JOURNAL OF OPTIMIZATION THEORY AND APPLICATIONS: Vol. 120, No. 3, pp. 503–531, March 2004 (©2004)

# Reflection-Projection Method for Convex Feasibility Problems with an Obtuse Cone[1,2]

H. H. Bauschke[3] and S. G. Kruk[4]

Communicated by P. Tseng

**Abstract.** The convex feasibility problem asks to find a point in the intersection of finitely many closed convex sets in Euclidean space. This problem is of fundamental importance in the mathematical and physical sciences, and it can be solved algorithmically by the classical method of cyclic projections.

In this paper, the case where one of the constraints is an obtuse cone is considered. Because the nonnegative orthant as well as the set of positive-semidefinite symmetric matrices form obtuse cones, we cover a large and substantial class of feasibility problems. Motivated by numerical experiments, the method of reflection-projection is proposed: it modifies the method of cyclic projections in that it replaces the projection onto the obtuse cone by the corresponding reflection.

This new method is not covered by the standard frameworks of projection algorithms because of the reflection. The main result states that the method does converge to a solution whenever the underlying convex feasibility problem is consistent. As prototypical applications, we discuss in detail the implementation of two-set feasibility problems aiming to find a nonnegative [resp. positive semidefinite] solution to linear constraints in $\mathbb{R}^n$ [resp. in $\mathbb{S}^n$, the space of symmetric $n \times n$ matrices] and we report on numerical experiments. The behavior of the method for two inconsistent constraints is analyzed as well.

---

[1]The research of the first author was supported by the Natural Sciences and Engineering Research Council of Canada. Part of this work was carried out while the first author was a visiting member of the Fields Institute in Winter 2002. He thanks the Fields Institute for hospitality and support.

[2]We thank Andrzej Cegielski and Mike Todd for sending us Refs. 14–18. We also thank two referees for their unusually careful reading and helpful comments.

[3]Assistant Professor, Department of Mathematics and Statistics, University of Guelph, Guelph, Ontario, Canada.

[4]Assistant Professor, Department of Mathematics and Statistics, Oakland University, Rochester, Michigan.

**Key Words.** Convex feasibility problems, obtuse cones, projection methods, self-dual cones.

## 1. Introduction

Throughout this paper, we assume that $\mathbb{X}$ is a Euclidean space, with inner product $\langle \cdot, \cdot \rangle$ and induced norm $\|\cdot\|$, and that $C_1, \ldots, C_N$ are closed convex sets in $\mathbb{X}$, with corresponding projectors $P_1, \ldots, P_N$. The projector corresponding to a closed convex set is explained in Definition 2.1. Moreover, we suppose that $K$ is a closed convex cone in $\mathbb{X}$, with reflector $R_K = 2P_K - I$. We will require that $K$ be obtuse, a notion made precise in Definition 2.3 and broad enough to cover many interesting cones arising in optimization, including the nonnegative orthant and the cone of positive semidefinite matrices.

Let

$$C := K \cap C_1 \cap \cdots \cap C_N.$$

Our aim is to solve the following convex feasibility problem:

$$\text{find} \quad x \in C,$$

where, for the most part of this paper, we assume that $C \neq \emptyset$. The convex feasibility problem is of fundamental importance in mathematics and the physical sciences and there exists a multitude of projection algorithms for solving it; see, for instance, Refs. 1–6.

The motivation for this paper stems from a method that works very well in numerical experiments, but falls outside the scope of the standard frameworks. Specifically, we propose the method of reflection-projection: after fixing a starting point $x_0$, it generates a sequence via

$$\begin{aligned} &x_0 \mapsto R_K x_0 \mapsto P_1 R_K x_0 \mapsto P_2 P_1 R_K x_0 \mapsto \cdots \mapsto P_N \cdots P_1 R_K x_0 =: x_1 \\ &\mapsto R_K x_1 \mapsto P_1 R_K x_1 \mapsto P_2 P_1 R_K x_1 \mapsto \cdots \mapsto P_N \cdots P_1 R_K x_1 =: x_2 \\ &\mapsto R_K x_2 \mapsto \cdots . \end{aligned}$$

The terms just displayed form a sequence which has $(x_k)$ as a subsequence. The update operation for $(x_k)$ can be described more concisely by

$$x_{k+1} := (P_N P_{N-1} \cdots P_1 R_K) x_k.$$

In Figure 1, we visualize the method of reflection-projection and contrast it with the classical method of cyclic projections (which arises when the reflection is replaced by the corresponding projection) for a two-set convex

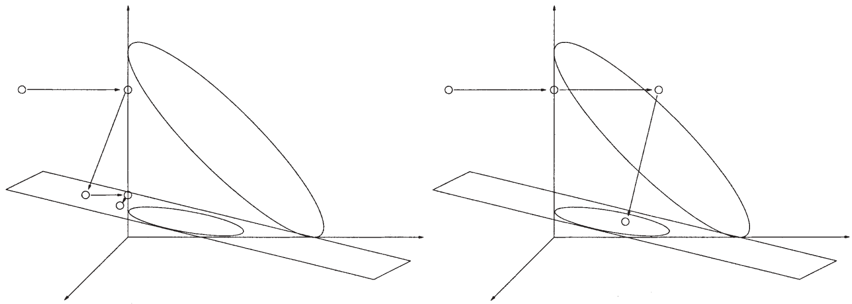

Fig. 1. Contrasting behavior of cyclic projections and the method of reflection-projection for the intersection of an icecream cone and a plane.

feasibility problem involving an icecream cone and a plane in $\mathbb{R}^3$. Of course, this particular example favors dramatically the method of reflection-projection and experiments described later bear out this advantage more generally. Nevertheless, we cannot expect the reflection-projection method to outperform always the cyclic projection method. If the starting point, for the example of Figure 1, is changed to the acute area on the lower right, both algorithms will behave in essentially the same manner because the reflection will not go deeply into the cone.

The aim of this paper is to show that the method of reflection-projection generates a sequence which converges to a solution of the convex feasibility problem. Moreover, experiments demonstrate that the method can yield a solution faster than other standard methods.

We point out that the standard theory is not applicable, since the reflector $R_K$ is nonexpansive [Lemma 2.1(ii)], but does not share any of the common properties (Remark 2.2) typically imposed on the operators in the general frameworks presented in Refs. 1, 2, 4, 6.

In fact, the only algorithmic schemes utilizing true reflections are classical and due to Motzkin and Schoenberg (Ref. 7) and Cimmino (Ref. 8). However, none of the convergence results associated with these methods cover the method of reflection-projection presented here.

The paper is organized as follows. Section 2 introduces the cones of interest along with classical convergence results based on the Fejér monotone sequences. Section 3 introduces abstractly our feasibility algorithm and the convergence proof in the consistent case. In Section 4, we review affine space projections and the Moore-Penrose inverse. This material is necessary for the practical implementations in $\mathbb{R}^n$ (see Section 5) and in $\mathbb{S}^n$ (see Section 6). Section 7 offers partial results on the inconsistent case and we conclude in Section 8. A summary of the algorithms is given in the Appendix (Section 9).

## 2. Preliminaries

### 2.1. Projections

**Definition 2.1.** Projection and Projector. Suppose that $S$ is a closed convex nonempty set in $\mathbb{X}$ and that $x\in\mathbb{X}$. Then, there exists a unique point in $S$ nearest to $x$, denoted $P_S(x)$ or $P_Sx$, and called the projection of $x$ onto $S$. Note that $P_Sx$ realizes the distance from $x$ to $S$,

$$\|x-P_Sx\| = d(x, S) := \min_{s\in S} \|x-s\|.$$

The induced map $P_S: \mathbb{X}\to S$ is called the projector.

**Fact 2.1.** The projection $P_Sx$ is characterized by $P_Sx\in S$ and $\sup\langle S-P_Sx,\ x-P_Sx\rangle\leq 0$. In particular, the projector $P_S$ is firmly nonexpansive, i.e.,

$$\|P_Sx-P_Sy\|^2 + \|(I-P_S)x-(I-P_S)y\|^2 \leq \|x-y\|^2, \qquad \forall x\in\mathbb{X},\ \forall y\in\mathbb{X}.$$

**Proof.** See Ref. 9, Chapter 12, or Ref. 10. □

### 2.2. Moreau Decomposition and Obtuse Cones

**Definition 2.2.** Polar Cone. Suppose that $K$ is a closed convex cone in $\mathbb{X}$. Then, $K^{\ominus} := \{x\in\mathbb{X}: \sup\langle x, K\rangle\leq 0\}$ is the negative polar cone of $K$. Also, $K^{\oplus} := -K^{\ominus}$ is the positive polar cone of $K$. Given $x\in\mathbb{X}$, we write $x^+ := P_Kx$ and $x^- := P_{K^{\ominus}}x$.

**Fact 2.2.** See Ref. 11 (Moreau). $P_{K^{\ominus}} = I-P_K$. Let $x\in\mathbb{X}$. Then, $x = x^+ + x^-$ and $\langle x^+, x^-\rangle = 0$.

**Proof.** See Ref. 11 or the discussion following Ref. 12, Theorem 31.5. □

**Definition 2.3.** Obtuse and Self-Dual Cones. A closed convex cone $K$ in $\mathbb{X}$ is obtuse [resp. self-dual], if $K^{\oplus}\subseteq K$ [resp. $K^{\oplus} = K$].

**Remark 2.1.** The notion of an obtuse cone was coined by Goffin (Ref. 13, Section 3.2) and implicitly used by Todd (Ref. 14, Corollary 4.1). An obtuse cone is large in the following sense:

(i) The affine span of a closed convex obtuse cone $K$ is equal to the entire space $\mathbb{X}$; in particular, $K$ has nonempty interior: indeed, let $\mathbb{Y}$

be the linear (equivalently, affine) span of $K$. Then, $\mathbb{Y}^{\perp} \subseteq K^{\oplus}$. On the one hand, this implies (multiply by $-1$) the inclusion $\mathbb{Y}^{\perp} \subseteq K^{\ominus}$. On the other hand, since $K$ is obtuse, we conclude that $\mathbb{Y}^{\perp} \subseteq K^{\oplus} \subseteq K$. Altogether, $\mathbb{Y}^{\perp} \subseteq K \cap K^{\ominus} = \{0\}$ and so $\mathbb{Y} = \mathbb{X}$.

(ii) See Theorem 3.2.1 in Ref. 13. Suppose that $K$ is a closed convex cone in $\mathbb{X}$. Then, $K$ is obtuse if and only if $K^{\oplus}$ is acute, i.e., $\inf\langle K^{\oplus}, K^{\oplus}\rangle = 0$. ("$\Rightarrow$" is easy to see; for "$\Leftarrow$", use a separation argument.)

The notions of an acute and an obtuse cone have proven quite useful in optimization; see, for instance, Refs. 13 and 15–20. The self-dual cones form an important subclass of the obtuse cones, as they include the nonnegative orthant as well as the cone of positive semidefinite matrices: these two cones are of central importance in modern interior-point methods (see Refs. 21, 22). We will discuss these cones in detail in Sections 5 and 6 below.

**Example 2.1.** Halfspaces with Zero in the Boundary. Fix $a \in \mathbb{X}\backslash\{0\}$ and let

$$K := \{x \in \mathbb{X}: \langle a, x\rangle \geq 0\}.$$

Then,

$$K^{\oplus} = \{\rho a: \rho \geq 0\}.$$

Hence, $K^{\oplus} \subseteq K$, and therefore $K$ is obtuse.

### 2.3. Ice Cream Cones

**Definition 2.4.** The ice cream cone with parameter $\alpha > 0$, denoted ice($\alpha$), is defined by

$$\text{ice}(\alpha) := \{(x, r) \in \mathbb{X} \times \mathbb{R}: \|x\| \leq \alpha r\}.$$

Note that ice($\alpha$) is a closed convex cone in $\mathbb{X} \times \mathbb{R}$. When $\alpha = 1$, one obtains the so-called second-order cone which has found important applications because of the recent successes of interior-point methods for convex programming (see Ref. 23). If $\mathbb{X} = \mathbb{R}^3$, the second-order cone becomes

$$\left\{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4: x_4 \geq \sqrt{x_1^2 + x_2^2 + x_3^2}\right\},$$

i.e., the future light cone or Lorentz cone from theoretical physics.

The dual cone and the projector of an ice cream cone are known explicitly.

**Fact 2.3.** Suppose that $\alpha > 0$ and $(x, r) \in \mathbb{X} \times \mathbb{R}$. Then, $\text{ice}^{\oplus}(\alpha) = \text{ice}(1/\alpha)$, and

$$P_{\text{ice}(\alpha)}(x, r) = \begin{cases} (x, r), & \text{if } \|x\| \leq \alpha r, \\ (0, 0), & \text{if } \alpha\|x\| \leq -r, \\ [(\alpha\|x\| + r)/(\alpha^2 + 1)](\alpha\, x/\|x\|, 1), & \text{otherwise.} \end{cases}$$

**Proof.** See Ref. 1, Theorem 3.3.6. □

**Corollary 2.1.** Suppose that $\alpha > 0$. Then $\text{ice}(\alpha)$ is obtuse $\Leftrightarrow$ $\alpha \geq 1$; $\text{ice}(\alpha)$ is self-dual $\Leftrightarrow$ $\alpha = 1$.

**Proof.** If $\beta > 0$, then $\text{ice}(\alpha) \subseteq \text{ice}(\beta) \Leftrightarrow \alpha \leq \beta$; this and Fact 2.3 readily yield the result. □

### 2.4. Reflector

**Definition 2.5.** Reflector. Suppose that $K$ is a closed convex set in $\mathbb{X}$. Then, the reflector corresponding to $K$ is defined by

$$R_K := 2P_K - I.$$

If $K$ is a cone and $x \in \mathbb{X}$, we write also

$$x^{++} := R_K x.$$

The following lemma collects various useful results on reflectors and obtuse cones.

**Lemma 2.1.** Suppose that $K$ is a closed convex cone in $\mathbb{X}$ and that $x$, $y$ are two points in $\mathbb{X}$. Then:

(i) $x^{++} = x^+ - x^-$.
(ii) $\|x - y\|^2 - \|x^{++} - y^{++}\|^2 = 4\langle x^+, -y^- \rangle + 4\langle y^+, -x^- \rangle \geq 0$.
(iii) The reflector $R_K$ is nonexpansive, $\|x^{++} - y^{++}\| \leq \|x - y\|$.
(iv) If $y \in K$, then $y^{++} = y$ and $\|x - y\| \geq \|x^{++} - y\|$.
(v) $K$ is obtuse if and only if $R_K$ maps $\mathbb{X}$ onto $K$.

**Proof.** In view of Fact 2.2, we have $x = x^+ + x^-$. Now, $x^- \in K^{\ominus}$; hence, $-x^- \in -K^{\ominus} = K^{\oplus}$; similarly, $-y^- \in K^{\oplus}$.

(i) $x^{++} = R_K x = (2P_K - I)(x) = 2x^+ - (x^+ + x^-) = x^+ - x^-$.

(ii) Using Fact 2.2, $\{x^+, y^+\} \subseteq K$ and $\{-x^-, -y^-\} \subseteq K^{\oplus}$, we obtain

$$\begin{aligned} &\|x-y\|^2 - \|x^{++} - y^{++}\| \\ &= \|(x^+ + x^-) - (y^+ + y^-)\|^2 - \|(x^+ - x^-) - (y^+ - y^-)\|^2 \\ &= \|(x^+ - y^+) + (x^- - y^-)\|^2 - \|(x^+ - y^+) - (x^- - y^-)\|^2 \\ &= 4\langle x^+ - y^+, x^- - y^- \rangle \\ &= 4\langle x^+, -y^- \rangle + 4\langle y^+, -x^- \rangle \\ &\geq 0. \end{aligned}$$

(iii) This is immediate from (ii).

(iv) If $y \in K$, then $y = P_K y$ and hence $y^+ = y$ and $y^- = 0$. By (i), $y^{++} = y^+ - y^- = y$. The result now follows from (iii).

(v) ($\Rightarrow$) By assumption on $K$, we have $x^{++} = x^+ + (-x^-) \in K + K^{\oplus} \subseteq K + K = K$.
($\Leftarrow$) Fix $x \in K^{\ominus}$. Then, $x^+ = 0$ and $x^- = x$. By assumption and (i), $x^{++} = x^+ - x^- \in K$; hence, $-x \in K$. Since $x$ was chosen arbitrarily in $K^{\ominus}$, we conclude that $-K^{\ominus} = K^{\oplus} \subseteq K$. □

**Remark 2.2.** The reflector $R_K$ [Lemma 2.1(iii)] is nonexpansive even when $K$ is merely assumed to be a closed convex nonempty set. The reason is that $P_K$ is firmly nonexpansive $\Leftrightarrow$ $R_K$ is nonexpansive; see, for instance, Ref. 9, Theorem 12.1. Let $R_K$ be the reflector corresponding to the nonnegative orthant in the Euclidean plane. Lemma 2.1(ii) shows not only that $R_K$ is nonexpansive, but it can be also used to demonstrate that $R_K$ does not satisfy any of the following stronger notions: strongly nonexpansive (Ref. 24); nonexpansive in the sense of De Pierro and Iusem (Ref. 25); firmly nonexpansive (Ref. 2); averaging (Ref. 2); strongly attracting (Ref. 2); attracting (Ref. 2).

It is this lack of additional good properties in the sense of nonexpansive mappings that makes the analysis of the method of reflection-projection impossible within standard frameworks.

### 2.5. Fejér Monotone Sequences

**Definition 2.6.** Suppose that $S$ is a closed convex nonempty set in $\mathbb{X}$ and that $(y_k)_{k \geq 0}$ is a sequence in $\mathbb{X}$. Then, $(y_k)$ is Fejér monotone with respect to $S$ if

$$\|y_{k+1} - s\| \leq \|y_k - s\|, \qquad \forall k \geq 0,\ \forall s \in S.$$

Fejér monotone sequences are very useful in the analysis of optimization algorithms; see for instance Refs. 2, 26, 27. We now record a selection of good properties that will be handy later.

**Fact 2.4.** Suppose that $S$ is a closed convex nonempty set in $\mathbb{X}$ and that $(y_k)_{k\geq 0}$ is Fejér monotone with respect to $S$. Then:

(i) $(y_k)$ is a bounded sequence.
(ii) $(d(y_k, S))$ is decreasing and nonnegative, hence convergent.
(iii) The sequence $(P_S y_k)$ converges to some point $\bar{s}\in S$.
(iv) $(y_k)$ converges to $\bar{s}$ if and only if all cluster points of $(y_k)$ belongs to $S$.

**Proof.** See, for instance, Refs. 2, 26, 27. □

## 3. Method of Reflection-Projection

The method of reflection-projection is formally expressed by Algorithm A1.

**Theorem 3.1.** Suppose that $C\neq\emptyset$ and $K$ is obtuse. Let $x_0\in\mathbb{X}$. Then, the sequence $(x_k)$ generated by Algorithm A1 converges to a point in $C$.

**Proof.** We proceed in several steps. Let

$$(y_k) := (x_0, x_0^{++}, P_1 x_0^{++}, \ldots, x_1, x_1^{++}, \ldots),$$

i.e., the sequence implicit in the generation of the sequence $(x_k)$ with all the intermediate terms.

Step 1. $(y_k)$ is Fejér monotone with respect to $C$. The reflector $R_K$ is nonexpansive [Lemma 2.1(ii)], and so are the projections $P_1, \ldots, P_N$ (Fact 2.1); moreover, the intersection of the fixed-point sets of these $N+1$ maps is precisely $C$. It follows that $(y_k)$ is Fejér monotone with respect to $C$.

Step 2. $(x_k^{++})$ is contained in $K$ and each $d(x_k^{++}, C_i)\to 0$. Since $K$ is obtuse, Lemma 2.1(v) implies that $(x_k^{++})$ lies entirely in $K$. Next, apply the firm nonexpansiveness of $P_1$ to the two points $x_k^{++}$, $P_C x_k^{++}$ to obtain

$$\|x_k^{++} - P_C x_k^{++}\|^2 \geq \|P_1 x_k^{++} - P_C x_k^{++}\|^2 + \|x_k^{++} - P_1 x_k^{++}\|^2.$$

This, Step 1, and Fact 2.4(ii) yield

$$d^2(x_k^{++}, C_1) \le d^2(x_k^{++}, C) - d^2(P_1 x_k^{++}, C) \to 0.$$

The firm nonexpansiveness of $P_2$ applied to the two points $P_1 x_k^{++}$, $P_C P_1 x_k^{++}$ results analogously in

$$d^2(P_1 x_k^{++}, C_2) \le d^2(P_1 x_k^{++}, C) - d^2(P_2 P_1 x_k^{++}, C) \to 0.$$

Continuing in this fashion yields $N-2$ further results, the last of which states that

$$d^2(P_{N-1} \cdots P_1 x_n^{++}, C_N) \le d^2(P_{N-1} \cdots P_1 x_k^{++}, C) - d^2(x_{k+1}, C) \to 0.$$

In particular,

(1) $x_k^{++} - P_1 x_k^{++} \to 0$,
(2) $P_1 x_k^{++} - P_2 P_1 x_k^{++} \to 0$,
(3) $P_2 P_1 x_k^{++} - P_3 P_2 P_1 x_k^{++} \to 0, \ldots,$
($N$) $P_{N-1} \cdots P_1 x_k^{++} - x_{k+1} \to 0$.

Now, fix $\nu \in \{1, \ldots, N\}$. Summing the null sequences of items (1) to ($\nu$), followed by telescoping and taking the norm, yields

$$0 \le d(x_k^{++}, C_\nu) \le \| x_k^{++} - P_\nu \ldots P_1 x_k^{++} \| \to 0.$$

Since $\nu$ was chosen arbitrarily, we have completed the proof of Step 2.

Step 3. Each cluster point of $(x_k^{++})$ lies in $C$. This is clear from Step 2 and the continuity of each distance function $d(\cdot, C_i)$.

Step 4. $(x_k^{++})$ converges to some point $\bar{c} \in C$. On the one hand, by Step 1, the sequence $(x_k^{++})$ is Fejér monotone with respect to $C$. On the other hand, by Step 3, all cluster points of $(x_k^{++})$ belong to $C$. Using Fact 2.4(iv), we conclude altogether that $(x_k^{++})$ converges to some point in $C$.

Step 5. The entire sequence $(y_k)$ converges to $\bar{c}$. Using Step 4 and the continuity of $P_1$ yields the convergence of $(P_1 x_k^{++})$ to $\bar{c}$. Applying the continuity of $P_2, \ldots, P_N$ successively in this fashion, we conclude altogether that $(y_k)$ converges to $\bar{c}$.

Final Step. $(x_k)$ converges to $\bar{c}$. This is immediate from Step 5, since $(x_k)$ is a subsequence of $(y_k)$. □

**Remark 3.1.** Various comments on Theorem 3.1 are in order.

(i) Theorem 3.1 may be extended routinely in various directions by incorporating weights, relaxation, and extrapolation parameters as in Refs. 1–4, 6. However, rather than obtaining a somewhat more general version, we opted to present a setting that not only shows clearly the usefulness of obtuseness, but that works also quite well in practice on the sample problems that we investigated numerically: in fact, in Section 5.4, we compare the method of reflection-projection to relaxed projections (the numerical results presented there strongly support the practical usefulness of the proposed algorithm).

(ii) Similarly, Theorem 3.1 and its proof extend to general Hilbert spaces as follows: the sequence $(x_k)$ converges weakly to some point in $C$, provided that each projector $P_i$ is weakly continuous. Thus, the method of reflection-projection can be used to solve convex feasibility problems with an obtuse cone constraint along with affine constraints (for which the corresponding projections are indeed weakly continuous).

(iii) In Theorem 3.1, it is impossible to strengthen the conclusion to handle two or more obtuse cones via reflectors: indeed, consider two neighboring quadrants in the Euclidean plane. The sequence of alternating reflections will not converge if we fix a starting point in the interior of one quadrant.

(iv) The condition $C \neq \emptyset$ is essential, as the algorithm may fail to converge in its absence: consider the nonnegative orthant in $\mathbb{R}^2$ and the halfspace $\{(\rho_1, \rho_2) \in \mathbb{R}^2 : \rho_1 + \rho_2 \leq -1\}$. For $x_0 := (0, 1)$, the method of reflection-projection cycles indefinitely: $x_n \equiv (0, (-1)^n)$. See, however, Section 7 for some positive results on the inconsistent case.

(v) The method of reflection-projection creates a sequence that is Fejér monotone with respect to $C$; hence, it fits the Combettes framework (see Ref. 27, Algorithm 4.1). However, if one wants to derive Theorem 3.1 from Ref. 27, Theorem 4.3(i), then one would have to check that all weak cluster points lie in $C$, which is exactly the bulk of the work in the proof of Theorem 3.1 and hence is not advantageous.

## 4. Affine Subspace Projector

Throughout this short section, we assume that $\mathbb{X}$ and $\mathbb{Y}$ are Euclidean spaces and that $A$ is a linear operator from $\mathbb{X}$ to $\mathbb{Y}$. Because we work with finite-dimensional spaces, the operator $A$ is continuous and its range,

$$\operatorname{ran} A := \{Ax \in \mathbb{Y} : x \in \mathbb{X}\},$$

is closed. We first summarize the fundamental properties of the Moore-Penrose inverse, taken from Chapter II of the Groetsch monograph (Ref. 28).

**Fact 4.1.** Moore-Penrose Inverse. There exists a unique (continuous) linear operator $A^\dagger$ from $\mathbb{Y}$ to $\mathbb{X}$ with $AA^\dagger = P_{\operatorname{ran} A}$ and $A^\dagger A = P_{\operatorname{ran} A^\dagger}$. The operator $A^\dagger$ is called the Moore-Penrose inverse of $A$. Moreover, $\operatorname{ran} A^\dagger = \operatorname{ran} A^*$, and the Moore-Penrose inverse can be computed via

$$A^\dagger = A^*(AA^*)^\dagger = A^*(AA^*|_{\operatorname{ran} A})^{-1} = (A^*A|_{\operatorname{ran} A^*})^{-1}A^* = (A^*A)^\dagger A^*.$$

Fix, $b \in \mathbb{Y}$, not necessarily in the range of $A$, and let $b_0 := P_{\operatorname{ran} A}(b)$. Then, $b_0 \in \operatorname{ran} A$, and $S := \{x \in \mathbb{X} | Ax = b_0\}$ is an affine subspace of $\mathbb{X}$.

**Lemma 4.1.** Affine Subspace Projector. $P_S(x) = x - A^\dagger(Ax - b)$, for every $x \in \mathbb{X}$.

**Proof.** Pick $x \in \mathbb{X}$ and let

$$s := x - A^\dagger(Ax - b).$$

In view of Fact 2.1, we need to show that

$$\text{(i)} \quad s \in S, \qquad \text{(ii)} \quad \sup\langle S - s, x - s\rangle \le 0.$$

Now, using Fact 4.1, we have

$$\begin{aligned} As &= A(x - A^\dagger(Ax - b)) \\ &= Ax - AA^\dagger(Ax) + AA^\dagger(b) \\ &= Ax - P_{\operatorname{ran} A}(Ax) + P_{\operatorname{ran} A}(b) \\ &= Ax - Ax + b_0 = b_0. \end{aligned}$$

Hence, $s \in S$ and (i) holds. Since $s \in S$, we have

$$S = s + \ker A,$$

where

$$\ker A := \{x \in \mathbb{X}: Ax = 0\}$$

is the kernel of $A$. By Fact 4.1,

$$A^\dagger(Ax - b) \in \operatorname{ran} A^\dagger = \operatorname{ran} A^*.$$

Hence,

$$A^\dagger(Ax - b) \in (\ker A)^\perp.$$

This implies that

$$0 = \langle \ker A, A^\dagger(Ax - b)\rangle = \langle S - s, x - s\rangle.$$

Therefore, (ii) is verified and we are done. □

As an illustration, let us rederive the well-known formula for the projection onto a hyperplane.

**Example 4.1.** Hyperplane Projection. Suppose that $a \in \mathbb{X} \setminus \{0\}$ and $b \in \mathbb{R}$. Let

$$S = \{x \in \mathbb{X}: \langle a, x \rangle = b\}.$$

Then,

$$P_S(x) = x - [(\langle a, x \rangle - b)/\|a\|^2]a, \qquad \text{for every } x \in \mathbb{X}.$$

**Proof.** Let $\mathbb{Y} := \mathbb{R}$, and define $A: \mathbb{X} \to \mathbb{Y}$ by $Ax = \langle a, x \rangle$. It is easy to see that

$$A^\dagger(y) = [y/\|a\|^2]a.$$

The result now follows from Lemma 4.1. □

The following remark discusses the complications arising from considering two or more hyperplanes.

**Remark 4.1.** Gram Matrix. Let $\mathbb{Y} = \mathbb{R}^m$ and $a_1, a_2, \ldots, a_m$ be $m$ vectors in $\mathbb{X}$. This induces a linear operator $A: \mathbb{X} \to \mathbb{Y}: x \mapsto (\langle a_i, x \rangle)_{i=1}^m$. Unless $m = 1$, there is no closed form available for $A^\dagger$, unfortunately. Note that $AA^*$ maps $\mathbb{R}^m$ to itself. Hence, after fixing a basis and switching to coordinates, $AA^*$ is represented by a matrix $G \in \mathbb{R}^{m \times m}$. It is not hard to see that $G$ is the Gram matrix of the vectors $a_1, \ldots, a_m$; i.e., $G_{i,j}$, the $(i, j)$-entry of $G$, is equal $\langle a_i, a_j \rangle$, the inner product of the vectors $a_i$ and $a_j$. Fact 4.1 results in

$$A^\dagger = A^*(AA^*)^\dagger = A^* G^\dagger.$$

Thus, finding the Moore-Penrose inverse of $A$ essentially boils down to computing the Moore-Penrose inverse of the Gram matrix $G$. See also the end of Chapter 8 in Deutsch's recent monograph (Ref. 29).

**Remark 4.2.** Everything that we recorded in this section holds true provided that $\mathbb{X}$ and $\mathbb{Y}$ are Hilbert spaces and provided that $A$ has closed range. This is so because all results cited from Ref. 28 hold in this setting. For various algorithms on computing the Moore-Penrose inverse, see Ref. 28, Sections 3–5 in Chapter II.

## 5. Nonnegative Orthant

Throughout this section, we assume that $\mathbb{X} := \mathbb{R}^n$ and that the obtuse cone is simply the positive orthant,

$$K := \mathbb{R}^n_+ := \{x \in \mathbb{R}^n : x_i \geq 0, \forall i\}.$$

It is easy to see that $K$ is a closed convex self-dual cone. We consider an additional affine constraint, derived as follows. Let $A \in \mathbb{R}^{m \times n}$ represent a linear operator from $\mathbb{X}$ to $\mathbb{Y} := \mathbb{R}^m$. Suppose that $b \in \mathbb{Y}$ and define

$$L := \{x \in \mathbb{X} : Ax = b\}.$$

Our interest concerns the basic two-set convex feasibility problem

$$\text{find} \quad x \in C := K \cap L.$$

This feasibility problem is of fundamental importance in various areas of mathematics, including medical imaging (Ref. 3).

**5.1. Implementation of Cone Projector and Reflector.** The projection onto the cone $K$ is simply

$$(P_K x)_i = x_i^+ = \max\{x_i, 0\}, \qquad \text{for every } i \in \{1, \ldots, n\};$$

thus the reflector

$$R_K = 2P_K - I$$

is given by

$$(R_K x)_i = |x_i| = \max\{x_i, -x_i\} = \text{abs}(x_i).$$

**5.2. Implementation of the Affine Subspace Projector.** Lemma 4.1 gives us a handle on computing $P_L$; the key step is to find an efficient and robust representation of $A^\dagger$, the Moore-Penrose inverse of $A$. Let $r = \text{rank}(A)$. We assume without loss of generality that $r \geq 1$. If $r = 0$, then $A = 0$ and hence $A^\dagger = 0 \in \mathbb{R}^{n \times m}$.

**Computation of $A^\dagger$ via QR Factorization.** Reference 30, Algorithm 5.4.1 describes an efficient implementation of the following factorization of $A^*$:

$$A^* = [Q, Q_0] \begin{bmatrix} R & D \\ 0 & 0 \end{bmatrix} P^*,$$

where both $Q \in \mathbb{R}^{n \times r}$ and $Q_0 \in \mathbb{R}^{n \times (n-r)}$ have orthonormal columns, $R \in \mathbb{R}^{r \times r}$ is upper triangular and of full rank, $D \in \mathbb{R}^{r \times (m-r)}$, and finally $P \in \mathbb{R}^{m \times m}$ is a permutation matrix, hence orthogonal. In practice, $Q_0$ is not computed, since

$$A^*P = Q[R, D].$$

By Fact 4.1, we have

$$\begin{aligned} A^\dagger &= A^*(AA^*)^\dagger \\ &= Q[R\,D]P^*\left[P\begin{bmatrix} R^* \\ \\ D^* \end{bmatrix} Q^*Q[R\,D]P^*\right]^\dagger \\ &= Q[R\,D]P^*\left[P\begin{bmatrix} R^*R & R^*D \\ \\ D^*R & D^*D \end{bmatrix} P^*\right]^\dagger. \end{aligned}$$

**Implementation of $P_L$ via QR Factors.** After permuting the rows of $[A, b]$ according to $P^*$ and removing redundant constraints if necessary, we assume without loss of generality that $P^* = I$ and $r = m \le n$. Then, $D$ disappears altogether and the previous expression for $A^\dagger$ simplifies to

$$A^\dagger = QR^{-*}.$$

We can detect whether $L$ is nonempty by comparing the column rank of $[A, b]$ to $r$. Assuming $L \neq \emptyset$ and utilizing Lemma 4.1, the projection of $x \in \mathbb{X}$ onto $L$ now becomes

$$\begin{aligned} P_L(x) &= x - A^\dagger(Ax - b) \\ &= x - QR^{-*}(Ax - b). \end{aligned}$$

In large-scale applications, one needs to store $Q$ and $R$ in compact form using Householder reflections; see Ref. 30, Section 5.2.1 for further information. Since $A = R^*Q^*$, this projection can be expressed also as

$$\begin{aligned} P_L(x) &= x - QR^{-*}(Ax - b) \\ &= x - QQ^*x + QR^{-*}b \\ &= (I - QQ^*)x + QR^{-*}b; \end{aligned}$$

however, especially when $A$ is sparse, this is not preferable in terms of cost or robustness because of the term involving $QQ^*$.

**5.3. Complete Implementation of the Algorithm.** We present now the method of reflection-projection (Algorithm A2) for an affine constraint and the nonnegative orthant, based on the material developed earlier in this section. An important feature is that we allow arbitrary input data $[A, b]$

with no restrictions on the relative size of the matrix $A \in \mathbb{R}^{m \times n}$ or on its rank. We compare first the (numerical) ranks of $A$ and $[A, b]$, to determine whether $L$ is nonempty. If it is, then the constraints are permuted and redundant constraints are removed.

The termination criteria in the actual implementation go somewhat further than the abstract formulation of Algorithm A1: Based on the analysis in Section 7, a heuristic attempt to detect a possible inconsistency of the feasibility problem, i.e., $C = K \cap L = \emptyset$ even though $L \neq \emptyset$. To deal with inconsistency, a stopping criterion based on the analysis in Section 7 is proposed; see Remark 7.5.

After the initial cost of the QR factorization, which is $O(n^3)$ in the dense matrix case (Ref. 30, Section 5.2.1), each iteration is fast, since the cost of a triangular solve and a matrix-vector multiplication is only $O(n^2)$; see Ref. 30, Section 3.1.

**5.4. Numerical Experiments.** Complete implementations of the algorithms described below, for both Octave[5] and Matlab[6] are available at Ref. 31. These implementations were used to generate all the numerical results in this paper.

To highlight some advantages of using the method of reflection-projection over (relaxed) projections, we devised an experiment whose details and results are illustrated now. We generated a large set of random feasible problems of varying size. On these problems, we ran a sequence of alternating relaxed projection algorithms. More precisely, representing the relaxed projection onto the cone by

$$x \mapsto ((1 - \alpha_K)I + \alpha_K P_K)x, \quad \alpha_K \in (0, 2],$$

and the relaxed projection onto the flat by

$$x \mapsto ((1 - \alpha_L)I + \alpha_L P_L)x, \quad \alpha_L \in (0, 2),$$

we measured the performance of iterating the map

$$((1 - \alpha_L)I + \alpha_L P_L)((1 - \alpha_K)I + \alpha_K P_K)$$

for a fixed pair of relaxation parameters $(\alpha_K, \alpha_L) \in (0, 2] \times (0, 2)$.

If the relaxation parameter is equal to 1, then the relaxed projection is actually an exact projection; similarly, if it is equal to 2, then we obtain a reflection. Thus, the method of alternating projections corresponds to the choice $(\alpha_K, \alpha_L) = (1, 1)$, whereas the new method of reflection-projection is obtained by setting $(\alpha_K, \alpha_L) = (2, 1)$. If $\alpha_K < 2$, then the iterates are known to

[5]Octave is freely redistributable software available at www.octave.org.
[6]Matlab is a registered trademark of The MathWorks, Inc.

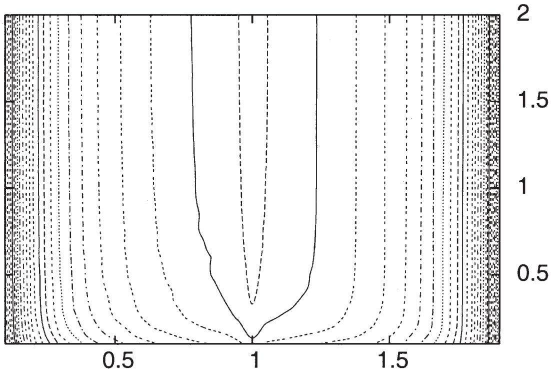


Fig. 2. Iteration count behavior for variations in the flat parameter $0<\alpha_L<2$ and the cone parameter $0<\alpha_k\leq 2$.

converge; see Refs. 1, 2, 4, 6. And if $\alpha_K=2$ but $\alpha_L\neq 1$, then a convergence result can be obtained easily by a straightforward modification of the proof of Theorem 3.1; see also Remark 3.1(i).

We searched experimentally for the optimal relaxation parameters by generating a set of several thousand problems, with fixed $n=64$ and various $m$ in the range $2<m<62$. We solved each of these problems by varying the relaxation parameters on the interval $[0, 2]$ in small increments. Figure 2 displays the average number of iterations for the problems, for every combination of the relaxation parameters. This experiment suggests that the optimal strategy is to project exactly on the flat ($\alpha_L=1$), but to reflect into the cone ($\alpha_K=2$): this corresponds precisely to the method of reflection-projection (Algorithm A2). Of course, a different set of problems may suggest a different combination of the relaxation parameters.

## 6. Positive-Semidefinite Cone

In this section, we consider the Euclidean space of all real symmetric $n\times n$ matrices,

$$\mathbb{X}:=\mathbb{S}^n:=\{X\in\mathbb{R}^{n\times n}: X=X^*\},$$

with

$$\langle X, Y\rangle := \text{trace}(XY), \qquad \text{for } X,\ Y\in\mathbb{X}.$$

For $X\in\mathbb{X}$, we write $X\succeq 0$ to indicate that $X$ is positive semidefinite, and we collect all such matrices in the set

$$K:=\mathbb{S}^n_+:=\{X\in\mathbb{X}: X\succeq 0\}.$$

The Fejér theorem states that $K$ is a closed convex self-dual cone; see Ref. 32, Corollary 7.5.4. This setting lies at the heart of modern optimization; see for instance Refs. 33, 34.

Building upon Section 5, we consider an affine constraint given by finitely many linearly independent vectors $A_1, \ldots, A_m$ in $\mathbb{X}$ and a vector $b \in \mathbb{R}^m$. Linear independence may be enforced as described in our discussion of Algorithm A2 in Section 5.3. Our assumption is equivalent to the surjectivity of the operator

$$A: \mathbb{X} \to \mathbb{R}^m: X \mapsto [\langle A_1, X\rangle \langle A_2, X\rangle \ldots \langle A_m, X\rangle]^*.$$

Hence,

$$L := \{X \in \mathbb{X}: A(X) = b\} \neq \emptyset$$

and we aim to solve the two-set feasibility problem

$$\text{find} \quad X \in C := K \cap L.$$

For simplicity, we assume consistency, i.e., $C \neq \emptyset$. The inconsistent case is quite subtle: the two constraints may have no points in common, yet their gap may be zero (see Remark 7.4); such behavior is impossible for the (polyhedral) setting of the previous Section 5.

The motivation for considering this problem stems from the interior-point methods for solving semidefinite programming problems. These algorithms fall into two disjoint classes: the so-called infeasible methods (which do not require a feasible starting point) and the feasible methods. The feasible starting point required for algorithms of the latter class is precisely a solution of the above feasibility problem. See Refs. 34, 35 and references therein for further details and background on semidefinite programming. As we illustrate in this section, the method of reflection-projection is well-suited to find such a feasible starting point.

**6.1. Implementation of the Cone Projector and Reflector.** Fix an arbitrary $X \in \mathbb{X}$. Since $X$ is symmetric, we can factor $X = U^* D U$, where $U$ is an orthogonal matrix whose columns are the eigenvectors $u_1, \ldots, u_n$ of $X$ and $D$ is a diagonal matrix whose diagonal entries $\lambda_i$ are the corresponding eigenvalues:

$$\lambda_i u_i = X u_i, \qquad \text{for all } i.$$

Denote by $D^+$ the diagonal matrix in $\mathbb{X}$ with

$$(D^+)_{i,i} = (\lambda_i)^+ = \max\{\lambda_i, 0\}.$$

Using the complete eigenvalue decomposition (see e.g. Ref. 30, Section 5.5.4), we have

$$P_K(X) = U^* D^+ U;$$

equivalently,

$$P_K(X) = \sum_{i:\lambda_i > 0} \lambda_i u_i u_i^*.$$

The bulk of the work in computing the projection $P_K(X)$ or the reflection

$$R_K(X) = 2P_K(X) - X$$

lies thus in the determination of the eigenvalues and eigenvectors of $X$.

The eigendecomposition of a symmetric matrix is an intricate but well-studied problem, and algorithms have been developed for which code is (sometimes freely) available. We refer the reader to the classical Ref. 36 and to the more recent treatment in Ref. 37. Note that, in order to compute $P_K(X)$, we do not need the complete decomposition; rather, the eigenpairs corresponding to either the positive or negative eigenvalues are sufficient. For actual numerical implementations, a Lanczos or Arnoldi process appears to be most appropriate, especially for large sparse matrices (Ref. 37). From now on, we consider the decomposition as a given black-box routine.

**6.2. Implementation of the Affine Subspace Projector.** The space of real symmetric $n \times n$ matrices is a proper subspace of the space of real $n \times n$ matrices: $\mathbb{S}^n \subsetneq \mathbb{R}^{n\times n}$; in fact, the dimension of $\mathbb{X} = \mathbb{S}^n$ is

$$t(n) := 1 + 2 + \cdots + n = n(n+1)/2,$$

the $n$th triangular number. From a numerical point of view, it is much faster and more memory efficient to work with corresponding vectors in $\mathbb{R}^{t(n)}$ rather than with (symmetric and hence redundant) matrices in $\mathbb{R}^{n\times n}$. Consequently, we start by describing the isometry

$$\text{svec}: \mathbb{S}^n \to \mathbb{R}^{t(n)},$$

which takes the first $i$ entries in column $i$, stacks them (proceeding from left to right) into a long column vector

$$[X_{1,1} \quad X_{1,2} \quad X_{2,2} \quad X_{1,3} \quad X_{2,3} \quad X_{3,3} \ldots X_{1,n} \quad X_{2,n} \ldots X_{n,n}]^*,$$

and finally multiplies each offdiagonal element by $\sqrt{2}$ to guarantee that the norm $\|X\|$, taken in $\mathbb{S}^n$, agrees with the norm $\|\text{svec}(X)\|$, taken in $\mathbb{R}^n$. More formally, define the following two index functions

$$\text{svecind}(i,j) := t(j-1) + i, \qquad 1 \le i \le j \le n,$$

and

$$\text{smatind}(k) := (k - j(j-1)/2, j),$$
$$j := \text{Ceiling of } [(1/2)(\sqrt{1+8k} - 1)] \quad \text{and} \quad 1 \le k \le t(n),$$

which are inverses of each other (Ref. 44). For $X \in \mathbb{S}^n$ and $1 \le k \le t(n)$, the isometry svec is described by

$$\text{svec}(X)_k = \begin{cases} X_{\text{smatind}(k)}, & \text{if } k \text{ is triangular;} \\ \sqrt{2} X_{\text{smatind}(k)}, & \text{otherwise.} \end{cases}$$

And, for $x \in \mathbb{R}^{t(n)}$ and $1 \le i, j \le n$, the inverse smat of svec is given explicitly by

$$\text{smat}(x)_{i,j} = \begin{cases} (1/\sqrt{2}) x_{\text{svecind}(i,j)}, & \text{if } i < j, \\ x_{\text{svecind}(i,j)}, & \text{if } i = j, \\ (1/\sqrt{2}) x_{\text{svecind}(j,i)}, & \text{if } i > j. \end{cases}$$

Now, define the $m \times t(n)$ matrix

$$\text{sop}(A) := [(\text{svec}(A_1))^* (\text{svec}(A_2))^* \ldots (\text{svec}(A_m))^*]^*.$$

For $X \in \mathbb{X}$, the affine constraint in $\mathbb{S}^n$ is thus reformulated equivalently in $\mathbb{R}^{t(n)}$ by

$$A(X) = b \quad \Leftrightarrow \quad [\text{sop}(A)]\text{svec}(X) = b.$$

Therefore, the computation of $P_L(X)$ is reduced to the case considered previously in Section 5.2.

**6.3. Formulation of the Algorithm.** For the sake of brevity, we shall omit the steps dealing with infeasibility and redundancy, since they are similar to those taken at the beginning of Algorithm A2 (see also Ref. 31).

Algorithm A3 produces sequences denoted $(X_k)$ and $(X_k^{++})$ and representing the successive iterates onto the flat and into the cone,

$$X_0^{++} \overset{P_L}{\mapsto} X_1 \overset{R_K}{\mapsto} X_1^{++} \overset{P_L}{\mapsto} X_2 \overset{R_K}{\mapsto} X_2^{++} \overset{P_L}{\mapsto} X_3 \overset{R_K}{\mapsto} \cdots,$$

where the loop invariant maintains the iterates within the positive definite cone. Therefore, the termination criterion involves only the distance to the flat. The reason for exiting the inner loop after a reflection is that the solution returned by the algorithm is numerically positive definite in most cases. This may be useful when strictly interior solutions are sought as is the case for feasible interior-point algorithms of semidefinite programming.

Table 1. Experiment on random semidefinite feasibility problems.

| $m$ | $\|x-P_L(x)\|$ | Iter | $m$ | $\|x-P_L(x)\|$ | Iter |
|---|---|---|---|---|---|
| 1 | 1.000799e-16 | 5 | 17 | 9.856825e-07 | 386 |
| 2 | 9.279782e-16 | 30 | 18 | 9.878404e-07 | 428 |
| 3 | 8.530127e-07 | 54 | 19 | 9.874863e-07 | 454 |
| 4 | 8.917017e-07 | 75 | 20 | 9.621321e-07 | 495 |
| 5 | 9.767166e-07 | 96 | 21 | 9.959647e-07 | 531 |
| 6 | 9.867892e-07 | 118 | 22 | 9.912171e-07 | 575 |
| 7 | 9.977883e-07 | 134 | 23 | 9.950764e-07 | 627 |
| 8 | 9.543684e-07 | 158 | 24 | 9.826544e-07 | 790 |
| 9 | 9.400719e-07 | 175 | 25 | 9.918322e-07 | 748 |
| 10 | 9.402718e-07 | 197 | 26 | 9.981378e-07 | 883 |
| 11 | 9.411396e-07 | 217 | 27 | 9.989317e-07 | 931 |
| 12 | 9.500699e-07 | 246 | 28 | 9.938861e-07 | 961 |
| 13 | 9.707506e-07 | 263 | 29 | 9.928170e-07 | 1079 |
| 14 | 9.703939e-07 | 296 | 30 | 9.959961e-07 | 1327 |
| 15 | 9.802466e-07 | 319 | 31 | 9.897021e-07 | 1406 |
| 16 | 9.923299e-07 | 351 | 32 | 9.973459e-07 | 1864 |

**6.4. Numerical Experiments.** Table 1 presents the results of Algorithm A3 on random semidefinite problems; the code used to generate these data is available at Ref. 31. Feasibility tolerance was set to $10^{-5}$. The actual distance to the flat is indicated by the column $\|x-P_L(x)\|$. We do not indicate the distance to the cone, since this is always 0. For these problems, we set $n=15$ and the entries in Table 1 average the number of iterations for each size after 40 runs. The reader will notice that the algorithm performs very well if the number of constraints is low, but the performance then degrades as the feasible set gets smaller.

## 7. Results for the Inconsistent Case

In this section, we discuss the behavior of the algorithm for two possibly nonintersecting constraints. The reason for this restriction is this: even for the mathematically easier case of cyclic projections, the geometry and behavior is only fully understood for two sets; see Ref. 38 for a survey. On the other hand, the results of this section do hold for two general closed convex sets, i.e., neither is assumed to be an obtuse cone. Nonintersecting constraints occur frequently in practical problems, due to measurement or roundoff errors. Hence, it is crucial to understand at least partially the behavior of the method of reflection-projection in the inconsistent case.

We start by reviewing the geometry of the problem, which is independent of the algorithm under consideration. For the rest of this section, we assume that $A$ and $B$ are two closed convex nonempty sets in $\mathbb{X}$. We let

$$v := P_{\mathrm{cl}(B-A)}(0) \quad \text{and} \quad \delta := \|v\| = \inf \|A - B\|$$

be the gap between $A$ and $B$. Here $\mathrm{cl}(B-A)$ denotes the closure of the Minkowski difference

$$B - A := \{b - a : a \in A, b \in B\}.$$

We collect the points in $A$ and $B$ where the gap is attained in the following sets:

$$E := \{a \in A : d(a, B) = \delta\},$$
$$F := \{b \in B : d(b, A) = \delta\}.$$

The sets $E$ and $F$ thus generalize the idea of the intersection of the two sets $A$ and $B$. Note, however, that $E$ and $F$ may be empty: consider in the Euclidean plane the horizontal axis and the epigraph of $\rho \mapsto 1/\rho$, i.e.,

$$\{(\rho_1, \rho_2) \in \mathbb{R}^2 : 0 < 1/\rho_1 \leq \rho_2\}.$$

**Definition 7.1.** Set of Fixed Points. If $T: \mathbb{X} \to \mathbb{X}$ is a map, then

$$\mathrm{Fix}(T) = \{x \in \mathbb{X} : T(x) = x\}$$

denotes the set of fixed points of $T$.

**Fact 7.1.**
(i) $E = \mathrm{Fix}(P_A P_B)$ and $F = \mathrm{Fix}(P_B P_A)$.
(ii) $E + v = F$, $E = A \cap (B - v)$, and $F = (A + v) \cap B$.
(iii) Suppose that $e \in E$ and $f \in F$. Then, $P_B e = e + v$ and $P_A f = f - v$.

**Proof.** For (i), see Ref. 39; for (ii) and (iii), see Ref. 40. □

The method of reflection-projection consists of computing the iterates of the maps $P_A R_B$ and $R_B P_A$; consequently, we are interested in the fixed-point sets of these compositions.

**Lemma 7.1.** $\mathrm{Fix}(P_A R_B) = E$ and $\mathrm{Fix}(R_B P_A) = F + v$.

**Proof.** Let $(\bar{a}, \bar{b})$ be a fixed point pair,

$$\bar{a} = P_A \bar{b} \quad \text{and} \quad \bar{b} = R_B \bar{a}.$$

Fix $a \in A$ and $b \in B$ arbitrarily. Then, using Fact 2.1, $\bar{a} = P_A \bar{b}$ is characterized by

$$\bar{a} \in A \quad \text{and} \quad \langle a - \bar{a}, \bar{b} - \bar{a} \rangle \leq 0.$$

Similarly, $\bar{b} = R_B \bar{a}$ is equivalent to

$$(\bar{a} + \bar{b})/2 \in B \quad \text{and} \quad \langle b - (\bar{a} + \bar{b})/2, \bar{a} - \bar{b} \rangle \leq 0.$$

Adding the inequalities yields

$$\langle b - a - ((\bar{a} + \bar{b})/2 - \bar{a}), \bar{a} - \bar{b} \rangle \leq 0,$$

which shows that

$$(\bar{a} + \bar{b})/2 - \bar{a} = P_{\mathrm{cl}(B-A)}(0) = v.$$

Fact 7.1 now shows that

$$\bar{a} \in E \quad \text{and} \quad \bar{b} = R_B \bar{a} = \bar{a} + 2v \in F + v.$$

Hence,

$$\mathrm{Fix}(P_A R_B) \subseteq E \quad \text{and} \quad \mathrm{Fix}(R_B P_A) \subseteq F + v.$$

The reverse inclusions are shown similarly, using once again Fact 7.1. □

**Remark 7.1.** Viewed from fixed-point theory, the compositions $R_B P_A$ and $P_A R_B$ are nonexpansive maps with little extra structure. For instance, they lack asymptotic regularity: indeed, in $\mathbb{X} = \mathbb{R}^2$, let $B$ be the nonnegative orthant and let $A$ be the line $\{(\rho, -\rho - 1): \rho \in \mathbb{R}\}$. Let $a_0 = (0, -1)$ be the starting point for the sequence $a_k := P_A R_B a_{k-1}$ generated by the method of reflection-projection. Then, the orbit $(a_k)$ consists of two distinct subsequences,

$$a_{2k} = (0, -1) \quad \text{and} \quad a_{2k+1} = (-1, 0).$$

Hence, $(a_k)$ does not converge even though the distance between the two sets is uniquely attained at $(-1/2, -1/2) \in A$ and $(0, 0) \in B$. In particular, $a_k - a_{k+1} \not\to 0$, which means that the sequence $(a_k)$ is not asymptotically regular.

On the other hand, we have the following positive result.

**Lemma 7.2.** Let $b_k := R_B a_k$ and $a_{k+1} := P_A b_k$ be the sequence generated by the method of reflection-projection, with starting point $a_0$. Then,

$$\|a_k - a\|^2 \geq \|a_{k+1} - P_A R_B a\|^2 + \|(b_k - a_{k+1}) - (R_B a - P_A R_B a)\|^2, \qquad \forall a \in A.$$

Furthermore, if the gap between $A$ and $B$ is realized, i.e., $E \neq \emptyset$, then:

(i) $\|a_k - e\|^2 \geq \|a_{k+1} - e\|^2 + \|(b_k - a_{k+1}) - 2v\|^2$, $\forall e \in E$. In particular, $(a_k)$ is Fejér monotone with respect to $E$.
(ii) $b_k - a_{k+1} \to 2v$.
(iii) Every cluster point of $(a_k + a_{k+1})/2$ belongs to $E$.
(iv) Every cluster point of $(P_B a_k)$ belongs to $F$.

**Proof.** The inequality follows, since $R_B$ is nonexpansive and $P_A$ is firmly nonexpansive.

(i) This is a special case of the inequality, since

$$R_B e = e + 2v \quad \text{and} \quad P_A R_B e = e;$$

see Lemma 7.1 and Fact 7.1.

(ii) This is clear from (i).

(iii) and (iv). Observe that (ii) is equivalent to $P_B a_k - (a_k + a_{k+1})/2 \to v$.

Now, the first term in this difference belongs to $B$ and the second one to $A$. The result now follows directly from Ref. 40, Lemma 2.3. □

**Remark 7.2.** We do not know whether the sequence of averages $((a_k + a_{k+1})/2)$ must converge to a point in $E$. This will happen if $E$ is a singleton, as is the case in the example presented in Remark 7.1.

**Remark 7.3.** Lack of Monotonicity. In the Euclidean plane, let

$$A := \{(\rho, -3\rho - 3) : \rho \in \mathbb{R}\}$$

and let $B$ be the nonnegative orthant. Further, set

$$a_0 := (0, -3) \in A$$

and define recursively

$$b_n := R_B(a_n) \text{ and } a_{n+1} := P_A(b_n), \qquad \text{for } n \geq 0.$$

It is easy to see that

$$\|a_1 - b_0\| < \|a_2 - b_1\|.$$

Hence the sequence $(\|a_{k+1} - b_k\|)$ is not decreasing. However, monotonicity properties of this kind lie at the heart of the analysis of the method of cyclic projections and also of the Dykstra algorithm; see Ref. 40, Lemma 4.4(ii) and Lemma 3.1(iv). The lack of this type of monotonicity appears to make the analysis of the inconsistent case much more difficult.

**Remark 7.4.** Attainment versus Nonattainment. Whether or not the gap between the constraints $A$ and $B$ is realized depends essentially on the relative geometry of the sets. Some sufficient conditions for attainment are discussed in Ref. 40, Section 5. Perhaps, the most important case in applications occurs when one constraint is affine and the other is either the nonnegative orthant or the cone of positive-definite matrices. We explicitly record the following.

(i) If $A$ is affine and $B$ is the nonnegative orthant, then the gap between $A$ and $B$ is always realized. The reason is that, if $A$ and $B$ are both polyhedral, then so is their difference; in particular, $B-A$ is closed. See Ref. 40, Facts 5.1(ii) and also Ref. 38 for additional information.

(ii) If $A$ is affine and $B$ is the cone of positive semidefinite matrices, then the gap between $A$ and $B$ need not be realized at a pair of points; see Refs. 41, 42 for concrete examples.

**Remark 7.5.** Stopping Criterion for Algorithm A2. In Section 5, two constraints are considered: adapting to the notation of this section, the obtuse cone is $B := \mathbb{R}^n_+$, and the other constraint $A$ is affine. Let

$$b_k := R_B a_k \quad \text{and} \quad a_{k+1} := P_A b_k,$$

for some starting point $a_0$ and let $k \geq 0$. This is precisely the iteration of Algorithm A2. In view of Remark 7.4(i), the gap between $A$ and $B$ is attained. Moreover, by Lemma 7.2(ii),

$$w_k := b_k - a_{k+1} \to 2v;$$

hence,

$$\|w_{k-1} - w_k\| / (1 + \|w_k\|) \to 0.$$

Consequently, Algorithm A2 will eventually stop even if $A$ and $B$ have no points in common. The final vector $w_k$ is an approximation of twice the gap vector.

The following remark is due to one of the referees.

**Remark 7.6.** Detecting Infeasibility. Consider the setting of Lemma 7.2. Assume that $A \cap B \neq \emptyset$ and that an upper bound $\rho$ of $d(a_0, A \cap B)$ is known. Lemma 7.2(i) and telescoping yields readily

$$\sum_k \|b_k - a_{k+1}\|^2 \leq d^2(a_0, A \cap B) \leq \rho^2.$$

Therefore, if one observes that a partial sum of $\sum_k \|b_k - a_{k+1}\|^2$ exceeds $\rho^2$, then one can infer that $A \cap B$ is empty.

## 8. Conclusions

We presented a new algorithm, the method of reflection-projection, for solving the convex feasibility problem. It aims to find a point in the intersection of finitely many closed convex sets, where one of these sets is an obtuse cone. The method is similar to cyclic projections, but it is falls outside the standard frameworks and hence requires a separate proof. Experimental results indicate better performance on some problem sets.

We have given detailed instructions for the feasibility problems involving affine constraints and either the nonnegative orthant or the positive semidefinite cone, both of which are of practical importance. In the former case, in addition to a convergence proof of the algorithm on consistent problems, we have a theoretical detection mechanism of inconsistency. This criterion leads to the implementation of an effective heuristics.

One particularity of the method of reflection-projection is that it yields easily a strictly positive solution in most instances if the last step of the inner loop is the reflection, as we have described and implemented. This is of particular interest in the case of the semidefinite feasibility problem, since the method may then be used as a preliminary phase for feasible interior-point methods of semidefinite programming. We are currently considering such a multiphase approach.

## 9. Appendix: Summary of Algorithms

**Algorithm A1.** Method of reflection-projection.

Step 0. Initialization. Starting point $x_0 \in \mathbb{X}$ and set $k = 0$.

Step 1. Termination. If distance of $x_k$ to each set $K, C_1, \ldots, C_N$ is 0, exit.

Step 2. Reflection into the Cone: $x_k^{++} := R_K x_k$.

Step 3. Cyclic Projections onto the Sets: $x_{k+1} := P_N P_{N-1} \cdots P_1 x_k^{++}$.

Step 4. Update the Iterate: Set $k := k + 1$ and go back to Step 1.

**Algorithm A2.** Method of reflection-projection for $\{x \in \mathbb{R}^n | Ax = b\} \cap \mathbb{R}^n_+$.

Step 0. Initialization. Obtain data $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$. Start $x_{\text{cone}} := \mathbb{1} \in \mathbb{R}^n$. Compute rank-revealing QR factorizations $[A, b]^* P_0 = Q_0 R_0$, and $A^* P = QR$. If the ranks differ, then quit; else, compute the index of nonzero diagonal elements $I := (|\text{diag}(R)| > 10^{-13} \|A\|)$ and permute $A := P(I)^* A$, $Q := Q(I)$, $R := R(I)$, $b := P(I)^* b$.

Step 1. Flat Projection. Compute the residual $r := Ax_{\text{cone}} - b$. Solve the triangular system $R^* y = r$. Compute the projection onto flat $x_{\text{flat}} := x_{\text{cone}} - Qy$ and the distance to flat $d := \|x_{\text{cone}} - x_{\text{flat}}\|$.

Step 2. Gap. Compute an estimate of the gap $w := x_{\text{cone}} - x_{\text{flat}}$.
Step 3. Cone Reflection. Compute $x_{\text{cone}} := \text{abs}(x_{\text{flat}})$.
Step 4. Termination. If distance to flat is below tolerance, exit with success. If $\|w - (x_{\text{cone}} - x_{\text{flat}})\|/(1 + \|w\|)$ is below tolerance exit with failure. Go back to Step 1.

**Algorithm A3.** Method of reflection-projection for $\{x \in \mathbb{S}^n | A(X) = b\} \cap \mathbb{S}^n_+$.
Step 0. Initialization. Given data $A_1, A_2, \ldots, A_m \in \mathbb{S}^n$, $b \in \mathbb{R}^m$. Factor $[\text{sop}(A)]^* = QR$ after handling infeasibility as in Algorithm A2. Set the initial point to the identity $X := I \in \mathbb{S}^n$; $x := \text{svec}(X)$.
Step 1. Flat Projection. Compute the residual $r := \text{sop}(A)x - b$. Solve the triangular system $R^* y = r$. Compute the projection onto flat $x_{\text{flat}} := x - Qy$, $\text{X} := \text{smat}(x_{\text{flat}})$, and distance to flat $d := \|x - x_{\text{flat}}\|$.
Step 2. Cone Reflection. Compute eigendecomposition $X = V \Lambda V^*$. Extract positive values $\Lambda^{++} := \text{abs}(\Lambda)$; compute reflection $X := V \Lambda^{++} V^*; x := \text{svec}(X)$.
Step 3. Termination. If distance to flat is below tolerance, exit with success. Go back to Step 1.

## References


1. Bauschke, H. H., *Projection Algorithms and Monotone Operators*, PhD Thesis, Simon Fraser University, 1996. Available at www.cecm.sfu.ca/preprints/1996pp.html as 96:080.
2. Bauschke, H. H., and Borwein, J. M., *On Projection Algorithms for Solving Convex Feasibility Problems*, SIAM Review, Vol. 38, pp. 367–426, 1996.
3. Censor, Y., and Zenios, S. A., *Parallel Optimization*, Oxford University Press, Oxford, UK, 1997.
4. Combettes, P. L., *Hilbertian Convex Feasibility Problem: Convergence of Projection Methods*, Applied Mathematics and Optimization, Vol. 35, pp. 311–330, 1997.
5. Crombez, G., *Parallel Algorithms for Finding Common Fixed Points of Paracontractions*, Numerical Functional Analysis and Optimization, Vol. 23, pp. 47–59, 2002.
6. Kiwiel, K. C., and Łopuch, B., *Surrogate Projection Methods for Finding Fixed Points of Firmly Nonexpansive Mappings*, SIAM Journal of Optimization, Vol. 7, pp. 1084–1102, 1997.
7. Motzkin, T. S., and Schoenberg, I. J., *The Relaxation Method for Linear Inequalities*, Canadian Journal of Mathematics, Vol. 6, pp. 393–404, 1954.

8. Cimmino, G., *Calcolo Approssimato per le Soluzioni dei Sistemi di Equazioni Lineari*, La Ricerca Scientifica ed il Progresso Tecnico nella Economia Nazionale (Roma), Vol. 1, pp. 326–333, 1938.
9. Goebel, K., and Kirk, W. A., *Topics in Metric Fixed-Point Theory*, Cambridge University Press, Cambridge, UK, 1990.
10. Zarantonello, E. H., *Projections on Convex Sets in Hilbert Space and Spectral Theory*, Contributions to Nonlinear Functional Analysis, Edited by E. H. Zarantonello, Academic Press, New York, NY, pp. 237–424, 1971.
11. Moreau, J. J., *Décomposition Orthogonale d'un Espace Hilbertien selon Deux Cônes Mutuellement Polaires*, Comptes Rendus des Séances de l'Académie des Sciences, Paris, Séries A–B, Vol. 255, pp. 238–240, 1962.
12. Rockafellar, R. T., *Convex Analysis*, Princeton University Press, Princeton, New Jersey, 1970.
13. Goffin, J. L., *The Relaxation Method for Solving Systems of Linear Inequalities*, Mathematics of Operations Research, Vol. 5, pp. 388–414, 1980.
14. Todd, M. J., *Some Remarks on the Relaxation Method for Linear Inequalities*, Technical Report 419, School of Operations Research and Industrial Engineering, Cornell University, 1979.
15. Cegielski, A., *Projection onto an Acute Cone and Convex Feasibility Problem*, System Modelling and Optimization (Compiègne, 1993); Lecture Notes in Control and Information Sciences, Edited by J. Henry and J. P. Yvon, Springer Verlag, London, UK, Vol. 197, pp. 187–194, 1994.
16. Cegielski, A., *A Method of Projection onto an Acute Cone with Level Control in Convex Minimization*, Mathematical Programming, Vol. 85, pp. 469–490, 1999.
17. Cegielski, A., *Obtuse Cones and Gram Matrices with Nonnegative Inverse*, Linear Algebra and Its Applications, Vol. 335, pp. 167–181, 2001.
18. Cegielski, A., and Dylewski, R., *Residual Selection in a Projection Method for Convex Minimization Problems*, Optimization, Vol. 52, pp. 211–220, 2003.
19. Kiwiel, K. C., *The Efficiency of Subgradient Projection Methods for Convex Optimization, Part 2: Implementations and Extensions*, SIAM Journal on Control and Optimization, Vol. 34, pp. 677–697, 1996.
20. Kiwiel, K. C., *Monotone Gram Matrices and Deepest Surrogate Inequalities in Accelerated Relaxation Methods for Convex Feasibility Problems*, Linear Algebra and Its Applications, Vol. 252, pp. 27–33, 1997.
21. Güler, O., *Barrier Functions in Interior-Point Methods*, Mathematics of Operations Research, Vol. 21, pp. 860–885, 1996.
22. Nesterov, Y. E., and Todd, M. J., *Self-Scaled Barriers and Interior-Point Methods for Convex Programming*, Mathematics of Operations Research, Vol. 22, pp. 1–42, 1997.
23. Lobo, M. S., Vandenberghe, L., Boyd, S., and Lebret, H., *Applications of Second-Order Cone Programming*, Linear Algebra and Its Applications, Vol. 284, pp. 193–228, 1998.
24. Bruck, R. E., and Reich, S. *Nonexpansive Projections and Resolvents of Accretive Operators in Banach Spaces*, Houston Journal of Mathematics, Vol. 3, pp. 459–470, 1977.

25. De Pierro, A. R., and Iusem, A. N., *On the Asymptotic Behavior of Some Alternate Smoothing Series Expansion Iterative Methods*, Linear Algebra and Its Applications, Vol. 130, pp. 3–24, 1990.
26. Bauschke, H. H., *Projection Algorithms: Results and Open Problems*, Inherently Parallel Algorithms in Feasibility and Optimization and Their Applications (Haifa, 2000); Edited by D. Butnariu, Y. Censor, and S. Reich, Elsevier, Amsterdam, Holland, pp. 11–22, 2001.
27. Combettes, P. L., *Quasi-Fejérian Analysis of Some Optimization Algorithms*, Inherently Parallel Algorithms in Feasibility and Optimization and Their Applications (Haifa, 2000); Edited by D. Butnariu, Y. Censor, and S. Reich, Elsevier, Amsterdam, Holland, pp. 115–152, 2001.
28. Groetsch, C. W., *Generalized Inverses of Linear Operators: Representation and Approximation*, Dekker, New York, NY, 1977.
29. Deutsch, F., *Best Approximation in Inner Product Spaces*, Springer, New York, NY, 2001.
30. Golub, G. H., and Van Loan, C. F., *Matrix Computations*, Johns Hopkins University Press, Baltimore, Maryland, 1996.
31. Kruk, S., *Implementation of the Algorithms Discussed in This Manuscript*, www.oakland.edu/˜kruk/research/ProjRefl.
32. Horn, R. A., and Johnson, C. R., *Matrix Analysis*, Cambridge University Press, Cambridge, UK, 1985.
33. Borwein, J. M., and Lewis, A. S., *Convex Analysis and Nonlinear Optimization*, Springer, New York, NY, 2000.
34. Wolkowicz, H., Saigal, R., and Vandenberghe, L., Editors, *Handbook of Semidefinite Programming: Theory, Algorithms, and Applications*, International Series in Operations Research and Management Science, Kluwer, Norwell, Massachusetts, Vol. 27, 2000.
35. Vandenberghe, L., and Boyd, S., *Semidefinite Programming*, SIAM Review, Vol. 38, pp. 49–95, 1996.
36. Parlett, B. N., *The Symmetric Eigenvalue Problem*, Society for Industrial and Applied Mathematics, Philadelphia, Pennsylvania, 1998 (Corrected Reprint of the 1980 Original).
37. Stewart, G. W., *Matrix Algorithms, Vol. 2: Eigensystems*, Society for Industrial and Applied Mathematics, Philadelphia, Pennsylvania, 2001.
38. Bauschke, H. H., Borwein, J. M., and Lewis, A. S., *The Method of Cyclic Projections for Closed Convex Sets in Hilbert Space*, Recent Developments in Optimization Theory and Nonlinear Analysis (Jerusalem, 1995); Contemporary Mathematics, American Mathematical Society, Providence, Rhode Island, Vol. 204, pp. 1–38, 1997.
39. Cheney, W., and Goldstein, A. A., *Proximity Maps for Convex Sets*, Proceedings of the American Mathematical Society, Vol. 10, pp. 448–450, 1959.
40. Bauschke, H. H., and Borwein, J. M., *Dykstra's Alternating Projection Algorithm for Two Sets*, Journal of Approximation Theory, Vol. 79, pp. 418–443, 1994.
41. de Klerk, E., Roos, C., and Terlaky, T., *Infeasible-Start Semidefinite Programming Algorithms via Self-Dual Embeddings*, Topics in Semidefinite and

Interior-Point Methods (Toronto, 1996); Edited by P. M. Pardalos and H. Wolkowicz, American Mathematical Society, Providence, Rhode Island, pp. 215–236, 1998.

42. Luo, Z. Q., Sturm, J. F., and Zhang, S., *Conic Convex Programming and Self-Dual Embedding*, Optimization Methods and Software, Vol. 14, pp. 169–218, 2000.
43. Strang, G., *Linear Algebra and Its Applications*, Academic Press, New York, NY, 1976.
44. Kruk, S., *High-Accuracy Algorithms for the Solutions of Linear Programs*, University of Waterloo, PhD Thesis, 2001.