Infeasibility detection in the alternating direction method of multipliers for convex optimization

Goran Banjac, Paul Goulart, Bartolomeo Stellato and Stephen Boyd

June 1, 2017

Abstract

The alternating direction method of multipliers (ADMM) is a powerful operator splitting technique for solving structured optimization problems. For convex optimization problems, it is well known that the iterates generated by ADMM converge to a solution provided that it exists. If a solution does not exist, then the ADMM iterates do not converge. Nevertheless, we show that the ADMM iterates yield conclusive information regarding problem infeasibility for a wide class of convex optimization problems including both quadratic and conic programs. In particular, we show that in the limit the ADMM iterates either satisfy a set of first-order optimality conditions or produce a certificate of either primal or dual infeasibility. Based on these results, we propose termination criteria for detecting primal and dual infeasibility in ADMM.

1 Introduction

Operator splitting methods can be used to solve structured optimization problems of the form

$$\text{minimize} \quad f(x) + g(x), \tag{1}$$

where f and g are convex, closed and proper functions. These methods encompass algorithms such as the proximal gradient method (PGM), Douglas-Rachford splitting (DRS) and the alternating direction method of multipliers (ADMM) [1], and have been applied to problems ranging from feasibility and best approximation problems [2, 3] to quadratic and conic programs [4, 5, 6]. Due to their relatively low per-iteration computational cost and ability to exploit sparsity in the problem data [6], splitting methods are suitable for embedded [7, 8] and large-scale optimization [9], and have increasingly been applied for solving problems arising in signal processing [10, 11], machine learning [12] and optimal control [13].

In order to solve problem (1), PGM requires differentiability of one of the two functions. If a fixed step size is used in the algorithm, then one also requires a bound on the Lipschitz constant of the function's gradient [9]. On the other hand, ADMM and DRS, which turn out to be equivalent methods, do not require any additional assumptions on the problem beyond convexity, making them more robust to the problem data.

The growing popularity of ADMM has triggered a strong interest in understanding its theoretical properties. Provided that problem (1) is solvable, both ADMM and DRS are known to converge to an optimal solution [12, 14]. The use of ADMM for solving convex quadratic programs (QPs) was analysed in [4] and was shown to admit an asymptotic linear convergence rate. The authors in [15] analyse global linear convergence of ADMM for solving convex QPs, and the authors in [16] extend these results to a wider class of optimization problems. A particularly convenient framework for analysing asymptotic behaviour of such a method is by representing it as a fixed-point iteration of an averaged nonexpansive operator [14, 17, 16].

On the other hand, the ability to detect infeasibility of an optimization problem is very important in many applications, *e.g.* in any embedded application or in mixed-integer optimization when branch-and-bound techniques are used [18]. For infeasible problems, the asymptotic behaviour of ADMM and DRS has been studied only in some special cases. DRS for solving feasibility problems involving two convex sets when the sets do not intersect was studied in [3]. The authors in [19] study the asymptotic behaviour of ADMM for solving convex QPs when the problem is infeasible, but impose some strong assumptions on the problem data. The authors in [5] apply ADMM to the homogeneous self-dual embedding of a convex conic program, thereby producing a larger problem which is always feasible and whose solutions can be used either to produce a primal-dual solution or a certificate of infeasibility for the original problem.

In this paper we consider a very general class of convex optimization problems that includes linear programs (LPs), QPs, second-order cone programs (SOCPs) and semidefinite programs (SDPs) as special cases. We use a particular version of ADMM introduced in [20] that imposes no conditions on the problem data such as strong convexity of the objective function or full rank of the constraint matrix. We show that the method either converges to primal-dual solution or produces a certificate of primal or dual infeasibility. These results are directly applicable to infeasibility detection in ADMM for a wide range of problems.

1.1 Contents

We introduce the problem of interest in Section 2 and present a particular ADMM algorithm for solving it in Section 3. Section 4 analyses the asymptotic behaviour of ADMM and shows that the algorithm can detect primal and dual infeasibility of the problem. Section 5 demonstrates these results on several small numerical examples. Finally, Section 6 concludes the paper.

1.2 Notation

We introduce some definitions and notation that will be used in the rest of the paper. All definitions here are standard, and can be found *e.g.* in [21, 14].

Let $\mathbb{R}$ denote the set of real numbers, $\mathbb{R}_+$ the set of nonnegative real numbers, $\tilde{\mathbb{R}} := \mathbb{R} \cup \{+\infty\}$ the extended real line and $\mathbb{R}^n$ the n-dimensional real space equipped with inner product $\langle\cdot,\cdot\rangle$, induced norm $\|\cdot\|$ and identity operator $\mathrm{Id} : x \to x$. We denote by $\mathbb{R}^{m\times n}$ the set of real m-by-n matrices and by $\mathbb{S}^n$ ($\mathbb{S}^n_+$) the set of real n-by-n symmetric (positive semidefinite) matrices. Let $\mathrm{vec} : \mathbb{S}^n \to \mathbb{R}^{n^2}$ be the operator mapping a matrix to the stack of its columns, $\mathrm{mat} = \mathrm{vec}^{-1}$ its inverse operator, and $\mathrm{diag} : \mathbb{R}^n \to \mathbb{S}^n$ the operator mapping a vector to a diagonal matrix. For a sequence $\{x^k\}$, define $\delta x^{k+1} := x^{k+1} - x^k$. The *proximal operator* of a convex, closed and proper function $f : \mathbb{R}^n \to \tilde{\mathbb{R}}$ is given by

$$\mathrm{prox}_f(x) := \operatorname*{argmin}_y \{f(y) + \tfrac{1}{2}\|y - x\|^2\}.$$

For a nonempty, closed and convex set $\mathcal{C} \subseteq \mathbb{R}^n$, we denote the *indicator function* of $\mathcal{C}$ by

$$\mathcal{I}_{\mathcal{C}}(x) \coloneqq \begin{cases} 0 & x \in \mathcal{C}, \\ +\infty & \text{otherwise,} \end{cases}$$

the *projection* of $x \in \mathbb{R}^n$ onto $\mathcal{C}$ by

$$\Pi_{\mathcal{C}}(x) \coloneqq \operatorname*{argmin}_{y \in \mathcal{C}} \|x - y\|,$$

the *support function* of $\mathcal{C}$ by

$$S_{\mathcal{C}}(x) \coloneqq \sup_{y \in \mathcal{C}} \langle x, y \rangle ,$$

the *recession cone* of $\mathcal{C}$ by

$$\mathcal{C}^{\infty} \coloneqq \{x \in \mathbb{R}^n \mid x + \mathcal{C} \subset \mathcal{C}\},$$

and the *normal cone* of $\mathcal{C}$ at $x \in \mathcal{C}$ by

$$\mathrm{N}_{\mathcal{C}}(x) \coloneqq \{y \in \mathbb{R}^n \mid \sup_{x' \in \mathcal{C}} \left\langle x' - x, y \right\rangle \leq 0\}.$$

Note that $\Pi_{\mathcal{C}}$ is the proximal operator of the indicator function of $\mathcal{C}$. For a convex cone $\mathcal{K} \subseteq \mathbb{R}^n$, we denote its *polar cone* by

$$\mathcal{K}^{\circ} \coloneqq \{y \in \mathbb{R}^n \mid \sup_{x \in \mathcal{K}} \langle x, y \rangle \leq 0\},$$

and for any $b \in \mathbb{R}^n$ we denote a translated cone by $\mathcal{K}_b \coloneqq \mathcal{K} + \{b\}$.

Let $\mathcal{D}$ be a nonempty subset of $\mathbb{R}^n$. We denote the closure of $\mathcal{D}$ by $\overline{\mathcal{D}}$. For an operator $T : \mathcal{D} \to \mathbb{R}^n$ we define its *fixed-point set* as $\operatorname{Fix} T \coloneqq \{x \in \mathcal{D} \,|\, x = Tx\}$ and denote its range by $\operatorname{ran}(T)$. We say that T is *nonexpansive* if $(\forall x \in \mathcal{D})(\forall y \in \mathcal{D})$

$$\|Tx - Ty\| \leq \|x - y\|.$$

T is α-*averaged* with $\alpha \in (0, 1)$ if there exists a nonexpansive operator $R : \mathcal{D} \to \mathbb{R}^n$ such that $T = (1 - \alpha) \operatorname{Id} + \alpha R$.

2 Problem description

Consider the following convex optimization problem:

$$\begin{array}{ll} \underset{x}{\text{minimize}} & \frac{1}{2}x^T P x + q^T x \\ \text{subject to} & Ax \in \mathcal{C}, \end{array} \tag{2}$$

with $P \in \mathbb{S}_+^n$, $q \in \mathbb{R}^n$, $A \in \mathbb{R}^{m\times n}$ and $\mathcal{C} \subseteq \mathbb{R}^m$ a nonempty, closed and convex set. The dual of this problem is

$$\begin{array}{ll} \underset{s,y}{\text{maximize}} & -\frac{1}{2}s^T P s - S_{\mathcal{C}}(y) \\ \text{subject to} & Ps + A^T y = -q, \quad y \in (\mathcal{C}^\infty)^\circ, \end{array} \tag{3}$$

where the conic constraint on y is just the restriction of y to the domain of $S_{\mathcal{C}}$ [22, p.112 and Cor. 14.2.1]. We are interested in finding either an optimal solution to problem (2) or a certificate of either primal or dual infeasibility.

We will find it convenient to rewrite problem (2) in an equivalent form by introducing a variable $z \in \mathbb{R}^m$ to obtain

$$\begin{array}{ll} \underset{x,z}{\text{minimize}} & \frac{1}{2}x^T P x + q^T x \\ \text{subject to} & Ax = z, \quad z \in \mathcal{C}. \end{array} \tag{4}$$

We can then write the optimality conditions for problem (4) as:

$$Ax = z, \tag{5a}$$

$$Px + q + A^T y = 0, \tag{5b}$$

$$z \in \mathcal{C}, \quad y \in \mathrm{N}_{\mathcal{C}}(z), \tag{5c}$$

where $y \in \mathbb{R}^m$ is a Lagrange multiplier associated with the constraint $Ax = z$. If there exist $x \in \mathbb{R}^n$, $z \in \mathbb{R}^m$ and $y \in \mathbb{R}^m$ that satisfy the conditions (5), then we say that (x, z) is a *primal* and y is a *dual* solution of problem (4). For completeness, we include a proof of this result in Lemma A.1 of the Appendix.

We will use the following pair of results to certify infeasibility of (2) in cases where it is either primal or dual infeasible.

Proposition 1.

(i) If there exists some $\bar{y} \in \mathbb{R}^m$ *such that*

$$A^T \bar{y} = 0 \quad \textit{and} \quad S_{\mathcal{C}}(\bar{y}) < 0, \tag{6}$$

then the primal problem (2) *is infeasible.*

(ii) If there exists some $\bar{x} \in \mathbb{R}^n$ *such that*

$$P\bar{x} = 0, \quad A\bar{x} \in \mathcal{C}^\infty \quad \textit{and} \quad \langle q, \bar{x} \rangle < 0, \tag{7}$$

then the dual problem (3) *is infeasible.*

Proof. (i): The first condition in (6) implies that for all $x \in \mathbb{R}^n$,

$$\langle \bar{y}, Ax \rangle = \left\langle A^T \bar{y}, x \right\rangle = 0.$$

Similarly, the second condition in (6) is equivalent to $\sup_{z \in \mathcal{C}} \langle \bar{y}, z \rangle < 0$. Therefore, $\langle \bar{y}, z \rangle = 0$ is a hyperplane that strictly separates the sets $\{Ax \,|\, x \in \mathbb{R}^n\}$ and $\mathcal{C}$, meaning that the problem (2) is infeasible.

(ii): Define the set $\mathcal{Q} := \{Ps + A^T y \,|\, (s, y) \in \mathbb{R}^n \times (\mathcal{C}^\infty)^\circ\}$. The first two conditions in (7) imply that for all $(s, y) \in \mathbb{R}^n \times (\mathcal{C}^\infty)^\circ$

$$\left\langle \bar{x}, Ps + A^T y \right\rangle = \langle P\bar{x}, s \rangle + \langle A\bar{x}, y \rangle \leq 0,$$

where we used the fact that the inner product between vectors in a cone and its polar is nonpositive. Since the third condition in (7) can be written as $\langle \bar{x}, -q \rangle > 0$, this means that $\langle \bar{x}, x \rangle = 0$ is a hyperplane that strictly separates the sets $\mathcal{Q}$ and $\{-q\}$, and thus the dual problem (3) is infeasible. □

Note that if the condition (6) in Proposition 1 holds, then $\bar{y}$ also represents an unbounded direction in the dual problem assuming it is feasible. Likewise, $\bar{x}$ in condition (7) represents an unbounded direction for the primal problem if it is feasible. However, since we cannot exclude the possibility of both primal and dual infeasibility, we will refer to the condition (6) as *primal infeasibility* rather than *dual unboundedness*, and vice-versa for (7).

In some cases, *e.g.* when $\mathcal{C}$ is compact or polyhedral, conditions (6) and (7) in Proposition 1 are also necessary for infeasibility, and we say that (6) and (7) are *strong alternatives* for primal and dual feasibility, respectively. When $\mathcal{C}$ is a convex cone, additional assumptions are required for having strong alternatives; see *e.g.* [23, §5.9.4].

3 Alternating Direction Method of Multipliers (ADMM)

The alternating direction method of multipliers (ADMM) [12] is an operator splitting method that can be used for solving structured optimization problems. The iterates for ADMM in application to problem (1) can be written as

$$\tilde{x}^{k+1} = \mathrm{prox}_f(x^k - u^k) \tag{8a}$$

$$x^{k+1} = \mathrm{prox}_g(\tilde{x}^{k+1} + u^k) \tag{8b}$$

$$u^{k+1} = u^k + \tilde{x}^{k+1} - x^{k+1}. \tag{8c}$$

If $\tilde{x}^{k+1}$ in (8b) and (8c) is replaced by $\alpha\tilde{x}^{k+1} + (1-\alpha)x^k$ where $\alpha \in (0, 2)$ is a *relaxation parameter*, then the resulting algorithm is called *relaxed* ADMM.

We can write problem (4) in the general form (1) by setting

$$\begin{aligned} f(x, z) &= \tfrac{1}{2}x^T P x + q^T x + \mathcal{I}_{Ax=z}(x, z), \\ g(x, z) &= \mathcal{I}_{\mathcal{C}}(z). \end{aligned}$$

If we use the norm $\|(x, z)\| = \sigma\|x\| + \rho\|z\|$ with $(\sigma, \rho) > 0$ in the proximal operators of functions $f(x, z)$ and $g(x, z)$, then relaxed ADMM reduces to Algorithm 1, which was first introduced in [20]. The scalars ρ and σ are called the *penalty parameters*. Note that the strict positivity of both ρ and σ ensure that the equality constrained QP in step 3 of Algorithm 1 has a unique solution for any $P \in \mathbb{S}^n_+$ and $A \in \mathbb{R}^{m\times n}$.

It is well-known that ADMM and DRS are equivalent methods [24]. The authors in [25] show that the ADMM algorithm can be described alternatively in terms of the averaged iteration of a nonexpansive operator. In particular, an iteration of Algorithm 1 is equivalent

Algorithm 1 ADMM for problem (2)

given initial values x^0, z^0, y^0 and parameters $\rho > 0, \sigma > 0, \alpha \in (0, 2)$

repeat

$\quad (\tilde{x}^{k+1}, \tilde{z}^{k+1}) = \operatorname{argmin}_{(\tilde{x},\tilde{z}):A\tilde{x}=\tilde{z}} \frac{1}{2}\tilde{x}^T P\tilde{x} + q^T\tilde{x} + \frac{\sigma}{2}\|\tilde{x} - x^k\|^2 + \frac{\rho}{2}\|\tilde{z} - z^k + \frac{1}{\rho}y^k\|^2$

$\quad x^{k+1} = \alpha\tilde{x}^{k+1} + (1-\alpha)x^k$

$\quad z^{k+1} = \Pi_{\mathcal{C}}\left(\alpha\tilde{z}^{k+1} + (1-\alpha)z^k + \frac{1}{\rho}y^k\right)$

$\quad y^{k+1} = y^k + \rho\left(\alpha\tilde{z}^{k+1} + (1-\alpha)z^k - z^{k+1}\right)$

until termination condition is satisfied

to

$$\begin{aligned}
(\tilde{x}^k, \tilde{z}^k) &= \operatorname*{argmin}_{(\tilde{x},\tilde{z}):A\tilde{x}=\tilde{z}} \frac{1}{2}\tilde{x}^T P\tilde{x} + q^T\tilde{x} + \frac{\sigma}{2}\|\tilde{x} - x^k\|^2 + \frac{\rho}{2}\|\tilde{z} - (2\Pi_{\mathcal{C}} - \operatorname{Id})(v^k)\|^2 && \text{(9a)} \\
x^{k+1} &= x^k + \alpha\left(\tilde{x}^k - x^k\right) && \text{(9b)} \\
v^{k+1} &= v^k + \alpha\left(\tilde{z}^k - \Pi_{\mathcal{C}}(v^k)\right) && \text{(9c)}
\end{aligned}$$

where

$$z^k = \Pi_{\mathcal{C}}(v^k) \quad \text{and} \quad y^k = \rho(\operatorname{Id} - \Pi_{\mathcal{C}})(v^k). \tag{10}$$

We will exploit the following result in the next section to analyse asymptotic behaviour of the algorithm.

Fact 1. *The iteration described in* (9) *amounts to* $(x^{k+1}, v^{k+1}) = T(x^k, v^k)$, *where* T *is an* $(\alpha/2)$-*averaged operator.*

Proof. Follows from [25, §IV-C]. □

Due to [14, Prop. 6.46], the identities in (10) imply that at every iteration the pair (z^k, y^k) satisfies the optimality condition (5c) by construction. The solution of the equality constrained QP in (9a) satisfies the pair of optimality conditions

$$\begin{aligned}
A\tilde{x}^k - \tilde{z}^k &= 0 && \text{(11a)} \\
(P + \sigma I)\tilde{x}^k + q - \sigma x^k + \rho A^T\left(\tilde{z}^k - (2\Pi_{\mathcal{C}} - \operatorname{Id})(v^k)\right) &= 0. && \text{(11b)}
\end{aligned}$$

If we rearrange (9b) and (9c) to isolate $\tilde{x}^k$ and $\tilde{z}^k$, *i.e.* write

$$\tilde{x}^k = x^k + \tfrac{1}{\alpha}\delta x^{k+1} \tag{12a}$$
$$\tilde{z}^k = z^k + \tfrac{1}{\alpha}\delta v^{k+1}, \tag{12b}$$

and substitute them into (11), we obtain the following relations between the iterates:

$$Ax^k - \Pi_{\mathcal{C}}(v^k) = -\tfrac{1}{\alpha}\left(A\delta x^{k+1} - \delta v^{k+1}\right) \tag{13a}$$
$$Px^k + q + \rho A^T(\mathrm{Id} - \Pi_{\mathcal{C}})(v^k) = -\tfrac{1}{\alpha}\left((P+\sigma I)\delta x^{k+1} + \rho A^T \delta v^{k+1}\right). \tag{13b}$$

Observe that the right-hand terms of (13) are a direct measure of how far the iterates (x^k, z^k, y^k) are from satisfying the optimality conditions (5a) and (5b). In the next section, we will provide conditions under which the successive differences $(\delta x^k, \delta v^k)$ appearing in the right hand side of (13) converge and can be used to test for primal or dual infeasibility.

4 Assymptotic behaviour of ADMM

In order to analyse the asymptotic behaviour of the iteration (9), which is equivalent to Algorithm 1, we will rely heavily on the following results:

Lemma 1. *Suppose that $T : \mathcal{D} \to \mathcal{D}$ is an averaged nonexpansive operator. Let $x^0 \in \mathcal{D}$, $x^k = T^k x^0$ and δx be the projection of the zero vector onto $\overline{\mathrm{ran}(T - \mathrm{Id})}$. Then*

(i) $\frac{1}{k}x^k \to \delta x$.

(ii) $\delta x^k \to \delta x$.

(iii) If $\mathrm{Fix}\, T \neq \emptyset$*, then* x^k *converges to a point in* $\mathrm{Fix}\, T$.

Proof. The first result is [26, Cor. 2] and the last two are [3, Fact 3.2]. □

Note that since $\mathrm{ran}(T - \mathrm{Id})$ is not necessarily closed, the projection onto this set may not exist, but the projection onto its closure always exists. Due to Fact 1, Lemma 1.(i) ensures that $(\frac{1}{k}x^k, \frac{1}{k}v^k) \to (\delta x, \delta v)$, while Lemma 1.(ii) ensures that $(\delta x^k, \delta v^k) \to (\delta x, \delta v)$.

We make the following assumption on the constraint set in problem (2).

Assumption 1. *The set $\mathcal{C}$ is the Cartesian product of a convex compact set $\mathcal{B} \subseteq \mathbb{R}^{m_1}$ and a translated closed convex cone $\mathcal{K}_b \subseteq \mathbb{R}^{m_2}$,* i.e. $\mathcal{C} = \mathcal{B} \times \mathcal{K}_b$.

Many convex problems of practical interest, including LPs, QPs, SOCPs and SDPs, can be written in the form of problem (2) with $\mathcal{C}$ satisfying the conditions of Assumption 1. We can now prove several useful results about the limits δx and δv.

The core of our results is contained within the following two Propositions, which establish various relationships between the limits δx and δv. We include several supporting results required to prove these results in the Appendix. Given these two results, it will then be straightforward to extract certificates of optimality or infeasibility in Section 4.1. For both of these central results and in the remainder of the paper, we define

$$\delta z := \Pi_{\mathcal{C}^\infty}(\delta v) \quad \text{and} \quad \delta y := \rho \Pi_{(\mathcal{C}^\infty)^\circ}(\delta v).$$

Proposition 2. *Suppose that Assumption 1 holds. Then the following relations hold between the limits δx and δv:*

(i) $A\delta x = \delta z$.

(ii) $P\delta x = 0$.

(iii) $A^T \delta y = 0$.

(iv) $\delta z^k \to \delta z$ *and* $\delta y^k \to \delta y$.

Proof. Commensurate with our partitioning of the constraint set as $\mathcal{C} = \mathcal{B} \times \mathcal{K}_b$, we partition the matrix A and the iterates $(v^k,\, z^k,\, y^k)$ into components of appropriate dimension. We use subscript 1 for those components associated with the set $\mathcal{B}$ and subscript 2 for those associated with the set $\mathcal{K}_b$, *e.g.* $z^k = (z_1^k, z_2^k)$ where $z_1 \in \mathcal{B}$ and $z_2 \in \mathcal{K}_b$ and the matrix $A = [A_1; A_2]$. Note throughout that $\mathcal{C}^\infty = \{0\} \times \mathcal{K}$ and $(\mathcal{C}^\infty)^\circ = \mathbb{R}^{m_1} \times \mathcal{K}^\circ$.

(i): Divide (13a) by k, take the limit and apply Lemma 1 to get

$$A\delta x = \lim_{k \to \infty} \tfrac{1}{k} \Pi_{\mathcal{C}}(v^k).$$

Due to Lemma A.3(i), we then obtain

$$\begin{aligned} A_1\delta x &= \lim_{k\to\infty} \tfrac{1}{k}\Pi_{\mathcal{B}}(v_1^k) = 0, \\ A_2\delta x &= \lim_{k\to\infty} \tfrac{1}{k}\Pi_{\mathcal{K}_b}(v_2^k) = \Pi_{\mathcal{K}}(\delta v_2). \end{aligned}$$

(ii): Divide (13b) by ρk, take the inner product of both sides with δx and take the limit to obtain

$$\begin{aligned} -\tfrac{1}{\rho}\delta x^T P \delta x &= \lim_{k\to\infty} \left\langle A\delta x, \tfrac{1}{k}v_k - \tfrac{1}{k}\Pi_{\mathcal{C}}(v^k)\right\rangle \\ &= \left\langle A_1\delta x, \delta v_1 - \lim_{k\to\infty} \tfrac{1}{k}\Pi_{\mathcal{B}}(v_1^k)\right\rangle + \left\langle A_2\delta x, \delta v_2 - \lim_{k\to\infty} \tfrac{1}{k}\Pi_{\mathcal{K}_b}(v_2^k)\right\rangle \\ &= \left\langle \Pi_{\mathcal{K}}(\delta v_2), \delta v_2 - \Pi_{\mathcal{K}}(\delta v_2)\right\rangle \\ &= \langle \Pi_{\mathcal{K}}(\delta v_2), \Pi_{\mathcal{K}^\circ}(\delta v_2)\rangle \\ &= 0, \end{aligned}$$

where we used $A_1\delta x = 0$, $A_2\delta x = \Pi_{\mathcal{K}}(\delta v_2)$, Lemma A.3(i) and the Moreau decomposition [14, Thm 6.29]. Then $P\delta x = 0$ since $P \in \mathbb{S}_+^n$.

(iii): Divide (13b) by ρk, take the limit and use $P\delta x = 0$ to obtain

$$0 = \lim_{k\to\infty} \tfrac{1}{k} A^T(\mathrm{Id} - \Pi_{\mathcal{C}})(v^k).$$

We therefore have

$$0 = A_1^T\left(\delta v_1 - \lim_{k\to\infty} \tfrac{1}{k}\Pi_{\mathcal{B}}(v_1^k)\right) = A_1^T \delta v_1,$$

and

$$0 = A_2^T(\delta v_2 - \Pi_{\mathcal{K}}(\delta v_2)) = A_2^T \Pi_{\mathcal{K}^\circ}(\delta v_2),$$

where we used Lemma 1, Lemma A.3(i) and the Moreau decomposition again.

(iv): From (12) we have

$$-\tfrac{1}{\alpha}\left(\delta x^{k+1} - \delta x^k\right) = \delta x^k - \delta\tilde{x}^k, \tag{14a}$$

$$-\tfrac{1}{\alpha}\left(\delta v^{k+1} - \delta v^k\right) = \delta z^k - \delta\tilde{z}^k. \tag{14b}$$

Take the limit of (14a) to obtain

$$\lim_{k\to\infty} \delta\tilde{x}^k = \lim_{k\to\infty} \delta x^k = \delta x.$$

From (11) we now have $\delta\tilde{z}^k = A\delta\tilde{x}^k \to A\delta x$. Take the limit of (14b) and use the result from (i) to obtain

$$\lim_{k\to\infty} \delta z^k = \lim_{k\to\infty} \delta\tilde{z}^k = A\delta x = \Pi_{\mathcal{C}^\infty}(\delta v).$$

Since $\frac{1}{\rho}y^k = v^k - z^k$, we have

$$\begin{aligned}\lim_{k\to\infty} \tfrac{1}{\rho}\delta y^k &= \lim_{k\to\infty} \delta v^k - \lim_{k\to\infty} \delta z^k \\ &= \delta v - \Pi_{\mathcal{C}^\infty}(\delta v) \\ &= \Pi_{(\mathcal{C}^\infty)^\circ}(\delta v).\end{aligned}$$

□

Proposition 2 shows that the limits δx and δy will always satisfy the subspace and conic constraints in the primal and dual infeasibility tests (6) and (7), respectively. We next consider the terms appearing in the inequalities in (6) and (7).

Proposition 3. *Suppose that Assumption 1 holds. Then the following identities hold for the limits δx and δy:*

(i) $\langle q, \delta x\rangle = -\frac{\sigma}{\alpha}\|\delta x\|^2 - \frac{\rho}{\alpha}\|A\delta x\|^2$.

(ii) $S_{\mathcal{C}}(\delta y) = -\frac{1}{\alpha\rho}\|\delta y\|^2$.

Proof. (i): Take the inner product of both sides of (13b) with δx to obtain

$$\left\langle P\delta x, x^k\right\rangle + \langle \delta x, q\rangle + \rho\left\langle A\delta x, (\mathrm{Id} - \Pi_{\mathcal{C}})(v^k)\right\rangle = -\tfrac{\sigma}{\alpha}\left\langle \delta x, \delta x^{k+1}\right\rangle - \tfrac{\rho}{\alpha}\left\langle A\delta x, \delta v^{k+1}\right\rangle.$$

Using $A_1\delta x = 0$ and $P\delta x = 0$ from Proposition 2(i)–(ii) and then taking the limit gives

$$\begin{aligned}\langle q, \delta x\rangle + \tfrac{\sigma}{\alpha}\|\delta x\|^2 &= -\tfrac{\rho}{\alpha}\left\langle A\delta x, \delta v\right\rangle - \lim_{k\to\infty} \rho\left\langle A_2\delta x, \Pi_{\mathcal{K}^\circ}(v_2^k - b)\right\rangle \\ &= -\tfrac{\rho}{\alpha}\left\langle \Pi_{\mathcal{C}^\infty}(\delta v), \delta v_2\right\rangle - \lim_{k\to\infty} \rho\left\langle \Pi_{\mathcal{K}}(\delta v_2), \Pi_{\mathcal{K}^\circ}(v_2^k - b)\right\rangle \\ &= -\tfrac{\rho}{\alpha}\|\Pi_{\mathcal{C}^\infty}(\delta v)\|^2 \\ &= -\tfrac{\rho}{\alpha}\|A\delta x\|^2,\end{aligned}$$

where we used Proposition 2(i), Lemma A.2(ii)–(iii) and Lemma A.3(ii).

(ii): Take the inner product of both sides of (13a) with $\Pi_{(\mathcal{C}^\infty)^\circ}(\delta v)$ to obtain

$$\left\langle A^T \Pi_{(\mathcal{C}^\infty)^\circ}(\delta v), x^k + \tfrac{1}{\alpha}\delta x^{k+1} \right\rangle - \left\langle \Pi_{(\mathcal{C}^\infty)^\circ}(\delta v), \Pi_{\mathcal{C}}(v^k) \right\rangle = \tfrac{1}{\alpha} \left\langle \Pi_{(\mathcal{C}^\infty)^\circ}(\delta v), \delta v^{k+1} \right\rangle.$$

According to Proposition 2(iii) the first inner product is zero, and taking the limit we obtain

$$\begin{aligned}\lim_{k\to\infty} \left\langle \Pi_{(\mathcal{C}^\infty)^\circ}(\delta v), \Pi_{\mathcal{C}}(v^k) \right\rangle &= -\tfrac{1}{\alpha} \left\langle \Pi_{(\mathcal{C}^\infty)^\circ}(\delta v), \delta v \right\rangle \\ &= -\tfrac{1}{\alpha} \|\Pi_{(\mathcal{C}^\infty)^\circ}(\delta v)\|^2,\end{aligned}$$

where the last equality follows from Lemma A.2(iii). The limit in the above identity can be expressed as

$$\begin{aligned}\lim_{k\to\infty} \left\langle \Pi_{(\mathcal{C}^\infty)^\circ}(\delta v), \Pi_{\mathcal{C}}(v^k) \right\rangle &= \lim_{k\to\infty} \left\langle \delta v_1, \Pi_{\mathcal{B}}(v_1^k) \right\rangle + \lim_{k\to\infty} \left\langle \Pi_{\mathcal{K}^\circ}(\delta v_2), \Pi_{\mathcal{K}_b}(v_2^k) \right\rangle \\ &= S_{\mathcal{B}}(\delta v_1) + \lim_{k\to\infty} \left\langle \Pi_{\mathcal{K}^\circ}(\delta v_2), b + \Pi_{\mathcal{K}}(v_2^k - b) \right\rangle \\ &= S_{\mathcal{B}}(\delta v_1) + \langle \Pi_{\mathcal{K}^\circ}(\delta v_2), b \rangle + \lim_{k\to\infty} \left\langle \Pi_{\mathcal{K}^\circ}(\delta v_2), \Pi_{\mathcal{K}}(v_2^k - b) \right\rangle \\ &= S_{\mathcal{B}}(\delta v_1) + \langle \Pi_{\mathcal{K}^\circ}(\delta v_2), b \rangle,\end{aligned}$$

where the second equality follows from Lemma A.4 and Lemma A.2(i), and the last equality from Lemma A.3(ii). The term $\langle \Pi_{\mathcal{K}^\circ}(\delta v_2), b \rangle$ can be written in terms of the support function of $\mathcal{K}_b$, *i.e.*

$$\langle b, \Pi_{\mathcal{K}^\circ}(\delta v_2) \rangle = \langle b, \Pi_{\mathcal{K}^\circ}(\delta v_2) \rangle + \sup_{z\in\mathcal{K}} \langle z, \Pi_{\mathcal{K}^\circ}(\delta v_2) \rangle = S_{\mathcal{K}_b}(\Pi_{\mathcal{K}^\circ}(\delta v_2)),$$

where the first equality follows from nonpositive product between vectors in a cone and its polar; since $0 \in \mathcal{K}$, the supremum is attained for $z = 0$.

Finally, we have

$$\begin{aligned}-\tfrac{1}{\alpha} \|\Pi_{(\mathcal{C}^\infty)^\circ}(\delta v)\|^2 &= S_{\mathcal{B}}(\delta v_1) + S_{\mathcal{K}_b}(\Pi_{\mathcal{K}^\circ}(\delta v_2)) \\ &= S_{\mathcal{C}}(\Pi_{(\mathcal{C}^\infty)^\circ}(\delta v)),\end{aligned}$$

and due to Proposition 2(iv)

$$S_{\mathcal{C}}(\tfrac{1}{\rho}\delta y) = -\tfrac{1}{\alpha}\|\tfrac{1}{\rho}\delta y\|^2,$$

The results then follows since both the support function and the norm are positive homogeneous. □

4.1 Optimality and infeasibility certificates

We are now in a position to prove that, in the limit, the iterates of Algorithm 1 satisfy either the optimality conditions (5) or produce a certificate of infeasibility. Recall that Fact 1, Lemma 1(ii) and Proposition 2(iv) ensure convergence of the sequence $\{\delta x^k, \delta z^k, \delta y^k\}$.

Proposition 4 (Optimality)**.** *If* $(\delta x^k, \delta z^k, \delta y^k) \to 0$*, then the optimality conditions* (5) *are satisfied in the limit,* i.e.

$$\|Px^k + q + A^T y^k\| \to 0 \quad \textit{and} \quad \|Ax^k - z^k\| \to 0. \tag{15}$$

Proof. Follows from (10) and (13). □

Lemma 1(iii) is sufficient to prove that if problem (2) is solvable then the sequence of iterates $\{x^k, z^k, y^k\}$ converges to its primal-dual solution. However, convergence of $\{\delta x^k, \delta z^k, \delta y^k\}$ to zero is not itself sufficient to prove convergence of $\{x^k, z^k, y^k\}$; we provide a numerical example in Subsection 5.3 to show when this scenario can occur. According to Proposition 4, in this case the optimality conditions are still satisfied in the limit. A meaningful criterion for detecting optimality is therefore that the norms in (15) are small.

We next show that if $\{\delta x^k, \delta z^k, \delta y^k\}$ converges to a nonzero value, then we can construct a certificate of primal and/or dual infeasibility. Note that due to Proposition 2(i), δz can be nonzero only if δx is nonzero.

Theorem 1. *Suppose that Assumption 1 holds.*

(i) If $\delta y \neq 0$*, then the problem* (2) *is infeasible and* δy *satisfies the primal infeasibility conditions* (6)*.*

(ii) If $\delta x \neq 0$*, then the problem* (3) *is infeasible and* δx *satisfies the dual infeasibility conditions* (7)*.*

(iii) If $\delta x \neq 0$ and $\delta y \neq 0$, then problems (2) *and* (3) *are simultaneously infeasible.*

Proof. (i): Follows from Proposition 2(iii) and Proposition 3(ii).

(ii): Follows from Proposition 2(i)–(ii) and Proposition 3(i).

(iii): Follows from (i) and (ii). □

Since $(\delta x^k, \delta y^k) \to (\delta x, \delta y)$, a meaningful criterion for detecting primal and dual infeasibility would be to use δy^k and δx^k to check the conditions (6) and (7), respectively. Infeasibility detection based on these vectors is used in OSQP [20, 27], an open-source operator splitting solver for quadratic programming.

5 Numerical examples

In this section we demonstrate via several numerical examples the different asymptotic behaviours of iterates generated by Algorithm 1 for solving optimization problems of the form (2).

5.1 Parametric QP

Consider the quadratic program

$$\begin{array}{ll} \underset{x_1,x_2}{\text{minimize}} & \frac{1}{2}x_1^2 + x_1 - x_2 \\ \text{subject to} & 0 \leq x_1 + a x_2 \leq u_1 \\ & 1 \leq x_1 \leq 3 \\ & 1 \leq x_2 \leq u_3, \end{array} \tag{16}$$

where $a \in \mathbb{R}$, $u_1 \in [0, +\infty\rangle$ and $u_3 \in [1, +\infty\rangle$ are parameters. Note that the above problem is an instance of problem (2) with

$$P = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad q = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad A = \begin{bmatrix} 1 & a \\ 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \mathcal{C} = [l, u], \quad l = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \quad u = \begin{bmatrix} u_1 \\ 3 \\ u_3 \end{bmatrix},$$

where $[l,u] \coloneqq \{x \in \mathbb{R}^n \,|\, l \le x \le u\}$. Depending on the values of parameters u_1 and u_3, the constraint set in (16) can be either bounded or unbounded. The projection onto a set $[l,u]$ can be evaluated as

$$\Pi_{[l,u]}(x) = \max\left(\min(x,u), l\right),$$

and the support function of the bounded set $\mathcal{B} = [l,u]$ as

$$S_{\mathcal{B}}(y) = \langle l, \min(y,0)\rangle + \langle u, \max(y,0)\rangle,$$

where min and max functions should be taken elementwise. The support function of the translated cone $\mathcal{K}_b$ is

$$S_{\mathcal{K}_b}(y) = \begin{cases} \langle b, y\rangle & y \in \mathcal{K}^\circ, \\ +\infty & \text{otherwise.} \end{cases}$$

In the sequel we will discuss four scenarios that can occur depending on the values of the parameters: (i) optimality, (ii) primal infeasibility, (iii) dual infeasibility, (iv) simultaneous primal and dual infeasibility, and will show that Algorithm 1 correctly produces certificates for all four scenarios. In all cases we set the parameters $\alpha = \rho = \sigma = 1$ and set the initial iterate $(x^0, z^0, y^0) = (0, 0, 0)$.

Optimality Consider the problem (16) with parameters

$$a = 1, \quad u_1 = 5, \quad u_3 = 3.$$

Algorithm 1 converges to $x^\star = (1,3)$, $z^\star = (4,1,3)$, $y^\star = (0,-2,1)$, for which the objective value equals -1.5, and we have

$$Ax^\star = z^\star \quad \text{and} \quad Px^\star + q + A^T y^\star = 0,$$

i.e. the pair $(x^\star, y^\star)$ is a primal-dual solution of problem (16). After 50 iterations the algorithm produces iterates (x^{50}, z^{50}, y^{50}) for which the objective value equals -1.4999 and the residuals are

$$\|Ax^{50} - z^{50}\| = 5.3964 \cdot 10^{-5} \quad \text{and} \quad \|Px^{50} + q + A^T y^{50}\| = 1.5430 \cdot 10^{-4}.$$

Recall that the iterates of the algorithm always satisfy the optimality conditions (5c).

Primal infeasibility We next set the parameters of problem (16) to

$$a = 1, \quad u_1 = 0, \quad u_3 = 3.$$

Note that in this case the constraint set is $\mathcal{C} = \mathcal{B} = \{0\} \times [1,3] \times [1,3]$. The sequence $\{\delta y^k\}$ generated by Algorithm 1 converges to $\delta y = (\frac{2}{3}, -\frac{2}{3}, -\frac{2}{3})$, and we have

$$A^T \delta y = 0 \quad \text{and} \quad S_{\mathcal{C}}(\delta y) = -\tfrac{4}{3} < 0.$$

According to Proposition 1(i), δy is a certificate of primal infeasibility for the problem. After 20 iterations the algorithm produces δy^{20}, for which we have

$$\|A^T \delta y^{20}\| = 9.9620 \cdot 10^{-5} \quad \text{and} \quad S_{\mathcal{C}}(\delta y) = -1.3333.$$

Dual infeasibility We set the parameters to

$$a = 0, \quad u_1 = 2, \quad u_3 = +\infty.$$

The constraint set has the form $\mathcal{C} = \mathcal{B} \times \mathcal{K}_b$ with

$$\mathcal{B} = [0,2] \times [1,3], \quad \mathcal{K} = \mathbb{R}_+, \quad b = 1,$$

and the constraint matrix A can be written as

$$A = [A_1; A_2] \quad \text{with} \quad A_1 = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \quad \text{and} \quad A_2 = \begin{bmatrix} 0 & 1 \end{bmatrix}. \tag{17}$$

The sequence $\{\delta x^k\}$ generated by Algorithm 1 converges to $\delta x = (0, \frac{1}{2})$, and we have

$$P\delta x = 0, \quad A_1 \delta x = 0, \quad A_2 \delta x = \tfrac{1}{2} \in \mathcal{K}, \quad \langle q, \delta x \rangle = -\tfrac{1}{2} < 0.$$

According to Proposition 1(ii), δx is a certificate of dual infeasibility of the problem. After 20 iterations the algorithm produces δx^{20}, for which we have

$$\|P\delta x^{20}\| = 7.8161 \cdot 10^{-6}, \quad \|A_1\delta x^{20}\| = 1.1054 \cdot 10^{-5},$$
$$A_2\delta x^{20} = 0.50000, \quad \left\langle q, \delta x^{20} \right\rangle = -0.50000.$$

Simultaneous primal and dual infeasibility We set

$$a = 0, \quad u_1 = 0, \quad u_3 = +\infty.$$

The constraint set has the form $\mathcal{C} = \mathcal{B} \times \mathcal{K}_b$ with

$$\mathcal{B} = \{0\} \times [1, 3], \quad \mathcal{K} = \mathbb{R}_+, \quad b = 1,$$

and the constraint matrix A can be written as in (17). The sequences $\{\delta x^k\}$ and $\{\delta y^k\}$ generated by Algorithm 1 converge to $\delta x = (0, \frac{1}{2})$ and $\delta y = (\frac{1}{2}, -\frac{1}{2}, 0)$, respectively. If we partition δy as $\delta y = (\delta y_1, \delta y_2)$ with $\delta y_1 = (\frac{1}{2}, -\frac{1}{2})$ and $\delta y_2 = 0$, then we have

$$A^T\delta y = 0, \quad S_{\mathcal{C}}(\delta y) = S_{\mathcal{B}}(\delta y_1) + S_{\mathcal{K}_b}(\delta y_2) = -\tfrac{1}{2} < 0,$$

and

$$P\delta x = 0, \quad A_1\delta x = 0, \quad A_2\delta x = \tfrac{1}{2} \in \mathcal{K}, \quad \langle q, \delta x \rangle = -\tfrac{1}{2} < 0.$$

Therefore, δx and δy are certificates that the problem is simultaneously primal and dual infeasible. After 20 iterations the algorithm produces δx^{20} and δy^{20}, for which we have

$$\|A^T\delta y^{20}\| = 1.3407 \cdot 10^{-7}, \quad S_{\mathcal{C}}(\delta y^{20}) = -0.50000,$$

and

$$\|P\delta x^{20}\| = 2.0546 \cdot 10^{-6}, \quad \|A_1\delta x^{20}\| = 2.9057 \cdot 10^{-6},$$
$$A_2\delta x^{20} = 0.50000, \quad \left\langle q, \delta x^{20} \right\rangle = -0.50000.$$

5.2 Infeasible SDPs from SDPLIB

We next demonstrate the asymptotic behaviour of Algorithm 1 on two infeasible SDPs from the benchmark library SDPLIB [28]. The problems are given in the following form

$$\begin{array}{ll} \underset{x,z}{\text{minimize}} & q^T x \\ \text{subject to} & Ax = z \\ & z \in \mathcal{S}_b^m, \end{array}$$

where $\mathcal{S}^m$ denotes the vectorized form of $\mathbb{S}_+^m$, *i.e.* $z \in \mathcal{S}^m$ is equivalent to $\text{mat}(z) \in \mathbb{S}_+^m$.

Let $X \in \mathbb{S}^m$ have the following eigenvalue decomposition

$$X = U \operatorname{diag}(\lambda_1, \ldots, \lambda_m) U^T.$$

Then the projection of X onto $\mathbb{S}_+^m$ is

$$\Pi_{\mathbb{S}_+^m}(X) = U \operatorname{diag}\left(\max(\lambda_1, 0), \ldots, \max(\lambda_m, 0)\right) U^T.$$

Primal infeasible SDP The primal infeasible problem `infp1` from SDPLIB has decision variables $x \in \mathbb{R}^{10}$ and $z \in \mathcal{S}^{30}$. We run Algorithm 1 with parameters $\alpha = 1$ and $\rho = \sigma = 0.1$ from the initial iterate $(x^0, z^0, y^0) = (0, 0, 0)$. After 500 iterations the algorithm produces δy^{500}, for which we have

$$\|A^T \delta y^{500}\| = 3.4515 \cdot 10^{-7}, \quad \text{mat}(\delta y^{500}) \preceq 4.0846 \cdot 10^{-5} I, \quad \left\langle b, \delta y^{500} \right\rangle = -21.924.$$

According to Proposition 1(i), δy^{500} is an approximate certificate of primal infeasibility of the problem.

Dual infeasible SDP Dual infeasible problem `infd1` from SDPLIB has decision variables $x \in \mathbb{R}^{10}$ and $z \in \mathcal{S}^{30}$. We run Algorithm 1 with parameters $\alpha = 1$ and $\rho = \sigma = 0.001$ from the initial iterate $(x^0, z^0, y^0) = (0, 0, 0)$. After 500 iterations the algorithm produces δx^{500}, for which we have

$$P \delta x = 0, \quad \text{mat}(A \delta x) \succeq -1.6479 \cdot 10^{-4} I, \quad \langle q, \delta x \rangle = -2.0311.$$

According to Proposition 1(ii), δx^{500} is an approximate certificate of dual infeasibility of the problem.

5.3 Infeasible SDP with no certificate

Consider the following feasibility problem [29, Ex. 5]

$$\begin{array}{ll} \underset{x_1, x_2}{\text{minimize}} & 0 \\ \text{subject to} & \begin{bmatrix} x_1 & 1 & 0 \\ 1 & x_2 & 0 \\ 0 & 0 & -x_1 \end{bmatrix} \succeq 0, \end{array} \tag{18}$$

noting that it is primal infeasible by inspection. If we write the constraint set in (18) as

$$\underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix}}_{A_1} x_1 + \underbrace{\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}}_{A_2} x_2 + \underbrace{\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}}_{A_0} \succeq 0$$

and denote by $A = [\operatorname{vec}(A_1) \;\; \operatorname{vec}(A_2)]$ and $b = -\operatorname{vec}(A_0)$, then the constraint can be written as $Ax \in \mathcal{S}_b^3$, where $\mathcal{S}^3$ denotes the vectorized form of $\mathbb{S}_+^3$. If we define $Y := \operatorname{mat}(y)$, then the primal infeasibility conditions (6) for the above problem amount to

$$Y_{11} - Y_{33} = 0, \quad Y_{22} = 0, \quad Y_{12} < 0, \quad Y \preceq 0,$$

where Y_{ij} denotes the element of $Y \in \mathbb{S}^3$ in the i-th row and j-th column. Given that $Y \preceq 0$ and $Y_{22} = 0$ imply $Y_{12} = 0$, the above system is infeasible as well. Note that $Y = 0$ is a feasible point for the dual of problem (18) and problem (18) is thus not dual infeasible.

We next show that $(\delta x^k, \delta Z^k, \delta Y^k) \to 0$, where $\delta Z^k := \operatorname{mat}(\delta z^k)$ and $\delta Y^k := \operatorname{mat}(\delta y^k)$. Set $x^k = \big((1 + \frac{\rho}{\sigma})\varepsilon, 1/\varepsilon\big)$ and $V^k := \operatorname{mat}(v^k) = \operatorname{diag}(\varepsilon, 1/\varepsilon, 0)$ where $\varepsilon > 0$. The iteration (9) then produces the following iterates

$$Z^k = V^k, \quad \tilde{x}^k = (\varepsilon, 1/\varepsilon), \quad \tilde{Z}^k = \operatorname{diag}(\varepsilon, 1/\varepsilon, -\varepsilon),$$

and therefore we have

$$\delta x^{k+1} = \alpha\,(\tilde{x}^k - x^k) = \alpha\,(-\tfrac{\rho}{\sigma}\varepsilon, 0),$$
$$\delta V^{k+1} = \alpha\,(\tilde{Z}^k - Z^k) = \alpha\ \mathrm{diag}(0, 0, -\varepsilon).$$

By taking ε arbitrarily small, we can make $(\delta x^{k+1}, \delta V^{k+1})$ arbitrarily close to zero, which according to Lemma 1 means that $(\delta x^k, \delta V^k) \to (\delta x, \delta V) = 0$, and according to Proposition 4 the optimality conditions (5) are satisfied in the limit. However, the sequence $\{x^k, Z^k, Y^k\}$ does not have a limit point; otherwise, such a point would be a certificate for optimality of the problem. Let T denote the fixed-point operator mapping (x^k, V^k) to (x^{k+1}, V^{k+1}). Since $(\delta x, \delta V) \in \overline{\mathrm{ran}(T - \mathrm{Id})}$ by definition, and $(\delta x, \delta V) \notin \mathrm{ran}(T - \mathrm{Id})$, this means that the set $\mathrm{ran}(T - \mathrm{Id})$ is open, and the distance from $(\delta x, \delta V)$ to $\mathrm{ran}(T - \mathrm{Id})$ is zero.

We run Algorithm 1 with parameters $\alpha = \rho = \sigma = 1$ from the initial iterate $(x^0, Z^0, Y^0) = (0, 0, 0)$. After $N = 10^4$ iterations the algorithm produces (x^N, z^N, y^N) for which the residuals are

$$\|Ax^N - z^N\| = 7.0727 \cdot 10^{-3} \quad \text{and} \quad \|A^T y^N\| = 1.0002 \cdot 10^{-2}.$$

6 Conclusions

We have analysed the asymptotic behaviour of the alternating direction method of multipliers for solving a wide class of convex optimization problems, and have shown that if the sequence of successive differences of the algorithm's iterates does not converge to zero then the problem is primal and/or dual infeasible. Based on these results, we have proposed termination criteria for detecting primal and dual infeasibility, providing for the first time a set of reliable and generic stopping criteria applicable to infeasible convex conic problems for ADMM. We have also provided numerical examples to demonstrate different asymptotic behaviours of the algorithm's iterates.

A Supporting results

Lemma A.1. *The first-order optimality conditions for problem* (4) *are the conditions* (5)*.*

Proof. Problem (4) is equivalent to the following Lagrange relaxation

$$\begin{array}{ll} \underset{x,y,z}{\text{minimize}} & \frac{1}{2}x^T Px + q^T x + y^T(Ax - z) \\ \text{subject to} & z \in \mathcal{C}. \end{array}$$

If we denote the objective function in the above problem by $F(x, y, z)$, then the optimality conditions can be written as [21, Thm. 6.12], [23]

$$\begin{aligned} & z \in \mathcal{C}, \\ & 0 = \nabla_x F(x, y, z) = Px + q + A^T y, \\ & 0 = \nabla_y F(x, y, z) = Ax - z, \\ & 0 \geq \sup_{z' \in \mathcal{C}} \left\langle -\nabla_z F(x, y, z), z' - z \right\rangle = \sup_{z' \in \mathcal{C}} \left\langle y, z' - z \right\rangle, \end{aligned}$$

where the last condition is equivalent to $y \in \mathrm{N}_{\mathcal{C}}(z)$. □

Lemma A.2. *For any vectors* (v, b) *and a closed convex cone* $\mathcal{K}$*,*

(i) $\Pi_{\mathcal{K}_b}(v) = b + \Pi_{\mathcal{K}}(v - b)$.

(ii) $(\mathrm{Id} - \Pi_{\mathcal{K}_b})(v) = \Pi_{\mathcal{K}^\circ}(v - b)$.

(iii) $\langle \Pi_{\mathcal{K}}(v), v \rangle = \|\Pi_{\mathcal{K}}(v)\|^2$.

Proof. Part (i) is from [14, Prop. 28.1(i)].

(ii): From part (i) we have

$$(\mathrm{Id} - \Pi_{\mathcal{K}_b})(v) = v - b - \Pi_{\mathcal{K}}(v - b) = \Pi_{\mathcal{K}^\circ}(v - b),$$

where the second equality follows from the Moreau decomposition.

(iii): From the Moreau decomposition, we have

$$\langle \Pi_{\mathcal{K}}(v), v \rangle = \langle \Pi_{\mathcal{K}}(v), \Pi_{\mathcal{K}}(v) + \Pi_{\mathcal{K}^\circ}(v) \rangle = \|\Pi_{\mathcal{K}}(v)\|^2.$$

□

Lemma A.3. *Suppose that $\mathcal{K}$ is a closed convex cone and for some sequence $\{v^k\}$ we have $\lim_{k\to\infty} \frac{1}{k} v^k = \delta v$. Then for any b*

(i) $\lim_{k\to\infty} \frac{1}{k}\Pi_{\mathcal{K}_b}(v^k) = \Pi_{\mathcal{K}}(\delta v)$.

(ii) $\lim_{k\to\infty} \left\langle \Pi_{\mathcal{K}^\circ}(\delta v), \Pi_{\mathcal{K}}(v^k - b)\right\rangle = 0$.

Proof. (i): Write the limit as

$$\begin{aligned}\lim_{k\to\infty} \tfrac{1}{k}\Pi_{\mathcal{K}_b}(v^k) &= \lim_{k\to\infty} \tfrac{1}{k}\left(b + \Pi_{\mathcal{K}}(v^k - b)\right) \\ &= \lim_{k\to\infty} \Pi_{\mathcal{K}}\left(\tfrac{1}{k}(v^k - b)\right) \\ &= \Pi_{\mathcal{K}}\left(\lim_{k\to\infty} \tfrac{1}{k} v^k\right),\end{aligned}$$

where the first equality uses Lemma A.2(i), and the second and third follow from the homogeneity and continuity of the projection operator, respectively.

(ii): Due to the Moreau decomposition, we have

$$\left\langle \Pi_{\mathcal{K}^\circ}(v^k - b), \Pi_{\mathcal{K}}(v^k - b)\right\rangle = 0,$$

and therefore

$$\begin{aligned}0 &= \tfrac{1}{k}\left\langle \Pi_{\mathcal{K}^\circ}(v^k - b), \Pi_{\mathcal{K}}(v^k - b)\right\rangle \\ &= \left\langle \Pi_{\mathcal{K}^\circ}\left(\tfrac{1}{k}(v^k - b)\right), \Pi_{\mathcal{K}}(v^k - b)\right\rangle,\end{aligned}$$

where the second line follows from the homogeneity of the projection operator. Taking the limit of the above equality and exploiting continuity of the projection operator gives the result. □

Lemma A.4. *Suppose that $\mathcal{B}$ is a convex compact set and for some sequence $\{v^k\}$ we have $\lim_{k\to\infty} \frac{1}{k} v^k = \delta v$. Then*

$$\lim_{k\to\infty} \left\langle \delta v, \Pi_{\mathcal{B}}(v^k)\right\rangle = S_{\mathcal{B}}(\delta v).$$

Proof. Let $z^k := \Pi_{\mathcal{B}}(v^k)$. We have the following inclusion [3, Prop. 6.46]

$$v^k - z^k \in \mathrm{N}_{\mathcal{B}}(z^k),$$

and thus

$$
\begin{aligned}
\tfrac{1}{k}(v^k - z^k) \in \mathrm{N}_{\mathcal{B}}(z^k) \ &\Rightarrow\ \sup_{z\in\mathcal{B}} \left\langle \tfrac{1}{k}(v^k - z^k), z - z^k \right\rangle \leq 0 \\
&\Rightarrow\ \left\langle \tfrac{1}{k}(v^k - z^k), z^k \right\rangle \geq \sup_{z\in\mathcal{B}} \left\langle \tfrac{1}{k}(v^k - z^k), z \right\rangle \\
&\Rightarrow\ \left\langle \tfrac{1}{k}(v^k - z^k), z^k \right\rangle = \sup_{z\in\mathcal{B}} \left\langle \tfrac{1}{k}(v^k - z^k), z \right\rangle = S_{\mathcal{B}}\left(\tfrac{1}{k}(v^k - z^k)\right),
\end{aligned}
$$

where the first equality in the last line holds since $z^k \in \mathcal{B}$. Taking the limit of the above identity, we obtain

$$
\lim_{k\to\infty} \left\langle \tfrac{1}{k}(v^k - z^k), z^k \right\rangle = \lim_{k\to\infty} S_{\mathcal{B}}\left(\tfrac{1}{k}(v^k - z^k)\right) = S_{\mathcal{B}}\big(\lim_{k\to\infty} \tfrac{1}{k}(v^k - z^k)\big),
$$

where the second equality follows from the continuity of the support function. Since $\mathcal{B}$ is compact, we have $\lim_{k\to\infty} \frac{1}{k} z^k = 0$ and thus

$$
\lim_{k\to\infty} \left\langle \delta v, z^k \right\rangle = S_{\mathcal{B}}(\delta v). \qquad \square
$$

References

[1] N. Parikh and S. Boyd. Proximal algorithms. *Foundations and Trends in Optimization*, 1(3):123–231, 2013.

[2] H. H. Bauschke and J. M. Borwein. On projection algorithms for solving convex feasibility problems. *SIAM Review*, 38(3):367–426, 1996.

[3] H. H. Bauschke, P. L. Combettes, and D. R. Luke. Finding best approximation pairs relative to two closed convex sets in Hilbert spaces. *Journal of Approximation Theory*, 127(2):178–192, 2004.

[4] D. Boley. Local linear convergence of the alternating direction method of multipliers on quadratic or linear programs. *SIAM Journal on Optimization*, 23(4):2183–2207, 2013.

[5] B. O'Donoghue, E. Chu, N. Parikh, and S. Boyd. Conic optimization via operator splitting and homogeneous self-dual embedding. *Journal of Optimization Theory and*

Applications, 169(3):1042–1068, 2016.

[6] Y. Zheng, G. Fantuzzi, A. Papachristodoulou, P. Goulart, and A. Wynn. Fast ADMM for semidefinite programs with chordal sparsity. In *American Control Conference (ACC)*, 2017.

[7] B. O'Donoghue, G. Stathopoulos, and S. Boyd. A splitting method for optimal control. *IEEE Transactions on Control Systems Technology*, 21(6):2432–2442, 2013.

[8] J. Jerez, P. Goulart, S. Richter, G. Constantinides, E. Kerrigan, and M. Morari. Embedded online optimization for model predictive control at megahertz rates. *IEEE Transactions on Automatic Control*, 59(12):3238–3251, 2014.

[9] A. Beck and M. Teboulle. Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems. *IEEE Transactions on Image Processing*, 18(11):2419–2434, 2009.

[10] P. L. Combettes and V. R. Wajs. Signal recovery by proximal forward-backward splitting. *Multiscale Modeling & Simulation*, 4(4):1168–1200, 2005.

[11] P. L. Combettes and J.-C. Pesquet. Proximal splitting methods in signal processing. *Springer Optimization and Its Applications*, 49:185–212, 2011.

[12] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends in Machine Learning*, 3(1):1–122, 2011.

[13] G. Stathopoulos, H. Shukla, A. Szucs, Y. Pu, and C. Jones. Operator splitting methods in control. *Foundations and Trends in Systems and Control*, 3(3):249–362, 2016.

[14] H. H. Bauschke and P. L. Combettes. *Convex analysis and monotone operator theory in Hilbert spaces*. Springer Science & Business Media, 2011.

[15] E. Ghadimi, A. Teixeira, I. Shames, and M. Johansson. Optimal parameter selection for the alternating direction method of multipliers (ADMM): Quadratic problems. *IEEE Transactions on Automatic Control*, 60(3):644–658, 2015.

[16] P. Giselsson and S. Boyd. Linear convergence and metric selection for Douglas-Rachford splitting and ADMM. *IEEE Transactions on Automatic Control*, 62(2):532–544, 2017.

[17] G. Banjac and P. Goulart. Global linear convergence in operator splitting methods. In *IEEE Conference on Decision and Control (CDC)*, pages 233–238, 2016.

[18] V. V. Naik and A. Bemporad. Embedded mixed-integer quadratic optimization using accelerated dual gradient projection. In *IFAC World Congress*, 2017.

[19] A. U. Raghunathan and S. Di Cairano. Infeasibility detection in alternating direction method of multipliers for convex quadratic programs. In *IEEE Conference on Decision and Control (CDC)*, pages 5819–5824, 2014.

[20] B. Stellato, G. Banjac, P. Goulart, A. Bemporad, and S. Boyd. OSQP: An operator splitting solver for quadratic programs. *Technical Report, University of Oxford*, 2017.

[21] R. T. Rockafellar, R. J.-B. Wets, and M. Wets. *Variational analysis*. Grundlehren der mathematischen Wissenschaften. Springer, 1998.

[22] R. T. Rockafellar. *Convex Analysis*. Princeton University Press, USA, 1970.

[23] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

[24] D. Gabay. Applications of the method of multipliers to variational inequalities. *Studies in Mathematics and its Applications*, 15(C):299–331, 1983.

[25] P. Giselsson, M. Fält, and S. Boyd. Line search for averaged operator iteration. In *IEEE Conference on Decision and Control (CDC)*, pages 1015–1022, 2016.

[26] A. Pazy. Asymptotic behavior of contractions in Hilbert space. *Israel Journal of Mathematics*, 9(2):235–240, 1971.

[27] B. Stellato and G. Banjac. OSQP: An operator splitting solver for quadratic programs, 2017.

[28] B. Borchers. SDPLIB 1.2, a library of semidefinite programming test problems. *Optimization Methods and Software*, 11(1):683–690, 1999.

[29] M. V. Ramana. An exact duality theory for semidefinite programming and its complexity implications. *Mathematical Programming*, 77(1):129–162, 1997.