B$ and $C$. Therefore, if we have $\{ b \in B : d(b, A) = d(A, B) \} = \{ b \in B : d(b, C) = d(A, B) \}$, then the alternating projections for two sets $A, B$ if converges, to a pair of points of nearest distance, will converge finitely. Generally, we only have $\{ b \in B : d(b, A) = d(A, B) \} \subset \{ b \in B : d(b, C) = d(A, B) \}$.

Note also that if two sets in $\mathbb{R}^n$ are convex and one of the set is bounded (i.e. compact), the method of alternating projections always converge to a pair of points of shortest distance.

![Diagram](image_url)

**Fig. 2** Finite convergence of the alternating projection method

The alternating projections for two polyhedrons, in general, are not necessary to converge finitely. The following example presents a simple counterexample.

**Example 3** Consider the three dimensional space $\mathbb{R}^3$. Suppose $A := \{(x_1, x_2, x_3) : x_2 = x_3 = 0 \}$ and $B := \{(x_1, x_2, x_3) : x_2 = 1, x_1 = x_2 \}$ are skew lines in $\mathbb{R}^3$. The method of alternating projections does not converge finitely for this setting.
We propose a projection method for solving linear programming (LP).

\[ \min \langle c, x \rangle \]
\[ x \in A, \]

where \( c \in \mathbb{R}^n \) and \( A \) is a polyhedron. We assume that LP is bounded with \( M \) is a lower bound. Set

\[ A := \{ x \in \mathbb{R}^n : Ax \leq b \} \quad \text{and} \quad B := \{ x \in \mathbb{R}^n : \langle c, x \rangle \leq M \}. \]

A solution of LP will be obtained by applying iteratively the following steps:

**Algorithm**

1. Let \( x_1 \in A \). Let \( y_1 = x_1 + \frac{M - \langle c, x_1 \rangle}{\|c\|^2}c. \)
2. While \( \|x_{n+1} - y_n\| < \|x_n - y_n\| \),
   - project \( x_n \) to \( B \):
     \[ y_n := x_n + \frac{M - \langle c, x_n \rangle}{\|c\|^2}c; \]
   - project \( y_n \) to \( A \):
     \[ x_{n+1} \in P_A(y_n); \]
   end while.
3. Print solution \( x_{n+1} \).

By Theorem 6, a solution of LP is the following projection when the new lower bound \( D \leq M \) is sufficiently small.

\[ y_1 = x_1 + \frac{D - \langle c, x_1 \rangle}{\|c\|^2}c, \quad D \leq 1 - \frac{\alpha^2}{\alpha^2} \langle c, x_1 \rangle - M, \quad \text{and} \quad \alpha := \frac{1}{2} \times \min_{i=1, \ldots, m; \langle a_i, c \rangle > 0} \left( \frac{a_i}{\|a_i\|} \right) \| \mathbb{R}_+^k(c) \|. \]

This is because \( d(A, B) \geq M - D \) and \( d(x_1, B) = \langle c, x_1 \rangle - D \). Thus, \( d(x_1, B) \leq \frac{d(A, B)}{1 - \alpha^2} \). Thanks to Theorem 6, \( d(P_A(y_1), B) = d(A, B) \), equivalently, \( P_A(y_1) \) are solutions of LP.

5 Acknowledgment

The authors are supported by the Australian Research Council through the Centre for Transforming Maintenance through Data Science (grant number IC180100030), funded by the Australian Government.

References

1. H. H. Bauschke and J. M. Borwein. On the convergence of von Neumann’s alternating projection algorithm for two sets. *Set-Valued Anal.*, 1(2):185–212, 1993.
2. U. Betke and P. Gritzmann. Projection algorithms for linear programming. *European Journal of Operational Research*, 60(3):287 – 295, 1992.
3. U. Betke and M. Henk. Linear programming by minimizing distances. *Zeitschrift für Operations Research*, 35:299–307, 1991.
4. L. M. Bregman. The method of successive projection for finding a common point of convex sets. *Sov. Math., Dokl.*, 6:688–692, 1965.
5. Yair Censor and Andzej Cegielski. Projection methods: an annotated bibliography of books and reviews. *Optimization*, 64(11):2343–2358, 2015.
6. Yair Censor and Maroun Zaknoon. Algorithms and convergence results of projection methods for inconsistent feasibility problems: A review. *arXiv: Optimization and Control*, 2018.
7. Wei Chen, Patrick L. Combettes, Ran Davidi, and Gabor T. Herman. On the effectiveness of projection methods for convex feasibility problems with linear inequality constraints. *Computational Optimization and Applications*, 51:1065–1088, 2012.
8. W. Cheney and A.A. Goldstein. Proximity maps for convex sets. *Proceedings of the American Mathematical Society*, 10:448–450, 1959.

9. P.L. Combettes. Inconsistent signal feasibility problems: Least-squares solutions in a product space. *USSR Computational Mathematics and Mathematical Physics*, 42:2955–2966, 1994.

10. P.L. Combettes and P. Bondon. Hard-constrained inconsistent signal feasibility problems. *IEEE Transactions on Signal Processing*, 47:2460–2468, 1999.

11. D. Drusvyatskiy, A. D. Ioffe, and A. S. Lewis. Transversality and alternating projections for nonconvex sets. *Found. Comput. Math.*, 15(6):1637–1651, 2015.

12. P. E. Gill, W. Murray, M. A. Saunders, J. A. Tomlin, and M. H. Wright. On projected newton barrier methods for linear programming and an equivalent to karmarkar’s projective method. *Mathematical Programming*, 36:183–209, 1986.

13. A. I. Golikov and Yu. G. Evtushenko. Finding the projection of a given point on the set of solutions of a linear programming problem. *Proceedings of the Steklov Institute of Mathematics*, 263:68–83, 2008.

14. E. Kopecka and S. Reich. A note on the von neumann alternating projections algorithm. *Journal of Nonlinear and Convex Analysis*, 5:379–386, 2004.

15. Alexander Y. Kruger and Nguyen H. Thao. Regularity of collections of sets and convergence of inexact alternating projections. *J. Convex Anal.*, 23(3):823–847, 2016.

16. A. S. Lewis, D. R. Luke, and J. Malick. Local linear convergence for alternating and averaged nonconvex projections. *Found. Comput. Math.*, 9(4):485–513, 2009.

17. Boris S. Mordukhovich. *Variational Analysis and Generalized Differentiation. I: Basic Theory*, volume 330 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer, Berlin, 2006.

18. Dominikus Noll and Aude Rondepierre. On local convergence of the method of alternating projections. *Found. Comput. Math.*, 16(2):425–455, 2016.

19. E. A. Nurminski. Single-projection procedure for linear optimization. *Journal of Global Optimization*, 66:95 – 110, 2016.

20. Lszl A. Vgh and Giacomo Zambelli. A polynomial projection-type algorithm for linear programming. *Operations Research Letters*, 42(1):91 – 96, 2014.