Primal-dual algorithms and infinite-dimensional Jordan algebras of finite rank

Leonid Faybusovich* T. Tsuchiya†

November, 2001 (Revised: December, 2001 and November, 2002)

Abstract

We consider primal-dual algorithms for certain types of infinite-dimensional optimization problems. Our approach is based on the generalization of the technique of finite-dimensional Euclidean Jordan algebras to the case of infinite-dimensional JB-algebras of finite rank. This generalization enables us to develop polynomial-time primal-dual algorithms for "infinite-dimensional second-order cone programs." We consider as an example a long-step primal-dual algorithm based on the Nesterov-Todd direction. It is shown that this algorithm can be generalized along with complexity estimates to the infinite-dimensional situation under consideration. An application is given to an important problem of control theory: multi-criteria analytic design of the linear regulator. The calculation of the Nesterov-Todd direction requires in this case solving one matrix differential Riccati equation plus solving a finite-dimensional system of linear algebraic equations on each iteration. The size of this algebraic system is $m+1$ by $m+1$, where m is a number of quadratic performance criteria.

Key words: Interior-point algorithms, primal-dual algorithms, second-order cone programming, infinite-dimensional problems, control theory

1 Introduction

Finite-dimensional Euclidean Jordan algebras proved to be very useful for the analysis of interior-point algorithms of optimization [2, 3, 4, 5, 11, 14]. In the present paper we analyze the possibility of using infinite-dimensional Jordan algebras of finite rank in a similar fashion for the analysis of an infinite-dimensional situation. In particular, we concentrate on primal-dual algorithms which constitute probably the most important class of interior-point algorithms though other classes of interior-point algorithms can be generalized following the pattern presented here.

Let $(V, \langle , \rangle)$ be a Hilbert space, $\Omega \subset V$ be an open convex cone in V, $a, b \in V$, $X \subset V$ be a closed vector subspace in V. Consider an optimization problem:

$$\langle a, z \rangle \to \min, \tag{1}$$
$$z \in (b + X) \cap \bar{\Omega} \tag{2}$$

*Department of Mathematics 255 Hurley Building University of Notre Dame Notre Dame, IN 46556, USA (Email: leonid.faybusovich.1@nd.edu). This author was supported in part by DMS98-03191 and DMS01-02698.

†The Institute of Statistical Mathematics, 4-6-7 Minami-Azabu, Minato-Ku, Tokyo, 106-8569, Japan. (Email: tsuchiya@sun312.ism.ac.jp). This author was supported in part by the Grant-in-Aid for Scientific Research (C) 08680478 of the Ministry of Education, Culture, Sports, Science and Technology of Japan.

and its dual

$$\langle b, w\rangle \to \min, \tag{3}$$
$$w \in (a + X^{\perp}) \cap \bar{\Omega}^*. \tag{4}$$

Here $\bar{\Omega}$ is the closure of Ω (in the topology induced by norm :$\|z\| = \sqrt{\langle z, z\rangle}$),

$$\bar{\Omega}^* = \{w \in V : \langle w, z\rangle \geq 0, \forall z \in \Omega\}, \tag{5}$$

$X^{\perp}$ is the orthogonal complement of X in V with respect to the scalar product $\langle , \rangle$.

Let

$$\mathcal{F} = [(b + X) \cap \bar{\Omega}] \times [(a + X^{\perp}) \cap \bar{\Omega}^*].$$

We will assume throughout this paper that

$$\mathrm{int}(\mathcal{F}) = [(b + X) \cap \Omega] \times [(a + X^{\perp}) \cap \mathrm{int}(\bar{\Omega}^*)] \neq \emptyset. \tag{6}$$

It is very easy to see that if the pair $\bar{z}, \bar{w}$ satisfy (2) and (4), respectively, and

$$\langle \bar{z}, \bar{w}\rangle = 0,$$

then $\bar{z}$ is an optimal solution to (1), (2) and $\bar{w}$ is an optimal solution to (3), (4), respectively. Given $(z, w) \in V \times V$, we introduce the so-called duality gap:

$$\mu(z, w) = \frac{\langle z, w\rangle}{r}, \tag{7}$$

where $r > 0$ is some positive constant which will be specified later.

A typical primal-dual algorithm generates a sequence $(z^{(k)}, w^{(k)}) \in \mathrm{int}(\mathcal{F})$, $k = 0, 1, ...$, such that:

$$\mu(z^{(k+1)}, w^{(k+1)}) \leq \left(1 - \frac{\delta}{r^{\omega}}\right) \mu(z^{(k)}, w^{(k)}), \tag{8}$$

for some positive constants δ and ω.

The following proposition is a direct consequence of (8).

Proposition 1.1 *Let $0 < \varepsilon < 1$ be given and a primal-dual algorithm generates a sequence satisfying (8). Then*

$$\mu(z^{(k)}, w^{(k)}) \leq \varepsilon$$

for

$$k \geq r^{\omega} \frac{\log\left(\frac{\mu(z^{(0)}, w^{(0)})}{\varepsilon}\right)}{\delta}$$

provided $\delta / r^{\omega} < 1$.

For a proof see e.g. [17]. Observe that the existence of a primal-dual sequence satisfying (8) for an arbitrary $0 < \varepsilon < 1$ is highly nontrivial in an infinite-dimensional situation and, in particular, implies that (1), (2) and (3), (4) have no duality gap.

In the present paper we consider a rather special but important situation where V is a JB-algebra of a finite rank, and Ω is the so-called "cone of squares." The classification of JB-algebras of finite rank is known (see e.g. [8]) and is briefly described in Section 2 of the paper. It turns out that each such an algebra is a direct sum of uniquely defined irreducible factors. Each factor is either an irreducible finite-dimensional Euclidean Jordan algebra or the so-called (infinite-dimensional) spin-factor. This enables us to reduce the analysis of interior-point algorithms to two cases: a) V is a finite-dimensional Euclidean Jordan algebra and b) V is a direct sum of a finite number of infinite-dimensional spin-factors. It is well-known that the cone of squares Ω for a) is the symmetric cones. The cone of squares for b) is infinite-dimensional second-order cones.

The case a) is very well understood by now (see e.g. [2, 3, 4, 5, 11, 14]). We analyze in detail the case b) and show that it has a lot of similarities with the second-order cone programming [9, 12, 16]. Specifically, we pick up the long-step path-following algorithm with the Nesterov-Todd direction as an example and show that the algorithm terminates in $O(r \log \mu^0/\varepsilon)$ iterations, where μ^0 is the initial duality gap and ε is the final duality gap, and r is the rank of the associated JB-algebra.

The crucial point in the implementation of primal-dual algorithms is the availability of an efficient procedure for the calculation of an appropriate "descent direction" which enables one to move from $(z^{(k)}, w^{(k)})$ to $(z^{(k+1)}, w^{(k+1)})$. In the infinite-dimensional setting this problem is reduced to solving an infinite-dimensional system of linear equations. In the present paper we consider a concrete example, a min-max optimization problem with linear constraints in a Hilbert space, and show that the corresponding infinite-dimensional system can be efficiently solved. This problem admits a natural control-theoretic interpretation as a multi-criteria problem of the analytic design of a linear regulator.

2 JB-algebras algebras of finite rank

The purpose of this section is to describe the classification of JB-algebras of finite rank. For further details see [8].

Let V be a real commutative algebra with the unit element e. Given $z \in V$, consider the multiplication operator $L(z) : V \to V$,

$$L(z)z_1 = z \circ z_1, \ \ z_1 \in V.$$

Definition 2.1 *We say that V is a Jordan algebra if the identity*

$$[L(z), L(z^2)] = L(z)L(z^2) - L(z^2)L(z) = 0 \tag{9}$$

holds for any $z \in V$.

We can introduce the so-called quadratic representation in an arbitrary Jordan algebra V. Given $z \in V$,

$$P(z) = 2L(z)^2 - L(z^2). \tag{10}$$

A direct computation shows:

Proposition 2.2 *Given $z_1, z_2 \in V$, we have:*

$$P(P(z_1)z_2) = P(z_1)P(z_2)P(z_1). \tag{11}$$

Let V be a Jordan algebra with the unit element e and the multiplication operator $\circ$.

Definition 2.3 *An element $z \in V$ is called invertible in V if there exists $w \in V$ such that $z \circ w = e$, $z^2 \circ w = z$. We denote such an element w by z^{-1}.*

Proposition 2.4 *An element $z \in V$ is invertible if and only if $P(z)$ is an invertible linear operator. Moreover, in this case*

$$z^{-1} = P(z)^{-1} z.$$

Proposition 2.5 *Given an invertible element $z \in V$, a subalgebra generated by z, z^{-1}, e is associative.*

Definition 2.6 *A JB-algebra is a Jordan algebra V with the unit element e endowed with a complete norm $\| \cdot \|$ such that:*

$$\|z_1 \circ z_2\| \leq \|z_1\| \|z_2\|, \quad \|z_1^2\| = \|z_1\|^2, \quad \|z_1^2 + z_2^2\| \geq \|z_1^2\|, \; \forall z_1, z_2 \in V.$$

Proposition 2.7 *In every JB-algebra V the set*

$$\bar{\Omega} = \{z^2 : z \in V\} \tag{12}$$

is a closed convex cone.

Example 2.8 *Let K be a compact set and $Cont(K)$ is the vector space of continuous real-valued functions on K endowed with the norm:*

$$\|f\| = \sup\{|f(t)| : t \in K\}, \; f \in Cont(K).$$

It is quite obvious that $Cont(K)$ is a JB-algebra. A Jordan-algebraic multiplication in this example is the pointwise multiplication of functions. The cone $\bar{\Omega}$ is the cone of nonnegative functions from $Cont(K)$.

Lemma 2.9 *For every element z in a JB-algebra V, the closed subalgebra $C(z)$ generated by z and e is associative.*

Proposition 2.10 *Let V be a JB-algebra and $\bar{\Omega}$ defined in (12) be its cone of squares. The interior of $\bar{\Omega}$, which we denote by Ω, has the following properties:*

i) Ω is a nonempty open convex cone.

ii) Ω is the connected component of the unit element e in the set of invertible elements of V.

Let $\mathcal{L}(V)$ be the Banach space of bounded linear operators on V. Let, further,

$$GL(\Omega) = \{g \in \mathcal{L}(V) : \; g(\Omega) = \Omega, \; g \text{ is invertible in } \mathcal{L}(V)\}. \tag{13}$$

Proposition 2.11 *The cone Ω is linear homogeneous, i.e., for any $z \in \Omega$ there exists $g \in GL(\Omega)$ such that $ge = z$.*

Denote by $Aut(V)$ the group of Jordan algebra isomorphisms of a JB-algebra V, i.e., the group of invertible linear maps on V which preserve the Jordan-algebraic operations.

Proposition 2.12 *Given $g \in Aut(V), \|g(x)\| = \|x\|, \forall x \in V$. In particular, $Aut(V) \subset \mathcal{L}(V)$. Every $g \in GL(\Omega)$ admits a unique representation of the form (the polar decomposition):*

$$g = P(x) g_1, \; x \in \Omega, \; g_1 \in Aut(V).$$

We are now in position to introduce the "JB-algebras of finite rank" and its classification.

Let $(Y, (\cdot|\cdot))$ be a real Hilbert space. Introduce a multiplication operator on the vector space $V = \mathbf{R} \oplus Y$ as follows:

$$(s, x) \circ (t, y) = (st + (x|y), sy + tx).$$

If we denote $(1, 0) \in V$ by e, we immediately see that:

$$e \circ z = z \circ e = z, \ \forall z \in V.$$

It is easy to verify by a direct calculation that (9) holds.

Let $p = 1 + \dim Y$, where $\dim Y$ is the cardinality of an orthonormal basis in V. We call V the spin-factor (notation: V_p). It is known that spin-factors are JB-algebras with the norm defined as follows:

$$\|(t, y)\| = |t| + \sqrt{(y|y)}, \ (t, y) \in V.$$

Proposition 2.13 *Let V be a JB-algebra. The following conditions are equivalent:*

i) for every $z \in V$ the operator $L(z)$ satisfies a polynomial equation in $\mathcal{L}(V)$ over $\mathbf{R}$.

ii) there exists a natural number r such that every $z \in V$ admits a representation:

$$z = \lambda_1 e_1 + \lambda_2 e_2 + ... + \lambda_r e_r, \tag{14}$$

where $e_i \circ e_j = \delta_{ij} e_i, \ \lambda_i \in \mathbf{R}, \ i, j = 1, 2, ..., r$.

Proposition 2.13 singles out a subclass of JB-algebras of *finite rank*. The number r in Proposition 2.13 is called the *rank* of V (notation: $r = r(V)$).

Theorem 2.14 *Every JB-algebra of a finite rank admits a unique direct sum decomposition:*

$$V = V_1 \oplus V_2 \oplus ... \oplus V_k, \ r(V) = r(V_1) + ... + r(V_k) \tag{15}$$

and each V_i is either a spin-factor of infinite cardinality or a finite-dimensional irreducible JB-algebra.

Remark: Since the class of finite-dimensional JB-algebras coincides with the class of Euclidean Jordan algebras, there is a complete classification of finite-dimensional JB-algebras (see e.g. [1]).

3 Some Jordan-algebraic properties of spin-factors

In what follows we restrict ourselves to the analysis of problems (1), (2) and (3), (4) for the case where V is a JB-algebra of a finite rank. In view of Theorem 2.14, the only new feature in the analysis of interior-point algorithms for solving (1), (2) and (3),(4) is a possible presence of infinite-dimensional spin-factors in the decomposition (15). In this section we describe some Jordan-algebraic aspects of a spin-factor $\mathbf{R} \times Y$ essential for future considerations.

Let $z = (s, y) \in \mathbf{R} \times Y$. We start with the description of the multiplication operator $L(z)$. It is convenient to introduce the following notation. We think of $(s, y) \in V$ as a column vector $\begin{pmatrix} s \\ y \end{pmatrix}$. Then each linear operator on $\mathbf{R} \times Y$ admits the following block partition:

$$\begin{pmatrix} \alpha & A \\ B & C \end{pmatrix},$$

where $\alpha \in \mathbf{R}$, $A : Y \to \mathbf{R}$, $B : \mathbf{R} \to Y$, $C : Y \to Y$. Then

$$\begin{pmatrix} \alpha & A \\ B & C \end{pmatrix} \begin{pmatrix} s \\ y \end{pmatrix} = \begin{pmatrix} \alpha s + Ay \\ Bs + Cy \end{pmatrix}.$$

Since Y is a Hilbert space, each continuous linear map $A : Y \to \mathbf{R}$ has the form:

$$Ay = (a|y)$$

for some $a \in Y$. Each map $B : \mathbf{R} \to Y$ has the form $Bs = sb, b = B1$. Given $y \in Y$, introduce notation:

$$l_y : Y \to \mathbf{R},\ l_y(y_1) = (y|y_1),\ y_1 \in Y.$$

Observe that $l_y^T : \mathbf{R} \to Y$ has the form:

$$l_y^T(s) = sy,\ s \in \mathbf{R}.$$

Here l_y^T is the transpose of l_y with respect to the given scalar product $(\cdot|\cdot)$ on Y and the standard scalar product on $\mathbf{R}$, i.e.,

$$sl_y(y_1) = (l_y^T(s)|y_1),\ s \in \mathbf{R},\ y_1 \in Y.$$

With this notation, we have

Proposition 3.1 *Let $z = (s, y) \in \mathbf{R} \times Y$. Then*

$$L(z) = \begin{pmatrix} s & l_y \\ l_y^T & sI_Y \end{pmatrix}. \tag{16}$$

Here I_Y is the identity operator on Y.

Proof. The result immediately follows from definitions. ■

Our next goal is to explicitly calculate the spectral decomposition (14) for the spin-factor $\mathbf{R} \times Y$.

Proposition 3.2 *Let $(s, y) \in \mathbf{R} \times Y, y \neq 0$. Consider*

$$e_1 = \frac{1}{2}\left(1, \frac{y}{\|y\|}\right),\ e_2 = \frac{1}{2}\left(1, -\frac{y}{\|y\|}\right)$$
$$\lambda_1 = s + \|y\|,\ \lambda_2 = s - \|y\|,\ \|y\| = \sqrt{(y|y)}. \tag{17}$$

Then

$$(s, y) = \lambda_1 e_1 + \lambda_2 e_2, \tag{18}$$
$$e_1^2 = e_1,\ e_2^2 = e_2,\ e_1 \circ e_2 = 0. \tag{19}$$

Proof. A direct calculation. ■

Proposition 3.3 *Let $z \in (s, y) \in \mathbf{R} \times Y$. Then*

$$z^2 - 2sz + (s^2 - (y|y))e = 0.$$

Here $e = (1, 0)$ is the unit element in the Jordan algebra $\mathbf{R} \times Y$.

Proof. A direct computation. ■

.

Remark: Following the standard terminology (see e.g. [1]), we introduce the following notation:

$$\mathrm{tr}(z) = 2s, \ \det(z) = s^2 - (y|y). \tag{20}$$

Comparing (17) with (20), we see that

$$\mathrm{tr}(z) = \lambda_1(z) + \lambda_2(z), \ \det(z) = \lambda_1(z)\lambda_2(z). \tag{21}$$

The next proposition describes the inverse of an element $z = (s, y)$ in a spin-factor $\mathbf{R} \times Y$.

Proposition 3.4 *An element $z \in \mathbf{R} \times Y$ is invertible if and only if $\det(z) \neq 0$. In this case*

$$z^{-1} = \frac{1}{\det(z)}(s, -y) = \lambda_1(z)^{-1}e_1 + \lambda_2(z)^{-1}e_2,$$

(see (17), (18)).

Proof. A direct computation. ■

We next describe the quadratic representation (see (10)) in a spin-factor $\mathbf{R} \times Y$. Given $y \in Y$, we introduce a linear operator $y \otimes y \in \mathcal{L}(Y)$ as follows:

$$y \otimes y(y_1) = (y|y_1)y, \ y_1 \in Y. \tag{22}$$

Proposition 3.5 *Let $z = (s, y) \in \mathbf{R} \times Y$. Then*

$$P(s, y) = \det(z)I_V + 2\begin{pmatrix} (y|y) & sl_y \\ sl_y^T & y \otimes y \end{pmatrix}.$$

Here I_V is the identity map on $V = \mathbf{R} \times Y$.

Proof. By Proposition 3.1

$$L(z) = sI_V + \begin{pmatrix} 0 & l_y \\ l_y^T & 0 \end{pmatrix}.$$

Hence,

$$L(z)^2 = s^2 I_V + 2s\begin{pmatrix} 0 & l_y \\ l_y^T & 0 \end{pmatrix} + \begin{pmatrix} 0 & l_y \\ l_y^T & 0 \end{pmatrix}^2.$$

But

$$\begin{pmatrix} 0 & l_y \\ l_y^T & 0 \end{pmatrix}^2 = \begin{pmatrix} l_y l_y^T & 0 \\ 0 & l_y^T l_y \end{pmatrix}.$$

Further,

$$l_y l_y^T(t) = l_y(ty) = t(y|y), \ t \in \mathbf{R}.$$

Hence, $l_y l_y^T = (y|y)$. On the other hand,

$$l_y^T l_y(y_1) = l_y^T((y|y_1)) = (y|y_1)y, \ y_1 \in Y,$$

i.e.,

$$L(z)^2 = s^2 I_V + 2s \begin{pmatrix} 0 & l_y \\ l_y^T & 0 \end{pmatrix} + \begin{pmatrix} (y|y) & 0 \\ 0 & y \otimes y \end{pmatrix}.$$

Now,

$$z^2 = (s^2 + (y|y), 2sy).$$

Hence, using Proposition 3.1 again, we obtain:

$$L(z^2) = [s^2 + (y|y)]I_V + 2s \begin{pmatrix} 0 & l_y \\ l_y^T & 0 \end{pmatrix}.$$

Finally, by (10),

$$P(z) = 2L(z)^2 - L(z^2) = \det(z) I_V + 2s \begin{pmatrix} 0 & l_y \\ l_y^T & 0 \end{pmatrix} + 2 \begin{pmatrix} (y|y) & 0 \\ 0 & y \otimes y \end{pmatrix}.$$

■

We now describe the cone of squares in the spin-factor $\mathbf{R} \times Y$.

Proposition 3.6 *We have:*

$$\begin{aligned} \Omega &= \{(s, y) \in \mathbf{R} \times Y : s > \|y\|\}, \\ \bar{\Omega} &= \{(s, y) \in \mathbf{R} \times Y : s \geq \|y\|\}, && (23) \\ \bar{\Omega}^* &= \bar{\Omega}, && (24) \end{aligned}$$

i.e., the cone $\bar{\Omega}$ is self-dual.

Proof. Let $z = (s, y)$ have the spectral decomposition (18). By (21) and Proposition 3.4, z is invertible if and only if $\lambda_1(z) \neq 0$, $\lambda_2(z) \neq 0$. Using (19), we immediately see that

$$z^2 = \lambda_1(z)^2 e_1 + \lambda_2(z)^2 e_2.$$

Hence, by Proposition 2.10 $w \in \Omega$ implies

$$\lambda_1(w) > 0, \lambda_2(w) > 0. \tag{25}$$

On the other hand, using (17), (25) is equivalent to $s > \|y\|$. Inversely, $\lambda_1(w) > 0, \lambda_2(w) > 0$ implies $w = u^2$,

$$u = \sqrt{\lambda_1(w)} e_1 + \sqrt{\lambda_2(w)} e_2.$$

It remains to prove (24). Let $(t, x) \in \bar{\Omega}^*$. Then (see (5))

$$st + (x|y) \geq 0, \ \forall (s, y) \in \bar{\Omega}. \tag{26}$$

Since by (23) $(s, 0) \in \bar{\Omega}$ for $s > 0$, we deduce from (26) that $t \geq 0$. Take $\tilde{y} = -x$, $\tilde{s} = \|x\| + \varepsilon$, $\varepsilon > 0$. Obviously, $(\tilde{s}, \tilde{y}) \in \Omega$ and (26) yields:

$$t(\|x\| + \varepsilon) - \|x\|^2 \geq 0 \text{ for } \varepsilon > 0.$$

Taking limit as $\varepsilon \to 0$, we conclude that $t\|x\| \geq \|x\|^2$, i.e., $t \geq \|x\|$ (in the case $\|x\| = 0$, we have already proven $t \geq 0$). Inversely, let $(t, x) \in \mathbf{R} \times Y$ and $t \geq \|x\|$. Given $(s, y) \in \bar{\Omega}$,

$$ts + (x|y) \geq t\|y\| + (x|y) \geq t\|y\| - \|x\|\|y\| = (t - \|x\|)\|y\| \geq 0.$$

Here we used the Cauchy-Schwarz inequality. ■

Proposition 3.7 *Given* $z_1, z_2 \in \mathbf{R} \times Y$,

$$\det(P(z_1)z_2) = [\det(z_1)]^2 \det z_2,$$

where $\det(z)$ *is defined in (20).*

Proof. A direct computation. ■

We introduce now a canonical scalar product on $\mathbf{R} \times Y$:

$$\langle z_1, z_2 \rangle = \mathrm{tr}(z_1 \circ z_2)$$

If $z_i = (s_i, y_i)$, $i = 1, 2$, then by (20):

$$\langle z_1, z_2 \rangle = 2(s_1 s_2 + (y_1|y_2)). \tag{27}$$

Proposition 3.8 *Given* $z \in \bar{\Omega}$, $L(z) \geq 0$, *i.e.,*

$$\langle L(z)z_1, z_1 \rangle \geq 0, \ \forall z_1 \in \mathbf{R} \times Y \tag{28}$$

Proof. Let $z = (s, y)$, $z_1 = (t, x)$. Since $(s, y) \in \bar{\Omega}$, We have $s \geq \sqrt{(y|y)}$ by (23). Evaluating (28), we see that we need to check that

$$st^2 + 2(x|y)t + s(x|x) \geq 0, \ \forall t \in \mathbf{R}, \ x \in Y.$$

We can assume without loss of generality that $s > 0$ (if $s = 0$, then $y = 0$). Thus, we need to check that the quadratic in t polynomial

$$t^2 + \frac{2(x|y)t}{s} + (x|x)$$

is everywhere nonnegative. But its discriminant has the form

$$\Delta = \frac{(x|y)^2}{s^2} - (x|x).$$

Using Cauchy-Schwarz inequality and $s^2 \geq (y|y)$, we obtain:

$$\Delta \leq \frac{(x|x)(y|y)}{s^2} - (x|x) \leq 0.$$

The result follows. ■

In the next section, we will extend the polynomial-time convergence proof of primal-dual algorithms developed in [5] for finite-dimensional symmetric cone programs to the current infinite-dimensional setting. For this purpose, we need the following theorem which is an analogue of the result by Sturm [15] and plays a fundamental role in the analysis of finite-dimensional case.

Theorem 3.9 *Let* $z \in \Omega$. *Then* $L(z)$ *is invertible in* $\mathcal{L}(\mathbf{R} \times Y)$ *(i.e.,* $L(z)^{-1}$ *is a bounded linear operator from* $\mathbf{R} \times Y$ *onto itself) and, moreover,*

$$L(z)^{-1}\Omega \subset \Omega.$$

Proof. Let $z = (s, y) \in \Omega$ and $(t, x) \in \mathbf{R} \times Y$. We claim that

$$\begin{aligned} L(z)^{-1}\begin{pmatrix} t \\ x \end{pmatrix} &= \begin{pmatrix} r \\ u \end{pmatrix}, \\ r &= \frac{st - (x|y)}{\det(z)}, \qquad (29) \\ u &= \frac{1}{s}\left(x + \frac{(x|y) - st}{\det(z)} y\right). \qquad (30) \end{aligned}$$

It suffices to check that

$$L(z)\begin{pmatrix} r \\ u \end{pmatrix} = \begin{pmatrix} t \\ x \end{pmatrix},$$

which is a direct computation by using (16).

In order to prove the theorem, given $s > \|y\|$ and $t > \|x\|$, we need to check that $r > \|u\|$ (see Proposition 3.6). Observe that $(s, -y) \in \Omega$. Hence, (29) and Proposition 3.6 imply that $r \geq 0$. Observe that by (29), (30):

$$u = \frac{1}{s}(x - ry)$$

and consequently

$$(u|u) = \frac{1}{s^2}((x|x) + r^2(y|y) - 2r(x|y)).$$

Thus, $r^2 > (u|u)$ is equivalent to:

$$r^2(s^2 - (y|y)) + 2r(x|y) > (x|x). \qquad (31)$$

Using (29), we can rewrite (31) in the form

$$(st - (x|y))^2 + 2(x|y)(st - (x|y)) > (x|x)\det(z)$$

or

$$s^2t^2 - 2st(x|y) + (x|y)^2 + 2st(x|y) - 2(x|y)^2 > (x|x)(s^2 - (y|y))$$

(Recall that $\det(z) = s^2 - (y|y)$). This can be further simplified to:

$$s^2(t^2 - (x|x)) > (x|y)^2 - (x|x)(y|y).$$

But the last inequality is obvious, since $t^2 > (x|x)$ and $|(x|y)|^2 \leq (x|x)(y|y)$ by Cauchy-Schwarz inequality. ∎

4 Primal-dual algorithms

We now return to our pair of dual problems (1), (2) and (3), (4). In the remaining part of the paper we will assume that V is a JB-algebra of a finite rank, Ω is the cone of squares in V and $r = rank(V)$ in the definition of the duality gap (7). We continue to assume that the condition (6) is satisfied. The vector space V is endowed with the canonical Hilbert space structure. First of all there exists a canonical linear form $\mathrm{tr} : V \to \mathbf{R}$. It is defined through the direct sum decomposition (15). If $\dim V_i < \infty$, then there is a standard way to define the trace operator [1]. Otherwise V_i is an infinite-dimensional spin-factor and we use (20).

The scalar product is then defined as:

$$\langle z, w \rangle = \mathrm{tr}(z \circ w), \; z, w \in V.$$

Proposition 3.6 (along with the standard properties of finite-dimensional Euclidean Jordan Algebras) enables us to conclude that

$$\bar{\Omega}^* = \bar{\Omega}.$$

The advantage of the Jordan-algebraic framework suggested in the present paper is that we can easily carry over literally all interior-point algorithms along with their complexity estimates to the infinite-dimensional situation. Let us illustrate this point by considering a long-step primal-dual algorithm based on the Nesterov-Todd direction [13].

The main ingredient in the construction of primal-dual algorithms is the choice of a "descent" direction which drives the duality gap μ to zero. The class of scaling-invariant "descent" directions is obtained by solving the following system of linear equations. Given $(z, w) \in \Omega \times \Omega$ and $g \in GL(\Omega)$ (see (13)), observe first of all that $g^{-T} \in GL(\Omega)$, since $\bar{\Omega}^* = \bar{\Omega}$. The system of linear equations has the form:

$$L(\tilde{z})\tilde{\xi} + L(\tilde{w})\tilde{\eta} = \gamma\mu(z, w)e - \tilde{z} \circ \tilde{w}, \tag{32}$$

$$\tilde{\xi} \in g(X), \; \tilde{\eta} \in g^{-T}(X^{\perp}), \tag{33}$$

$$\tilde{z} = g(z), \; \tilde{w} = g^{-T}(w). \tag{34}$$

Here $0 < \gamma < 1$ is a real parameter and $(\tilde{\xi}, \tilde{\eta})$ is a scaled "descent direction." For a motivation of this construction see e.g. [4, 10, 16]. We consider a special choice of the cone automorphism g.

Proposition 4.1 *Given $(z_1, z_2) \in \Omega \times \Omega$, there exists a unique $z_3 \in \Omega$ such that*

$$P(z_3)z_1 = z_2. \tag{35}$$

Proof. The decomposition (15) leads to the corresponding decomposition of the cone of squares Ω:

$$\Omega = \Omega_1 \oplus \Omega_2 \oplus ... \oplus \Omega_k, \tag{36}$$

where Ω_i is the cone of squares in V_i, $i = 1, 2, ..., k$.

Hence, to prove (35) it suffices to consider two cases: a) $\dim V < \infty$ and b) $V = \mathbf{R} \times Y$ is a spin-factor. For the case a) we refer to [4]. We derive an explicit formula for the case b). The derivation below is a simplified modification of the one given in [16] for the analysis of the Nesterov-Todd direction for the finite-dimensional second-order cone programming.

Let $z_1 = (s, y)$, $z_2 = (t, x)$, $z_3 = (r, u)$. Consider, first, the case $\det(z_1) = \det(z_2) = 1$. By Proposition 3.7 we should have $\det(z_3) = 1$ or

$$r^2 - (u|u) = 1. \tag{37}$$

Using Proposition 3.5, we can rewrite (35) in the form:

$$y + u\langle (r, u), (s, y) \rangle = x, \tag{38}$$

$$s + 2(u|u)s + 2r(u|y) = t, \tag{39}$$

where

$$\langle (r, u), (s, y) \rangle = 2rs + 2(u|y)$$

(Compare with (27)). We can eliminate $(u|u)$ from (39), using (37). We obtain:

$$r\langle (r,u),(s,y)\rangle = t+s \tag{40}$$

Now (38) can be rewritten in the form:

$$u\langle (r,u),(s,y)\rangle = x-y. \tag{41}$$

From (40) and (41), we obtain:

$$r=\frac{s+t}{\delta},\ u=\frac{x-y}{\delta},\ \delta=\langle (r,u),(s,y)\rangle. \tag{42}$$

Substituting (41), (42) into (37), we obtain:

$$\delta^2=(t+s)^2-(x-y|x-y)=2+\langle (t,x),(s,y)\rangle, \tag{43}$$

where we used $\det(z_1)=\det(z_2)=1$. The formulas (42), (43) give explicit expressions for (r,u), proving the uniqueness of z_3 in (35).

A direct substitution of (42), (43) into (38), (39) shows that $z_3=(r,u)$ solves (35). The general case can be reduced to the considered case as follows. Let

$$\mu_i=\frac{1}{\sqrt{\det(z_i)}},\ i=1,2.$$

Then $\det(\mu_i z_i)=1$. Let $\tilde{z}_3\in\Omega$ be such that $P(\tilde{z}_3)(\mu_1 z_1)=\mu_2 z_2$. Then

$$P\left(\sqrt{\frac{\mu_1}{\mu_2}}\tilde{z}_3\right)z_1=z_2,$$

i.e,

$$z_3=\sqrt{\frac{\mu_1}{\mu_2}}\tilde{z}_3. \tag{44}$$

This completes the proof ■

Combining (42)–(44) we obtain

Corollary 4.2 *Let $z_1,z_2\in\Omega$, $z_1=(s,y), z_2=(t,x)$. Consider $z_3=(r,u)$ with*

$$r=\sqrt{\frac{\mu_1}{\mu_2}}\frac{\mu_1 s+\mu_2 t}{\sqrt{2+\mu_1\mu_2\langle (s,y),(t,x)\rangle}},$$
$$u=\sqrt{\frac{\mu_1}{\mu_2}}\frac{\mu_2 x-\mu_1 y}{\sqrt{2+\mu_1\mu_2\langle (s,y),(t,x)\rangle}},$$
$$\mu_i=\frac{1}{\sqrt{\det(z_i)}},\ i=1,2.$$

Then $z_3\in\Omega$ is a unique solution to (35) for the case $V=\mathbf{R}\times Y$.

Proposition 4.3 *Let Ω be the cone of squares in the spin-factor $\mathbf{R}\times Y$ and $(s,y)\in\Omega$. Consider*

$$z=\left(\frac{\mu}{2},\frac{y}{\mu}\right),\ \mu=\sqrt{s+\|y\|}+\sqrt{s-\|y\|}.$$

Then $z\in\Omega$ and $z^2=(s,y)$. Moreover, if

$$(s,y)=\lambda_1 e_1+\lambda_2 e_2$$

be the spectral decomposition of (s,y), then

$$z=\sqrt{\lambda_1}e_1+\sqrt{\lambda_2}e_2.$$

Proof. A direct computation. ∎

Remark: We denote z by $(s,y)^{1/2}$. Given $z \in \Omega$, we have $P(z^{1/2})^2 = P(z)$. It easily follows from (11). Thus

$$P(z^{1/2}) = P(z)^{1/2}. \tag{45}$$

Observe that (45) holds for an arbitrary JB-algebra V of a finite rank. It follows from decomposition (15) and the validity of (45) in the case $\dim V < \infty$ (see [1]).

We use Proposition 4.1 to introduce the so-called Nesterov-Todd direction in the infinite-dimensional setting. Given $z_1, z_2 \in \Omega$, let $z_3 \in \Omega$ be such that (35) holds. Take $g = P(z_3^{1/2}) \in GL(\Omega)$. Then

$$gz_1 = g^{-T}z_2 = v$$

and equations (32)–(34) takes the form

$$\tilde{\xi} + \tilde{\eta} = \gamma\mu(v,v)v^{-1} - v, \tag{46}$$

$$\tilde{\xi} \in P(z_3^{1/2})X, \ \tilde{\eta} \in P(z_3^{-1/2})(X^{\perp}), \tag{47}$$

$$v = P(z_3^{1/2})z_1 = P(z_3^{-1/2})z_2. \tag{48}$$

Observe that in the original variables, (46)–(48) has the form:

$$\xi + P(z_3)^{-1}\eta = \gamma\mu(z_1,z_2)z_2^{-1} - z_1. \tag{49}$$

$$\xi \in X, \ \eta \in X^{\perp} \tag{50}$$

It is obvious from (46)–(48) that the Nesterov-Todd direction exists and unique. Indeed, (46) and (47) show that $\tilde{\xi}$ is the orthogonal projection of the vector $\gamma\mu(v,v)v^{-1} - v$ onto the closed vector subspace $P(z_3^{1/2})X$. The existence and uniqueness of other popular directions (e.g., HRVW/KSH/M direction [16]) can be shown in a similar fashion.

As an example, consider a long-step primal-dual algorithm based on the Nesterov-Todd direction. Given $(z_1, z_2) \in \Omega \times \Omega$, let $z_4 \in \Omega$ be such that

$$v = v(z_1, z_2) = P(z_4^{1/2})z_2 = P(z_4)^{-1/2}z_1.$$

(Observe that $z_4 = z_3^{-1}$ in our previous notation.) Given $0 < \beta < 1$, introduce the so-called wide-neighborhood in $\Omega \times \Omega$:

$$N_{-\infty}(\beta) = \{(z_1, z_2) \in \Omega \times \Omega : \lambda_{\min}(v(z_1,z_2)^2) \geq (1-\beta)\mu(z_1,z_2)\}.$$

Here $\lambda_{\min}(z) = \min\{\lambda_i : i = 1, 2, ..., r\}$ in the decomposition (14). We can show that this neighborhood is scaling invariant in exactly the same way as in the case of finite-dimensional Euclidean Jordan algebra [5]. Note that the duality gap μ is also scaling invariant.

Fix $\varepsilon > 0$. Suppose that $(z_1^{(0)}, z_2^{(0)}) \in \mathrm{int}(\mathcal{F}) \cap N_{-\infty}(\beta)$ (see (6)). Let (ξ_k, η_k) be the Nesterov-Todd direction at the point $(z_1^{(k)}, z_2^{(k)})$ defined as in (49), (50). Let $\bar{t}$ be the largest value of $t \in [0,1]$ such that $z_1^{(k)} + t\xi^{(k)}, z_2^{(k)} + t\eta^{(k)}) \in N_{-\infty}(\beta)$. Set $(z_1^{(k+1)}, z_2^{(k+1)}) = (z_1^{(k)} + \bar{t}\xi_k, z_2^{(k)} + \bar{t}\eta_k)$. We stop the iteration when $\mu(z_1^{(k)}, z_2^{(k)}) \leq \varepsilon$.

Theorem 4.4 *For the primal-dual algorithm described above, we have:*

$$\mu(z_1^{(k)}, z_2^{(k)}) \leq \varepsilon$$

for

$$k \geq r \frac{\log\left(\dfrac{\mu(z_1^{(0)}, z_2^{(0)})}{\varepsilon}\right)}{(1-\gamma)\delta}$$

provided $\sqrt{\beta(1-\beta)} \leq 1-\gamma$ *and*

$$\delta(\beta,\gamma) = \frac{2\beta\gamma}{\beta\gamma^2/(1-\beta) + (1-\gamma)^2}.$$

The proof of this theorem is exactly as [5] where the case of general symmetric finite-dimensional cone programming have been considered. Observe that it is essential that we have Theorem 3.9 at our disposal. A direct proof of the analogous theorem for finite-dimensional second-order cone programs developed in [16] is also extended in a straightforward way to prove the theorem under the restriction that Ω is the direct sum of several finite/infinite-dimensional second-order cones.

Corollary 4.5 *There exists a sequence* $(z_1^{(k)}, z_2^{(k)}) \in \mathrm{int}(\mathcal{F})$ *such that*

$$\mu(z_1^{(k)}, z_2^{(k)}) \to 0$$

where $k \to \infty$.

The next theorem provides an infinite-dimensional generalization of the optimality criterion for (1), (2), and (3), (4) (see e.g. [2]).

Theorem 4.6 *Suppose that* V *is a JB-algebra of a finite rank,* Ω *is a cone of squares in* V *and (6) is satisfied. Then problems (1), (2), and (3), (4) both have optimal solutions. The sets of optimal solutions for both problems are bounded closed convex sets. If* z^* *(respectively,* w^**) is an optimal solution to (1), (2), (respectively, (3), (4)), then*

$$\langle z^*, w^* \rangle = 0. \tag{51}$$

Inversely, if z^* *satisfies (2),* w^* *satisfies (4) and (51) holds, then* z^* *is an optimal solution to (1), (2), and* w^* *is an optimal solution to (3), (4).*

Proof. Consider the sequence $(z^{(k)}, w^{(k)}) \in \mathrm{int}(\mathcal{F})$ such that $\langle z^{(k)}, w^{(k)} \rangle \to 0$, $k \to +\infty$. Without loss of generality, we can assume that

$$\langle z^{(k)}, w^{(k)} \rangle \leq \langle z^{(0)}, w^{(0)} \rangle,\ \ k = 0, 1, ...$$

Since $z^{(k)} - z^{(0)} \in X$, $w^{(k)} - w^{(0)} \in X^{\perp}$, we have:

$$\langle z^{(k)} - z^{(0)}, w^{(k)} - w^{(0)} \rangle = 0,\ \ k = 0, 1, ...$$

Hence,

$$\langle z^{(k)}, w^{(0)} \rangle + \langle z^{(0)}, w^{(k)} \rangle = \langle z^{(0)}, w^{(0)} \rangle + \langle z^{(k)}, w^{(k)} \rangle \leq 2\langle z^{(0)}, w^{(0)} \rangle,\ \ k = 0, 1, ... \tag{52}$$

Observe that (52) implies that $(z^{(k)}, w^{(k)}), k = 0, 1, ...$, is bounded. Indeed, due to decomposition (36), it suffices to consider the case where V is irreducible. If $\dim V < +\infty$, the result is well-known (see e.g. [1]). Let $V = \mathbf{R} \times Y$ be a spin-factor. Let $(t, x) \in \mathbf{R} \times Y$, $t > \|x\|$. Given $\alpha > 0$, consider the set

$$B_\alpha = \{(s, y) \in \mathbf{R} \times Y : s \geq \|y\|,\ st + (y|x) \leq \alpha\}.$$

If $(s, y) \in B_\alpha$, then by Cauchy-Schwarz inequality:

$$st + (y|x) \geq st - \|y\|\|x\| = s(t - \|x\|) + \|x\|(s - \|y\|) \geq s(t - \|x\|).$$

Hence,

$$\|y\| \leq s \leq \frac{\alpha}{t - \|x\|}.$$

Thus, the set B_α is bounded.

Since $(z^{(k)}, w^{(k)})$ is bounded, it follows that there is a subsequence $(z^{(k_l)}, w^{(k_l)})$, $l = 0, 1, ...$ which converges weakly to a feasible point (z^*, w^*). Observe that the feasible region $\mathcal{F}$ is convex and closed and, hence, weakly closed. Let us show that

$$\langle z^*, w^* \rangle = 0. \tag{53}$$

To simplify the notation, assume that $(z^{(k)}, w^{(k)})$ weakly converges to (z^*, w^*) when $k \to \infty$. We have:

$$\langle b - z^{(k)}, a - w^{(k)} \rangle = 0 \text{ or } \langle a, b \rangle + \langle z^{(k)}, w^{(k)} \rangle = \langle a, z^{(k)} \rangle + \langle b, w^{(k)} \rangle. \tag{54}$$

Taking limit in (54), when $k \to \infty$ and using $\langle z^{(k)}, w^{(k)} \rangle \to 0$, $z^{(k)} \to z^*$ (weakly), $w^{(k)} \to w^*$ (weakly), we obtain:

$$\langle a, b \rangle = \langle a, z^* \rangle + \langle b, w^* \rangle. \tag{55}$$

On the other hand,

$$\langle a - w^*, b - z^* \rangle = 0.$$

Comparing this with (55), we conclude that (53) holds. Let us show that each $(\tilde{z}, \tilde{w}) \in \mathcal{F}$ such that $\langle \tilde{z}, \tilde{w} \rangle = 0$ is a pair of optimal solutions for (1), (2), and (3), (4), respectively. Let z_1 be feasible for (1), (2). Then

$$\begin{aligned} \langle a, b \rangle &= \langle b, \tilde{w} \rangle + \langle a, \tilde{z} \rangle, \\ \langle a, b \rangle + \langle z_1, \tilde{w} \rangle &= \langle b, \tilde{w} \rangle + \langle a, z_1 \rangle, \end{aligned}$$

Using $\langle z, \tilde{w} \rangle \geq 0$, we obtain:

$$\langle b, \tilde{w} \rangle + \langle a, z_1 \rangle \geq \langle b, \tilde{w} \rangle + \langle a, \tilde{z} \rangle,$$

i.e., $\langle a, z_1 \rangle \geq \langle a, \tilde{z} \rangle$. Thus $\tilde{z}$ is an optimal solution to (1), (2). Similarly, we show that $\tilde{w}$ is an optimal solution to (3), (4). In particular, (z^*, w^*) constructed above is the pair of optimal solutions to (1), (2) and (3), (4), respectively. Besides, $\langle z^*, w^* \rangle = 0$. We then immediately see as above that if $\langle z, w \rangle > 0$ for a feasible pair (z, w),then (z, w) is not a pair of optimal solutions. Take any $(z, w) \in \text{int}(\mathcal{F})$. Then for any optimal pair (z^*, w^*), the condition (53) implies:

$$\langle w, z^* \rangle + \langle z, w^* \rangle = \langle z, w \rangle.$$

Reasoning as in the proof of boundedness of the sequence $(z^{(k)}, w^{(k)})$ above, we conclude that the set of optimal pairs is bounded. ■

5 Example

Consider the following optimization problem:

$$\max_{i\le i\le m} \|W_i y\| \to \min, \tag{56}$$

$$y \in c + Z. \tag{57}$$

Here $W_i : Y \to Y, i = 1, 2, ..., m$, are bounded linear operators on Y, Z is a closed vector subspace in the Hilbert space Y. Recall that $(\cdot|\cdot)$ the inner product associated with Y.

We can rewrite (56) and (57) in the form:

$$t \to \min, \tag{58}$$

$$\|W_i y\| \le t,\ \ i = 1, ..., m, \tag{59}$$

$$y \in c + Z. \tag{60}$$

Our immediate goal is to rewrite (58)–(60) in the form (1), (2).

Let $V_1 = \mathbf{R} \times Y$, $V = V_1 \times .... \times V_1$ (m times), $\Omega_1 = \{(s, y) \in \mathbf{R} \times Y : s > \|y\|\}$, $\Omega = \Omega_1 \times ... \times \Omega_1$. Consider a linear operator

$$\Lambda : V_1 \to V,$$
$$\Lambda(\mu, \zeta) = ((\mu, W_1\zeta), (\mu, W_2\zeta), ..., (\mu, W_m\zeta)).$$

Let, further, $a = ((1, 0), (0, 0), ..., (0, 0)) \in V$, $b = ((0, W_1c), (0, W_2c), ..., (0, W_mc)) \in V$, $z = (z_1, ..., z_m)$, $z_i = (t_i, x_i) \in V_1, i = 1, ..., m$. The scalar product in V is defined as follows:

$$\langle((t_1^{(1)}, x_1^{(1)}), ..., (t_m^{(1)}, x_m^{(1)})), ((t_1^{(2)}, x_1^{(2)}), ..., (t_m^{(2)}, x_m^{(2)}))\rangle = \sum_{i=1}^{m}[t_i^{(1)}t_i^{(2)} + (x_i^{(1)}|x_i^{(2)})].$$

We now can rewrite (58)–(60) in the form:

$$\langle a, z\rangle \to \min,$$
$$z \in (b + X) \cap \bar{\Omega},$$

where

$$X = \Lambda(\mathbf{R} \times Z). \tag{61}$$

An easy calculation shows that the orthogonal complement $X^\perp$ of X in V has the form:

$$X^\perp = \{((r_1, u_1), ..., (r_m, u_m)) \in V : r_1 + r_2 + ...r_m = 0,\ \sum_{i=1}^{m} W_i^* u_i \in Z^\perp\},$$

where $Z^\perp$ is the orthogonal complement of Z in Y and W_i^* is the adjoint of W_i for each i. According to (3), (4), its dual will be of the form

$$\sum_{i=1}^{m}(W_i c|u_i) \to \min$$
$$\sum_{i=1}^{m} r_i = m,\ \ \|u_i\| \le r_i,\ \ i = 1, 2, ..., m,$$
$$W_1^* u_1 + W_2^* u_2 + ... + W_m^* u_m \in Z^\perp.$$

It is easy to see that the condition (6) is satisfied. We can apply Theorem 4.6 in this example. Consider the Nesterov-Todd direction for our problem. Let $(m_1, m_2) \in \Omega \times \Omega$. According to (49) and (50) we need to find $(\xi, \eta) \in X \times X^{\perp}$ such that

$$P(z)\xi + \eta = \Delta. \tag{62}$$

Here $z \in \Omega$ is the scaling point uniquely determined from the equation $P(z)m_1 = m_2$ and $\Delta \in V$ is a known vector, depending on m_1, m_2.

We can rewrite (62) in the form:

$$P(z)\xi - \Delta \in X^{\perp},\ \xi \in X, \tag{63}$$

which is equivalent to:

$$\frac{\langle P(z)\xi, \xi\rangle}{2} - \langle \xi, \Delta\rangle \to \ \min, \tag{64}$$

$$\xi \in X. \tag{65}$$

Using the parameterization (61), we can write (64), (65) in the form:

$$\begin{aligned}&\rho(\mu, \zeta) = \frac{\langle P(z)\xi, \xi\rangle}{2} - \langle \xi, \Delta\rangle \to \ \min,\\ &(\mu, \zeta) \in \mathbf{R} \times Z.\end{aligned}$$

Observe that ρ is a convex quadratic function in variables (μ, ζ). Let $z = (z_1, ..., z_m)$, $z_i = (t_i, x_i) \in \Omega_1$, $\Delta = ((r_1^*, u_1^*), ..., (r_m^*, u_m^*)) \in V$, $\xi = (\xi_1, ..., \xi_m) \in X$. We obviously have:

$$P(z)\xi = (P(z_1)\xi_1, ..., P(z_m)\xi_m).$$

Using Proposition 3.5, we can easily calculate that

$$\begin{aligned}\rho(\mu, \zeta) &= \frac{1}{2}\sum_{i=1}^{m}(t_i^2 - \|x_i\|^2)\|W_i\zeta\|^2 + \sum_{i=1}^{m}(x_i|W_i\zeta)^2 - \sum_{i=1}^{m}(u_i^*|W_i\zeta) + \frac{\nu_1\mu^2}{2} + \nu_2\mu,\\ \nu_1 &= \sum_{i=1}^{m}(t_i^2 + \|x_i\|^2),\ \nu_2 = 2\sum_{i=1}^{m} t_i(x_i|W_i\zeta) - \sum_{i=1}^{m} r_i^*.\end{aligned}$$

Hence,

$$\begin{aligned}\phi(\zeta) &= \min\{\rho(\mu, \zeta) : \mu \in \mathbf{R}\}\\ &= \frac{(\zeta, M\zeta)}{2} + \frac{1}{2}(\zeta|(\sum_{i=1}^{m+1} \varepsilon_i(v_i \otimes v_i))\zeta) + (v_0|\zeta) - \frac{(\sum_{i=1}^{m} r_i^*)^2}{2\nu_1}\\ &= \frac{(\zeta, M\zeta)}{2} + \frac{1}{2}\sum_{i=1}^{m+1} \varepsilon_i(v_i|\zeta)^2 + (v_0|\zeta) - \frac{(\sum_{i=1}^{m} r_i^*)^2}{2\nu_1},\end{aligned} \tag{66}$$

where

$$\begin{aligned}M &= \sum_{i=1}^{m}(t_i^2 - \|x_i\|^2)W_i^*W_i,\\ v_i &= \sqrt{2}W_i^*x_i,\ i = 1, 2, ..., m,\\ v_{m+1} &= \frac{2}{\sqrt{\nu_1}}\sum_{i=1}^{m} t_iW_i^*x_i,\\ \varepsilon_i &= 1,\ i = 1, ..., m,\ \varepsilon_{m+1} = -1,\\ v_0 &= \sqrt{\frac{1}{\nu_1}}(\sum_{i=1}^{m} r_i^*)v_{m+1} - \sum_{i=1}^{m} W_i^*u_i^*,\end{aligned}$$

and $v_i \otimes v_i$ is defined as in (22).

Assume that

$$(M\zeta|\zeta) \geq \delta\|\zeta\|^2, \quad \forall \zeta \in Z \tag{67}$$

for some $\delta > 0$. Under this condition, we can show that the problem

$$\phi(\zeta) \to \min, \ \zeta \in Z, \tag{68}$$

where ϕ is described in (66) can be reduced to solving $(m+1)$ problems of the form

$$\frac{1}{2}(M\zeta|\zeta) + (v|\zeta) \to \min, \ \zeta \in Z \tag{69}$$

for appropriate choices of $v \in Y$, and the system of $(m+1) \times (m+1)$ linear algebraic equations. This observation makes sense because in some applications we have nice efficient algorithms to solve (69). Below we describe the procedure.

Let $\zeta_0 \in Z$ be the optimal solution to the problem

$$\frac{(\zeta|M\zeta)}{2} + (v_0|\zeta) \to \min, \ \zeta \in Z, \tag{70}$$

and $\zeta_i \in Z$, $i = 1, .., m+1$ be the optimal solutions to the problems

$$\frac{(\zeta|M\zeta)}{2} + (\varepsilon_i v_i|\zeta) \to \min, \ \zeta \in Z. \tag{71}$$

Let $S = [s_{ij}]$, $s_{ij} = (\zeta_i|v_j)$, $i, j = 1, 2, ..., m+1$, and

$$(I - S)\begin{pmatrix} \delta_1 \\ \vdots \\ \delta_{m+1} \end{pmatrix} = \begin{pmatrix} (v_0|\zeta_1) \\ \vdots \\ (v_0|\zeta_{m+1}) \end{pmatrix}, \tag{72}$$

then

$$\zeta(\delta) = \zeta_0 + \sum_{i=1}^{m+1} \delta_i \zeta_i$$

is an optimal solution to the problem (68). The procedure is a simple modification of the argument in [7] which deals with a version of the Sherman-Morrison-Woodbury formula in the infinite-dimensional setting. We give a derivation of (72) in Appendix.

Remark: It is easy to see that if (67) holds at an interior feasible solution z, then the linear operator $W : Y \to Y \times Y \times ... \times Y$ (m times) defined as $Wu \equiv (W_1 u, ..., W_m u)$ is invertible. Then y ((56) and (57)) is determined uniquely from z.

Consider now a more concrete situation in control theory which is similar to [6]. Denote by $L_2^n[0, T]$ the vector space of square integrable functions $f : [0, T] \to \mathbf{R}^n$. Let

$$Y = L_2^n[0, T] \times L_2^l[0, T], \ T > 0,$$

and

$$\begin{aligned} Z &= \{(\alpha, \beta) \in Y : \alpha \text{ is absolutely continuous on } [0, T], \\ &\quad \alpha(0) = 0, \dot{\alpha}(t) = A(t)\alpha(t) + B(t)\beta(t), \ t \in [0, T]\}. \end{aligned}$$

Here $A(t)$ (respectively $B(t)$) is an n by n (respectively, n by l) continuous matrix-valued function. Observe that

$$((\alpha_1, \beta_1)|(\alpha_2, \beta_2)) = \int_0^T [\alpha_1^T(t)\alpha_2(t) + \beta_1^T(t)\beta_2^T(t)]dt, \ (\alpha_i, \beta_i) \in Y, \ i = 1, 2.$$

In this case, $Z^\perp$ is easily calculated:

$$Z^\perp = \{(\dot{p} + A^T p, B^T p) : p \text{ is absolutely continous on } [0, T], \dot{p} \in L_2^n[0, T], p(T) = 0\}.$$

In the following, we deal with the following min-max optimization problem:

$$\max_i \int_0^T [(\alpha(t) - \bar{\alpha}_i(t))^T Q_i(\alpha(t) - \bar{\alpha}_i(t)) + (\beta(t) - \bar{\beta}_i)^T R_i(\beta(t) - \bar{\beta}_i(t))]dt, \to \min \tag{73}$$

where $Q_i(t)$ (respectively $R_i(t)$) is a continuous matrix-valued function such that $Q_i(t) = Q_i^T(t)$, $R_i(t) = R_i^T(t)$ are positive-definite symmetric matrices for any $t \in (0, T]$ and $(\bar{\alpha}(t), \bar{\beta}(t)) \in Y$. This problem is a very important problem in control theory, namely, a problem of multi-criteria design of the analytic regulator. This problem can be solved with our algorithm as follows.

For $i = 1, ..., m$, let $L_{Q_i}(t)$ and $L_{R_i}(t)$ be the lower triangular matrices obtained with the Cholesky factorizations of $Q_i(t)$ and $R_i(t)$, respectively. Letting

$$W_i \equiv \begin{pmatrix} L_{Q_i}^T & 0 \\ 0 & L_{R_i}^T \end{pmatrix}, \ i = 1, 2, ..., m$$

in (58)–(60), we obtain the problem equivalent to (73). In this case, we have

$$M = \begin{pmatrix} Q & 0 \\ 0 & R \end{pmatrix},$$

where

$$Q = \sum_{i=1}^m (t_i^2 - \|x_i\|^2) Q_i, \quad R = \sum_{i=1}^m (t_i^2 - \|x_i\|^2) R_i.$$

It is readily seen that (67) is satisfied here. The major part in computation of the Nesterov-Todd search direction is solution of (69) (with different v) to obtain ζ_i ($i = 0, ..., m+1$). Interestingly, this can be done as follows just by solving a matrix differential Riccati equation.

Let $v = (\gamma, \theta)$ in (69). Observe that the optimality condition is

$$M\zeta + v \in Z^\perp.$$

we are done if we can find $\zeta = (\alpha, \beta)$ satisfying the following condition:

$$Q\alpha + \gamma = -\dot{p} - A^T p, \ \alpha(0) = 0, \tag{74}$$

$$R\beta + \theta = -B^T p, \ p(T) = 0, \tag{75}$$

$$\dot{\alpha} = A\alpha + B\beta. \tag{76}$$

Let us try to find p in the form:

$$p = K\alpha + \rho, \ K(T) = 0, \ \rho(T) = 0. \tag{77}$$

where $K = K(t)$ is $n \times n$ matrix-valued function. Then

$$\dot{p} = \dot{K}\alpha + K\dot{\alpha} + \dot{\rho}. \tag{78}$$

Substituting this into (74), (75), we arrive at the following system of equations.

$$\dot{K} + A^T K + KA - KBR^{-1}B^T K + Q = 0, \;\; K(T) = 0, \tag{79}$$
$$\dot{\rho} + (A^T - KBR^{-1}B^T)\rho = -\gamma + KB\theta, \;\; \rho(T) = 0. \tag{80}$$

The system (79) is a matrix differential Riccati equation which admits a unique solution on the interval $[0, T]$ under natural control-theoretic constraints on the pair (A, B). To find $\zeta = (\alpha, \beta)$, we need to solve a linear system (80) and then find α and β using (74)–(78). Observe that the matrix differential Riccati equation (79) does not depend on $v = (\gamma, \theta)$, which is $\varepsilon_i v_i$ $(i = 0, 1, .., m+1)$ in our case. This means that (79) needs to be solved just once in one computation of the Nesterov-Todd direction, and (80) needs to be integrated $m + 2$ times.

6 Concluding Remarks

In the present paper we have considered infinite-dimensional generalization of interior-point algorithms using the framework of infinite-dimensional Jordan algebras of finite rank. Specifically, we developed a framework for primal-dual interior-point algorithms associated with the infinite-dimensional spin-factors and established a polynomial convergence result using the Nesterov-Todd direction. Though we have analyzed in detail only one primal-dual algorithm based on the Nesterov-Todd direction, it is pretty clear how to generalize other interior-point algorithms analyzed earlier in the finite-dimensional setting of Euclidean Jordan algebras.

We showed by considering an important example of a control problem that Nesterov-Todd direction can be calculated in an efficient way. Other popular directions (e.g., HRVW/KSH/M direction) can be analyzed in a similar fashion.

Acknowledgements

The authors would like to thank Prof. A. Nemirovski for his comment on the first version of this paper which motivated generalization of our original example in revision. This paper has been completed while the first author was a visiting professor in the Institute of Statistical Mathematics, Tokyo, Japan, and he wishes to thank Professors Takashi Tsuchiya, Kunio Tanabe and Yoshiaki Itoh for their hospitality.

Appendix: Derivation of (72)

First we observe that the functional which gives the optimal solution of (69) is linear with respect to v. Let M_Z be the restriction of superposition of the orthogonal projection to Z with M, to Z. Since M is positive definite on Z, there exists inverse of M_Z. We denote the inverse by M_Z^{-1}. Furthermore, let Π_Z be the orthogonal projection from X onto Z. Then the optimal solution of (69) is given as

$$\zeta = -M_Z^{-1}\Pi_Z v. \tag{81}$$

The optimality condition of (68) is

$$M\zeta + v_0 + \sum_{i=1}^{m+1} (v_i|\zeta)\varepsilon_i v_i \in Z^{\perp}.$$

Now, $(v_i|\zeta)$ is not yet known, but let δ_i be $(v_i|\zeta)$, and we continue as if we know δ. Then, we see that ζ is an optimal solution to (69) with

$$v = v_0 + \sum_{i=1}^{m+1} \delta_i \varepsilon_i v_i,$$

Due to (81), we see that the optimal solution of (69) is written as linear combination of the optimal solutions ζ_0 of (70) and ζ_i, $i = 1, ..., m+1$ of (71). Substituting $\zeta(\delta)$ into $(v_i|\zeta)$, we obtain

$$\delta_i = (v_i|\zeta(\delta)) \quad i = 1, ..., m+1. \tag{82}$$

This relation is obviously equivalent to (72), and is a necessary condition for $\zeta(\delta)$ to be the optimal solution of (68). Observe that such δ_i is ensured to exist due to solvability of (68). In the following, we show that (82) is sufficient for $\zeta(\delta)$ to be an optimal solution of (68). Let δ be the solution of (72) (and, equivalently, (82)). Due to the definition of ζ_i, we have

$$M\zeta(\delta) + v_0 + \sum_{i=1}^{m+1} \varepsilon_i v_i (v_i|\zeta(\delta)) = M\zeta_0 + v_0 + \sum_{i=1}^{m+1} \delta_i (M\zeta_i + \varepsilon_i v_i) \in Z^{\perp}.$$

This yields that $\zeta(\delta)$ is indeed the solution of (68).

Therefore, δ_i, $i = 1, ..., m+1$ determines the optimal solution of (68) if and only if (72) is satisfied. Since (68) has a unique optimal solution, (72) has a unique solution.

References

[1] J. Faraut and A. Koráyi: *Analysis on Symmetric Cones*, Oxford University Press, New York, 1994.

[2] Leonid Faybusovich: Euclidean Jordan algebras and interior-point algorithms. *Positivity*, Vol. 1 (1997), pp. 331-357.

[3] Leonid Faybusovich: Linear systems in Jordan algebras and primal-dual interior-point algorithms; *Journal of Computational and Applied Mathematics*, Vol. 86(1997), pp.149-175.

[4] Leonid Faybusovich: A Jordan-algebraic approach to potential-reduction algorithms. (To appear in Math. Zeitschr.)

[5] L. Faybusovich and R. Arana: A long-step primal-dual algorithm for the symmetric programming problem. *Systems & Control Letters*, Vol. 43 (2001), 3–7.

[6] L. Faybusovich and J. B. Moore: Infinite-dimensional quadratic optimization: interior-point methods and control applications. *Applied Mathematics and Optimization*, Vol. 36 (1997), pp. 43–66.

[7] L. Faybusovich and T. Mouktonglang: Finite-Rank Perturbation of the Linear-Quadratic control problem, preprint 2002.

[8] Wilhelm Kaup: Jordan algebras and holomorphy. *Functional analysis, holomorphy, and approximation theory* (Proc. Sem., Univ. Fed. Rio de Janeiro, Rio de Janeiro, 1978), pp. 341–365, *Lecture Notes in Mathematics*, Vol. 843, Springer, Berlin, 1981.

[9] R. D. C. Monteiro and T. Tsuchiya: Polynomial convergence of primal-dual algorithms for the second-order cone program based on the MZ-family of directions. Mathematical Programming, Vol. 88 (2000), No. 1, pp. 61–83.

[10] R. D. C. Monteiro and Y. Zhang: A unified analysis for a class of long-step primal-dual path-following interior-point algorithms for semidefinite programming. *Mathematical Programming*, Vol. 81 (1998), No. 3, pp. 281–299.

[11] Masakazu Muramatsu: On a commutative class of search directions for linear programming over symmetric cones. (To appear in *Jornal of Optimization Theory and its Applications*, Vol. 112, No.3(2002))

[12] A. Nemirovskii and K. Scheinberg: Extension of Karmarkar's algorithm onto convex quadratically constrained quadratic problems. Mathematical Programming, Vol. 72 (1996), No. 3, pp. 273–289.

[13] Y. E. Nesterov and M. Todd: Self-scaled barriers and interior-point methods for convex programming; *Mathematics of Operations Research*, Vol. 22 (1997), pp. 1-42.

[14] S. Schmieta and F. Alizadeh: Extensions of primal-dual interior point algorithms to symmetric cones. Report RRR 13-99, RUTCOR, Rutgers University, 640 Bartholomew Road, Piscataway, NJ 08854-8003, 1999.

[15] J. F. Sturm: Similarity and other spectral relations for symmetric cones. *Linear Algebra and its Applications*, Vol. 312 (2000), no. 1-3, pp. 135–154.

[16] T. Tsuchiya: A convergence analysis of the scaling-invariant primal-dual path-following algorithms for second-order cone programming. *Optimization Methods and Software*, Vol. 11&12 (1999), pp. 141–182.

[17] S. Wright: *Primal–dual Interior Point Algorithms*, SIAM Publications, Philadelphia, Pennsylvania, USA, 1997.