KYBERNETIKA — VOLUME *49* (2013), NUMBER 6, PAGES 883–896

NEW COMPLEXITY ANALYSIS OF A FULL NESTEROV–TODD STEP INFEASIBLE INTERIOR-POINT ALGORITHM FOR SYMMETRIC OPTIMIZATION

Behrouz Kheirfam and Nezam Mahdavi-Amiri

A full Nesterov–Todd step infeasible interior-point algorithm is proposed for solving linear programming problems over symmetric cones by using the Euclidean Jordan algebra. Using a new approach, we also provide a search direction and show that the iteration bound coincides with the best known bound for infeasible interior-point methods.

Keywords: interior-point methods, symmetric cone optimization, Euclidean Jordan algebra, polynomial complexity

Classification: 90C51

1. INTRODUCTION

In recent years, there have been extensive investigations concerning the analysis of interior-point methods (IPMs) for symmetric cone optimization (SCO). SCO includes solving problems such as linear optimization (LO), semidefinite optimization (SDO) and second-order cone optimization (SOCO). The foundation for solving these problems was laid by Nesterov and Nemirovskii [7]. Nesterov and Todd [8] proposed symmetric interior-point algorithms on a special class of cones called self-scaled cones. The early work connecting Jordan algebras and optimization is due to Güler [5]. He observed that the family of the self-scaled cones is identical to the set of symmetric cones for which there exists a complete classification theory. Faybusovich [2] first extended primal-dual IPMs for semidefinite optimization (SDO) to SCO by using Euclidean Jordan algebras. Muramatsu [6] presented a commutative class of search directions for SCO and analyzed the complexities of primal-dual IPMs for SCO. Rangarajan [9] proved the polynomial-time convergence of infeasible IPMs (IIPMs) for conic programming over symmetric cones using a wide neighborhood of the central path for a commutative family of search directions. Subsequently, Schmieta and Alizadeh [12] introduced primal-dual IPMs for SCO extensively under the framework of Euclidean Jordan algebra. In [10], an IIPM for LO was proposed by Roos. It differs from the classical IIPMs, since the new method uses only full steps which has the advantage that no line searches are needed. Recently, Gu et al. [4] have extended Roos' full-Newton step IIPM for LO to full Nesterov–Todd step (NT-step) IIPM for SCO by using Jordan algebras.

Motivated by the works mentioned above, we propose another version of Roos' algorithm for SCO which differs in the feasibility step. We show, by a simple analytical development, that the iteration bound coincides with the best-known iteration bound for IIPMs. In our approach, as in [4, 10], the size of the residual vectors reduces with the same speed as the duality gap.

The remainder of our work is organized as follows: In Section 2, we first briefly recall the properties of symmetric cones and its associated Euclidean Jordan algebras. Then, after giving the problem's background, we propose an algorithm with the modified feasibility steps. Section 3 is devoted to the analysis of the new feasibility steps, which is the main part of our work. There, the final iteration bound is also derived. Finally, we conclude in Section 4.

2. PRELIMINARIES

2.1. Euclidean Jordan algebra

Let V be an n-dimensional vector space over R. $(V, \circ, \langle ., . \rangle)$ is called an n-dimensional Euclidean Jordan algebra with rank r if there exists a bilinear map $\circ : V \times V \to V$ such that $x \circ y = y \circ x, x \circ (x^2 \circ y) = x^2 \circ (x \circ y)$, where $x^2 = x \circ x$, for all $x, y \in V$ with an inner product $\langle ., . \rangle$ which is associative. A Jordan algebra has an identity element, if there exists a unique element $e \in V$ such that $x \circ e = e \circ x = x$ holds, for all $x \in V$. The set $\mathcal{K} = \{x^2 : \ x \in V\}$ is called the cone of squares of Euclidean Jordan algebra $(V, \circ, \langle ., . \rangle)$. A cone is symmetric if and only if it is the cone of squares of some Euclidean Jordan algebra. An element $c \in V$ is idempotent if $c \circ c = c$. Two elements x and y are orthogonal if $x \circ y = 0$. An idempotent c is primitive if it is nonzero and can not be expressed by sum of two other nonzero idempotents. A set of primitive idempotents $\{c_1, c_2, \ldots, c_k\}$ is called a Jordan frame if $c_i \circ c_j = 0$, for any $i \neq j \in \{1, 2, \ldots, k\}$ and $\sum_{i=1}^{k} c_i = e$. For any $x \in V$, let r be the smallest positive integer such that $\{e, x, x^2, \cdots, x^r\}$ is linearly dependent; r is called the degree of x and is denoted by $deg(x)$. The rank of V, denoted by $rank(V)$, is defined as the maximum of $deg(x)$ over all $x \in V$. An element $x \in V$ is called regular iff its degree equals the rank of V. Since "$\circ$" is a bilinear map, for every $x \in V$, there exists a matrix $L(x)$ such that for every $y \in V, x \circ y = L(x)y$. For each $x \in V$, define

$$P(x) := 2L(x)^2 - L(x^2),$$

where, $L(x)^2 = L(x)L(x)$. The map $P(x)$ is called the quadratic representation of V.

For a regular element $x \in V$, since $\{e, x, x^2, \ldots, x^r\}$ is linearly dependent, there are real numbers $a_1(x), \ldots, a_r(x)$ such that the minimal polynomial of every regular element x is given

$$f(\lambda; x) = \lambda^r - a_1(x)\lambda^{r-1} + \ldots + (-1)^r a_r(x),$$

which is the characteristic polynomial of the regular element x. The coefficient $a_1(x)$ is called the trace of x, denoted as $\operatorname{tr}(x)$. The coefficient $a_r(x)$ is called the determinant of x, denoted as $\det(x)$.

Theorem 2.1. (Spectral decomposition, Theorem III.1.2 in [1]) Let $x \in V$. Then there exists a Jordan frame $\{c_1, c_2, \dots, c_r\}$ and real numbers $\lambda_1(x), \dots, \lambda_r(x)$ such that

$$x = \sum_{i=1}^{r} \lambda_i(x) c_i.$$

The numbers $\lambda_i(x)$ (with their multiplicities) are the eigenvalues of x. Furthermore,

$$\mathrm{tr}(x) = \sum_{i=1}^{r} \lambda_i(x), \quad \det(x) = \prod_{i=1}^{r} \lambda_i(x).$$

By using eigenvalues, we may extend the definition of any real valued, continuous univariate function to elements of a Euclidean Jordan algebra. Particularly, we have

Inverse: $x^{-1} := \sum_{i=1}^{r} \lambda_i^{-1}(x) c_i$, whenever all $\lambda_i \neq 0$ and undefined otherwise;

Square root: $x^{\frac{1}{2}} := \sum_{i=1}^{r} \lambda_i^{\frac{1}{2}}(x) c_i$, whenever all $\lambda_i \geq 0$ and undefined otherwise;

Square: $x^2 := \sum_{i=1}^{r} \lambda_i^2(x) c_i$.

The inner product $\langle \cdot, \cdot \rangle$ is defined by $\langle x, y \rangle = tr(x \circ y)$ for any $x, y \in V$. Thus, we can define norm on V by $\|x\|_F = \sqrt{tr(x^2)} = \sqrt{\sum_{i=1}^{r} \lambda_i(x)^2} = \|\lambda(x)\|$. Note that, since $e = c_1 + c_2 + \ldots + c_r$ has eigenvalue 1, with multiplicity r, it follows that $\mathrm{tr}(e) = r, \det(e) = 1$ and $\|e\|_F = \sqrt{r}$.

We say that two elements x and s in V are similar, denoted as $x \sim s$, if x and s share the same set of eigenvalues. We say $x \in \mathcal{K}$ if and only if $\lambda_i \geq 0$, for all $i = 1, 2, \cdots, r$, and $x \in \mathrm{int}\mathcal{K}$ if and only if $\lambda_i > 0$, for all $i = 1, 2, \cdots, r$.

For a comprehensive study on Jordan algebra and symmetric cones, the reader is referred to [1, 14]. Here, we outline some needed main results on Euclidean Jordan algebra and symmetric cones for the analysis of our algorithm.

Lemma 2.2. (Lemma 14 in [12]) Let $x, s \in V$. Then, the eigenvalues of $x + s$ are bounded as follows:

$$\lambda_{\min}(x + s) \geq \lambda_{\min}(x) - \|s\|_F, \quad \lambda_{\max}(x + s) \leq \lambda_{\max}(x) + \|s\|_F,$$

where $\lambda_{\min}(x)$ and $\lambda_{\max}(x)$ denote the smallest eigenvalue and the largest eigenvalue of x respectively.

Lemma 2.3. (Theorem III.2.1 in [1]) Let V be a Euclidean Jordan algebra. Then, $\mathcal{K}$ is a symmetric cone and is the set of elements x in V for which $L(x)$ is positive semidefinite. Furthermore, if x is invertible, then $P(x)\mathrm{int}\mathcal{K} = \mathrm{int}\mathcal{K}$.

Lemma 2.4. (Lemma 2.15 in [4]) If $x \circ s \in \mathrm{int}\mathcal{K}$, then $\det(x) \neq 0$.

Lemma 2.5. (Proposition 21 in [12]) Let $x, s, u \in \mathrm{int}\mathcal{K}$. Then,

(i) $P(x^{\frac{1}{2}})s \sim P(s^{\frac{1}{2}})x$.

(ii) $P\Big((P(u)x)^{\frac{1}{2}}\Big)P(u^{-1})s \sim P(x^{\frac{1}{2}})s$.

Lemma 2.6. (Proposition 3.2.4 in [14]) Let $x, s \in \mathrm{int}\mathcal{K}$, and w be the scaling point of x and s. Then, $\big(P(x^{\frac{1}{2}})s\big)^{\frac{1}{2}} \sim P(w^{\frac{1}{2}})s$.

Lemma 2.7. (Theorem 4 in [13]) Let $x, s \in \mathrm{int}\mathcal{K}$. Then,

$$\lambda_{\min}(P(x)^{\frac{1}{2}}s) \geq \lambda_{\min}(x \circ s).$$

Lemma 2.8. (Lemma 30 in [12]) Let $x, s \in \mathrm{int}\mathcal{K}$. Then,

$$\|P(x)^{\frac{1}{2}}s - e\|_F \leq \|x \circ s - e\|_F.$$

Lemma 2.9. (Lemma 2.9 in [9]) Given $x \in \mathrm{int}\mathcal{K}$, we have

$$\|x - x^{-1}\|_F \leq \frac{\|x^2 - e\|_F}{\lambda_{\min}(x)}.$$

2.2. Problem background

Let V be a Euclidean Jordan algebra of dimension n with rank r, and $\mathcal{K}$ be its associated cone of squares. Consider the following primal and dual problems:

$$\min \ \big\{\langle c, x\rangle : \ Ax = b, \quad x \in \mathcal{K}\big\}, \qquad \text{(SP)}$$

and

$$\max \ \big\{b^T y : \ A^T y + s = c, \quad s \in \mathcal{K}\big\}, \qquad \text{(SD)}$$

where, c and the rows of A lie in V, and $b \in R^m$. Without loss of generality, we assume that the rows of A are linearly independent.

As usual for IIPMs, we use the starting point as in [4] that one knows a positive scalar ξ such that $x^* + s^* \preceq_{\mathcal{K}} \xi e$ for some optimal solution (x^*, y^*, s^*) corresponding to (SP) and (SD) such that $tr(x^* \circ s^*) = 0$ and the initial iterates are $(x^0, y^0, s^0) = \xi(e, 0, e)$, with e being an identity element of V. Using $tr(x^0 \circ s^0) = r\xi^2$, the total number of iterations for the algorithm in [4] is bounded above by

$$16r \log \frac{\max\{r\xi^2, \|r_p^0\|_F, \|r_d^0\|_F\}}{\epsilon}, \tag{1}$$

where, $r_p^0 = b - Ax^0$ and $r_d^0 = c - A^T y^0 - s^0$, that is, r_p^0 and r_d^0 are the initial values of the primal and dual residuals.

To describe our aim, we recall the main ideas underlying the algorithm in [4]. For any ν with $0 < \nu \leq 1$, we consider the perturbed problem

$$\min \ \big\{\langle c - \nu r_d^0, x\rangle : \ \ b - Ax = \nu r_p^0, \ \ x \in \mathcal{K}\big\}, \qquad (P_\nu)$$

and its dual

$$\max \ \big\{(b - \nu r_p^0)^T y : \ \ c - A^T y - s = \nu r_d^0, \ \ s \in \mathcal{K}\big\}. \qquad (D_\nu)$$

Note that if $\nu = 1$, then $x = x^0$ is a strictly feasible solution of (P_ν) and $(y, s) = (y^0, s^0)$ is a strictly feasible solution of (D_ν). We conclude that if $\nu = 1$, then (P_ν) and (D_ν) are strictly feasible, meaning that both perturbed problems (P_ν) and (D_ν) satisfy the well-known interior-point condition (IPC). More generally, one has the following result (Lemma 4.1 in [4]).

Lemma 2.10. Let (SP) and (SD) be feasible and $0 < \nu \leq 1$. Then, the perturbed problems (P_ν) and (D_ν) satisfy the IPC.

Assuming that (SP) and (SD) are feasible, it follows from Lemma 2.10 that the problems (P_ν) and (D_ν) satisfy the IPC, for each $0 < \nu \leq 1$. Then, their central paths exist, meaning that the system

$$b - Ax = \nu r_p^0, \qquad x \in \mathcal{K}, \tag{2}$$

$$c - A^T y - s = \nu r_d^0, \qquad s \in \mathcal{K}, \tag{3}$$

$$x \circ s = \mu e, \tag{4}$$

has a unique solution, for any $\mu > 0$. This solution is denoted as $(x(\mu,\nu), y(\mu,\nu), s(\mu,\nu))$. These are the μ-centers of the perturbed problems (P_ν) and (D_ν). Now, if $\mu = \nu\mu^0$, then the solution should be denoted by $(x(\nu), y(\nu), s(\nu))$. With this notation, one has $(x(1), y(1), s(1)) = (x^0, y^0, s^0) = (\xi e, 0, \xi e)$ with $\mu^0 = \xi^2$. We measure the proximity of iterate (x, y, s) to the μ-centers of the perturbed problems (P_ν) and (D_ν) by the quantity

$$\delta(x, s; \mu) := \delta(v) := \tfrac{1}{2}\|v^{-1} - v\|_F = \tfrac{1}{2}\sqrt{\sum_{i=1}^{r}\left(\lambda_i^{-1}(v) - \lambda_i(v)\right)^2}, \tag{5}$$

where

$$v := \frac{P(w)^{-\frac{1}{2}}x}{\sqrt{\mu}} \;\Big[= \frac{P(w)^{\frac{1}{2}}s}{\sqrt{\mu}}\Big], \tag{6}$$

and $w = P(x)^{\frac{1}{2}}\left(P(x)^{\frac{1}{2}}s\right)^{-\frac{1}{2}} \left[= P(s)^{-\frac{1}{2}}\left(P(s)^{\frac{1}{2}}x\right)^{\frac{1}{2}}\right]$ is the Nesterov–Todd scaling point of x and s (see Lemma 3.2 in [3]).

Initially, we have $\delta(x, s; \mu) = 0$. In the sequel, we assume that at the start of each iteration, $\delta(x, s; \mu)$ is smaller than or equal to a threshold value $\tau > 0$. This certainly holds at the start of the first iteration.

We now describe one main iteration of the algorithm given in [4]. The algorithm begins with an infeasible interior-point (x, y, s) such that (x, y, s) is feasible for the perturbed problems (P_ν) and (D_ν), $tr(x \circ s) = r\mu$ and $\delta(x, s; \mu) \leq \tau$. Each main iteration consists of a feasibility step and a few centering steps. The feasibility step serves to get iterates (x^f, y^f, s^f) that are strictly feasible for (P_{ν^+}) and (D_{ν^+}) with $\nu^+ := (1-\theta)\nu$, $\theta \in (0, 1)$, and such that $\delta(x^f, s^f; \mu^+) \leq \frac{1}{\sqrt{2}}$, i.e., (x^f, y^f, s^f) lies in the quadratic convergence neighborhood with respect to the μ^+-center of (P_{ν^+}) and (D_{ν^+}). Then, by performing a few centering steps, starting from (x^f, y^f, s^f) and targeting at the μ^+-center of (P_{ν^+}) and (D_{ν^+}), we get the iterate (x^+, y^+, s^+) that is feasible for (P_{ν^+}) and (D_{ν^+}) and such that $\mathrm{tr}(x^+ \circ s^+) = \mu^+ r$ and $\delta(x^+, s^+; \mu^+) \leq \tau$. This process is repeated until the norms of the residuals and $\mathrm{tr}(x \circ s)$ are less than some prescribed accuracy parameter ϵ.

For the feasibility step, the search directions $\Delta^f x, \Delta^f y$ and $\Delta^f s$ are defined by the system

$$A\Delta^f x = \theta\nu r_p^0 \tag{7}$$

$$A^T\Delta^f y + \Delta^f s = \theta\nu r_d^0 \tag{8}$$

$$P(w^{-\frac{1}{2}})x \circ P(w^{\frac{1}{2}})\Delta^f s + P(w^{\frac{1}{2}})s \circ P(w^{-\frac{1}{2}})\Delta^f x = (1-\theta)\mu e - P(w^{-\frac{1}{2}})x \circ P(w^{\frac{1}{2}})s, \tag{9}$$

where, $\theta \in (0,1)$.

After a feasibility step, the new iterates

$$x^f = x + \triangle^f x, \quad y^f = y + \Delta^f y, \quad s^f = s + \triangle^f s,$$

are still interior points, i. e., $x^f, s^f \in \mathrm{int}\mathcal{K}$, since θ is small enough and $\delta(x^f, s^f; \mu^+) \leq \frac{1}{\sqrt{2}}$. In the centering step, starting at the iterates $(x, y, s) = (x^f, y^f, s^f)$ and targeting at the μ-centers, the search directions $\Delta x, \Delta y$ and Δs are the usual Nesterov–Todd (NT) directions, defined by

$$\begin{aligned} A\Delta x &= 0 \\ A^T\Delta y + \Delta s &= 0 \\ P(w^{-\frac{1}{2}})x \circ P(w^{\frac{1}{2}})\Delta s + P(w^{\frac{1}{2}})s \circ P(w^{-\frac{1}{2}})\Delta x &= \mu e - P(w^{-\frac{1}{2}})x \circ P(w^{\frac{1}{2}})s. \end{aligned}$$

Denoting the components of the iterate after a centering step by x^+, y^+ and s^+, we recall the following result [4].

Lemma 2.11. If $\delta(v) < 1$, then the full Nesterov–Todd step is strictly feasible, i. e., x^+ and s^+ are positive, and $\langle x^+, s^+\rangle = \mu tr(e)$. Moreover, if $\delta(v) \leq \frac{1}{\sqrt{2}}$, then $\delta(v^+) \leq \delta(v)^2$.

Now, define

$$d_x^f := \frac{P(w)^{-\frac{1}{2}}\Delta^f x}{\sqrt{\mu}}, \qquad d_s^f := \frac{P(w)^{\frac{1}{2}}\Delta^f s}{\sqrt{\mu}}. \tag{10}$$

The system defining the search directions $\Delta^f x, \Delta^f y$ and $\Delta^f s$ can be expressed in terms of the scaled directions d_x^f and d_s^f as follows

$$\begin{aligned} \bar{A}d_x^f &= \theta\nu r_p^0, \\ \bar{A}^T\frac{\Delta y}{\mu} + d_s^f &= \tfrac{1}{\sqrt{\mu}}\theta\nu P(w)^{\frac{1}{2}}r_d^0, \\ d_x^f + d_s^f &= (1-\theta)v^{-1} - v, \end{aligned} \tag{11}$$

where, $\bar{A} = \sqrt{\mu}AP(w)^{\frac{1}{2}}$.

The main contribution of our work here is a modification of the feasibility step. We present a different algorithm, obtained by changing the definition of the feasibility step via replacing the third equation of (11) by $d_x^f + d_s^f = 0$. We will see this simplifies the analysis of the algorithm, whereas the iteration bound essentially remains the same. We now give a more formal description of Algorithm 1 below.

Algorithm 1: A full Nesterov–Todd step IIPM for SCO

Input :
accuracy parameter $\epsilon > 0$,
barrier update parameter θ, $0 < \theta < 1$,
and threshold parameter $0 < \tau \leq \frac{1}{\sqrt{2}}$.
begin
$x := \xi e$; $y := 0$; $s := \xi e$; $\mu := \mu^0 = \xi^2$; $\nu = 1$;
while $\max\left(r\mu, \|r_p\|_F, \|r_d\|_F\right) > \epsilon$ **do**
feasibility step :
$(x, s, y) := (x, s, y) + (\Delta^f x, \Delta^f s, \Delta^f y)$;
$\mu -$ update :
$\mu := (1 - \theta)\mu$;
centering step :
while $\delta(x, s; \mu) \geq \tau$ **do**
$(x, y, s) := (x, y, s) + (\Delta x, \Delta y, \Delta s)$
end while
end while
end.

---

3. ANALYSIS OF THE FEASIBILITY STEP

Let x, y and s denote the components of the iterate at the start of an iteration, with $tr(x \circ s) = r\mu$ and $\delta(x, s; \mu) \leq \tau$. These certainly hold at the start of the first iteration, because we have $tr(x^0 \circ s^0) = r\xi^2$ and $\delta(x^0, s^0; \mu^0) = 0$.

3.1. Effect of the feasibility step

According to Lemma 2.11, we need to show that $\delta(x^f, s^f; \mu^+) \leq \frac{1}{\sqrt{2}}$ after the feasibility step, that is, the new iterates are within the region where the NT process targeting at the μ^+-centers of (P_{ν^+}) and (D_{ν^+}) is quadratically convergent. Using (10), we have

$$\begin{aligned} x^f &= x + \Delta^f x = \sqrt{\mu}(P(w)^{\frac{1}{2}}(v + d_x^f), \\ s^f &= s + \Delta^f s = \sqrt{\mu}P(w)^{-\frac{1}{2}}(v + d_s^f). \end{aligned} \tag{12}$$

Therefore, by $d_x^f + d_s^f = 0$, we have

$$(v + d_x^f) \circ (v + d_s^f) = v^2 + v \circ (d_x^f + d_s^f) + d_x^f \circ d_s^f = v^2 - (d_x^f)^2. \tag{13}$$

Since $P(w)^{\frac{1}{2}}$ and $P(w)^{-\frac{1}{2}}$ are automorphisms of $\mathrm{int}\mathcal{K}$ (Lemma 2.3), x^f and s^f belong to $\mathrm{int}\mathcal{K}$ if and only if $v + d_x^f$ and $v + d_s^f$ belong to $\mathrm{int}\mathcal{K}$. The next lemma gives conditions for strict feasibility of the full NT-step.

Lemma 3.1. The iterate (x^f, y^f, s^f) is strictly feasible if $v \pm d_x^f \in \mathrm{int}\mathcal{K}$.

Proof. Introduce a step length α with $\alpha \in [0,1]$ and define

$$v_x^\alpha = v + \alpha d_x^f, \qquad v_s^\alpha = v - \alpha d_x^f.$$

We thus have $v_x^0 = v, v_s^0 = v, v_x^1 = v + d_x^f$ and $v_s^1 = v - d_x^f$. It follows that

$$v_x^\alpha \circ v_s^\alpha = (v + \alpha d_x^f) \circ (v - \alpha d_x^f) = v^2 - (\alpha d_x^f)^2. \tag{14}$$

If $v \pm d_x^f \in \text{int}\mathcal{K}$, then we have $(d_x^f)^2 \prec_\mathcal{K} v^2$. Substituting this into (14), we get

$$v_x^\alpha \circ v_s^\alpha \succ_\mathcal{K} (1 - \alpha^2) v^2.$$

Since $v \in \text{int}\mathcal{K}$, we have

$$v_x^\alpha \circ v_s^\alpha \succ_\mathcal{K} 0.$$

By Lemma 2.4, it follows that $\det(v_x^\alpha) \neq 0$ and $\det(v_s^\alpha) \neq 0$, for $\alpha \in [0,1]$. Since $\det(v_x^0) = \det(v_s^0) = \det(v) > 0$, by continuity, $\det(v_x^\alpha)$ and $\det(v_s^\alpha)$ stay positive, for all $\alpha \in [0,1]$. Moreover, by Theorem 2.1, this implies that all the eigenvalues of v_x^α and v_s^α stay positive for all $\alpha \in [0,1]$. Hence, we can conclude that all the eigenvalues of v_x^1 and v_s^1 are nonnegative. Therefore, $v \pm d_x^f \in \text{int}\mathcal{K}$, completing the proof. □

Lemma 3.2. (Lemma II.60 in [11]) If $\delta := \delta(v)$ is defined as (5), then

$$\frac{1}{\rho(\delta)} \leq \lambda_i(v) \leq \rho(\delta), \ i = 1, 2, \ldots, r,$$

where, $\rho(\delta) := \delta + \sqrt{1 + \delta^2}$.

Assuming $v \pm d_x^f \succ_\mathcal{K} 0$, which according to Lemma 3.1 implies that the iterate (x^f, y^f, s^f) is strictly feasible, we proceed by deriving an upper bound for $\delta(x^f, s^f; \mu^+)$. Let w^f be the scaling point of x^f and s^f. Denoting the v-vector after the feasibility step with respect to the μ^+-center as v^f, we have, according to (5) and (6),

$$\delta(x^f, s^f; \mu^+) := \delta(v^f) := \frac{1}{2} \|(v^f)^{-1} - v^f\|_F, \tag{15}$$

where,

$$v^f := \frac{P(w^f)^{-\frac{1}{2}} x^f}{\sqrt{\mu(1-\theta)}} \Big[= \frac{P(w^f)^{\frac{1}{2}} s^f}{\sqrt{\mu(1-\theta)}} \Big]. \tag{16}$$

Lemma 3.3. (Lemma 3.4 in [4]) One has

$$(v^f)^2 \sim \frac{P(v + d_x^f)^{\frac{1}{2}} (v + d_s^f)}{1 - \theta}.$$

Lemma 3.4. We have

$$\lambda_{\min}\Big((v^f)^2\Big) \geq \frac{1}{1-\theta} \Big(\frac{1}{\rho(\delta)^2} - \|d_x^f\|_F^2 \Big).$$

P r o o f. By using Lemmas 3.3, 2.7, (13), Lemma 2.2 and Lemma 3.2, we get

$$\begin{aligned}
\lambda_{\min}\big((v^f)^2\big) &= \frac{1}{1-\theta}\lambda_{\min}\Big(P(v+d_x^f)^{\frac{1}{2}}(v+d_s^f\Big) \\
&\geq \frac{1}{1-\theta}\lambda_{\min}\Big((v+d_x^f)\circ(v+d_s^f)\Big) \\
&= \frac{1}{1-\theta}\lambda_{\min}\Big(v^2-(d_x^f)^2\Big) \\
&\geq \frac{1}{1-\theta}\Big(\lambda_{\min}(v^2)-\|d_x^f\|_F^2\Big) \\
&\geq \frac{1}{1-\theta}\Big(\frac{1}{\rho(\delta)^2}-\|d_x^f\|_F^2\Big).
\end{aligned}$$

This completes the proof. □

Lemma 3.5. We have

$$\|e-(v^f)^2\|_F \leq \frac{\sqrt{r}\theta+2(1+\rho(\delta))\delta+\|d_x^f\|_F^2}{1-\theta}.$$

P r o o f. By using Lemmas 3.3, 2.8 and properties of the Frobenius norm, we have

$$\begin{aligned}
\|e-(v^f)^2\|_F &= \Big\|e-P\big(\frac{v+d_x^f}{\sqrt{1-\theta}}\big)^{\frac{1}{2}}\big(\frac{v+d_s^f}{\sqrt{1-\theta}}\big)\Big\|_F \\
&\leq \Big\|\big(\frac{v+d_x^f}{\sqrt{1-\theta}}\big)\circ\big(\frac{v+d_s^f}{\sqrt{1-\theta}}\big)-e\Big\|_F \\
&= \Big\|\frac{v^2-(d_x^f)^2}{1-\theta}-e\Big\|_F \\
&\leq \frac{1}{1-\theta}\big(\sqrt{r}\theta+\|e-v^2\|_F+\|d_x^f\|_F^2\big), \qquad (17)
\end{aligned}$$

where, the last inequality follows from the triangle inequality. On the other hand, using (5), the inequality $|1-t|\leq|t^{-1}-t|$, for all $t>0$, and Lemma 3.2, we obtain

$$\|e-v^2\|_F \leq \|e-v\|_F+\|v-v^2\|_F \leq 2\delta+\lambda_{\max}(v)2\delta \leq 2\delta(1+\rho(\delta)). \qquad (18)$$

Substituting the bound of (18) into (17), the result follows. □

Using lemmas 3.4, 3.5 and 2.9, we have the following result.

Lemma 3.6. We have

$$2\delta(v^f) \leq \frac{\rho(\delta)\big(\sqrt{r}\theta+2(1+\rho(\delta))\delta+\|d_x^f\|_F^2\big)}{\sqrt{(1-\theta)\big(1-\rho(\delta)^2\|d_x^f\|_F^2\big)}}.$$

Needing $\delta(v^f) \leq \frac{1}{\sqrt{2}}$, it follows from Lemma 3.6 that it is sufficient to have

$$\frac{\rho(\delta)\big(\sqrt{r}\theta + 2(1+\rho(\delta))\delta + \|d_x^f\|_F^2\big)}{\sqrt{(1-\theta)\big(1-\rho(\delta)^2\|d_x^f\|_F^2\big)}} \leq \sqrt{2}.$$

At this stage, we choose

$$\tau = \frac{1}{16}, \quad \theta = \frac{\alpha}{2\sqrt{r}}, \quad \alpha \leq 1. \tag{19}$$

Then for $r \geq 2$ and $\delta \leq \tau$ we have $\rho(\delta) \leq \rho(\tau) \leq \frac{11}{10}$ and $1-\theta \geq 1 - \frac{1}{2\sqrt{r}} \geq \frac{3}{5}$. In this case, we obtain

$$\frac{\rho(\delta)\big(\sqrt{r}\theta + 2(1+\rho(\delta))\delta + \|d_x^f\|_F^2\big)}{\sqrt{(1-\theta)\big(1-\rho(\delta)^2\|d_x^f\|_F^2\big)}} \leq \frac{\frac{671}{800} + \frac{11}{10}\|d_x^f\|_F^2}{\sqrt{\frac{3}{5}\big(1-\frac{121}{100}\|d_x^f\|_F^2\big)}}.$$

It is sufficient to have

$$\frac{\frac{671}{800} + \frac{11}{10}\|d_x^f\|_F^2}{\sqrt{\frac{3}{5}\big(1-\frac{121}{100}\|d_x^f\|_F^2\big)}} \leq \sqrt{2},$$

which implies

$$\|d_x^f\|_F^4 + 2.725\|d_x^f\|_F^2 - 0.4103 \leq 0.$$

The above inequality follows that

$$\|d_x^f\|_F \leq \frac{1}{2\sqrt{2}}. \tag{20}$$

3.2. An upper bound for $\|d_x^f\|_F$

Here, we obtain an upper bound for $\|d_x^f\|_F$, which enables us to find a default value for θ. Consider the system

$$\begin{aligned} \bar{A}d_x^f &= \theta\nu r_p^0, \\ \bar{A}^T\frac{\Delta y}{\mu} + d_s^f &= \tfrac{1}{\sqrt{\mu}}\theta\nu P(w)^{\frac{1}{2}} r_d^0, \\ d_x^f + d_s^f &= 0. \end{aligned} \tag{21}$$

By $\xi := -\frac{\Delta y}{\mu}$ and eliminating d_s^f, we get

$$\begin{aligned} \bar{A}d_x^f &= \theta\nu r_p^0, \\ \bar{A}^T\xi + d_x^f &= -\tfrac{1}{\sqrt{\mu}}\theta\nu P(w)^{\frac{1}{2}} r_d^0. \end{aligned} \tag{22}$$

Multiplying both sides of the second equation in (22) from the left with $\bar{A}$ and using the first equation of (22), it follows that

$$\xi = (\bar{A}\bar{A}^T)^{-1}\Big(-\theta\nu r_p^0 - \frac{1}{\sqrt{\mu}}\theta\nu\bar{A}P(w)^{\frac{1}{2}} r_d^0\Big). \tag{23}$$

Substitution (23) into the second equation of (22), we get

$$d_x^f = -\frac{1}{\sqrt{\mu}}\theta\nu\Big(I - \bar{A}^T(\bar{A}\bar{A}^T)^{-1}\bar{A}\Big)P(w)^{\frac{1}{2}}r_d^0 + \theta\nu\bar{A}^T(\bar{A}\bar{A}^T)^{-1}r_p^0.$$

Let (x^*, y^*, s^*) be the optimal solution satisfying $x^* + s^* \preceq_{\mathcal{K}} \xi e$. Then, we may write

$$r_p^0 = A(x^* - x^0), \quad r_d^0 = A^T(y^* - y^0) + (s^* - s^0).$$

Substituting the expressions for r_p^0 and r_d^0 into the expression for d_x^f, we obtain

$$\begin{aligned} d_x^f =& -\frac{1}{\sqrt{\mu}}\theta\nu\Big(I - \bar{A}^T(\bar{A}\bar{A}^T)^{-1}\bar{A}\Big)P(w)^{\frac{1}{2}}\big(A^T(y^* - y^0) + s^* - s^0\big) \\ &+ \frac{\theta\nu}{\sqrt{\mu}}\bar{A}^T(\bar{A}\bar{A}^T)^{-1}\bar{A}P(w)^{-\frac{1}{2}}(x^* - x^0) \\ =& -\frac{1}{\sqrt{\mu}}\theta\nu\Big(I - \bar{A}^T(\bar{A}\bar{A}^T)^{-1}\bar{A}\Big)P(w)^{\frac{1}{2}}(s^* - s^0) \\ &+ \frac{\theta\nu}{\sqrt{\mu}}\bar{A}^T(\bar{A}\bar{A}^T)^{-1}\bar{A}P(w)^{-\frac{1}{2}}(x^* - x^0), \end{aligned}$$

the last equality follows by using $I - \bar{A}^T(\bar{A}\bar{A}^T)^{-1}\bar{A}$ is the orthogonal projection to the null space of $\frac{1}{\sqrt{\mu}}AP(w)^{\frac{1}{2}}$. On the the hand, $\frac{\theta\nu}{\sqrt{\mu}}\bar{A}^T(\bar{A}\bar{A}^T)^{-1}\bar{A}P(w)^{-\frac{1}{2}}(x^* - x^0)$ is the orthogonal projection of $\frac{\theta\nu}{\sqrt{\mu}}P(w)^{-\frac{1}{2}}(x^* - x^0)$ onto the row space of $\bar{A}$. Hence, by the triangle inequality, it follows that

$$\|d_x^f\|_F \leq \frac{\theta\nu}{\sqrt{\mu}}\Big(\|P(w)^{\frac{1}{2}}(s^* - s^0)\|_F + \|P(w)^{-\frac{1}{2}}(x^* - x^0)\|_F\Big). \tag{24}$$

Using $0 \preceq_{\mathcal{K}} x^* \preceq_{\mathcal{K}} x^* + s^* \preceq_{\mathcal{K}} \xi e$, $0 \preceq_{\mathcal{K}} s^* \preceq_{\mathcal{K}} x^* + s^* \preceq_{\mathcal{K}} \xi e$ and the iterate $(x^0, y^0, s^0) = (\xi e, 0, \xi e)$, we have

$$0 \preceq_{\mathcal{K}} x^0 - x^* \preceq_{\mathcal{K}} \xi e, \ 0 \preceq_{\mathcal{K}} s^0 - s^* \preceq_{\mathcal{K}} \xi e.$$

By using almost the same arguments made in [4], we obtain

$$\|P(w)^{\frac{1}{2}}(s^* - s^0)\|_F \leq \xi\sqrt{tr(w^2)} \leq \xi\frac{\sqrt{tr(x^2)}}{\sqrt{\mu}\lambda_{\min}(v)} \leq \frac{\xi tr(x)}{\sqrt{\mu}\lambda_{\min}(v)}. \tag{25}$$

Similarly, we have

$$\|P(w)^{-\frac{1}{2}}(x^* - x^0)\|_F \leq \xi\sqrt{tr(w^{-2})} \leq \xi\frac{\sqrt{tr(s^2)}}{\sqrt{\mu}\lambda_{\min}(v)} \leq \frac{\xi tr(s)}{\sqrt{\mu}\lambda_{\min}(v)}. \tag{26}$$

Substituting (25) and (26) into (24) and using $\mu = \nu\xi^2$ and Lemma 3.2, we obtain

$$\|d_x^f\|_F \leq \frac{\theta\nu\xi tr(x + s)}{\mu\lambda_{\min}(v)} = \frac{\theta tr(x + s)}{\xi\lambda_{\min}(v)} \leq \frac{\theta\rho(\delta)}{\xi}tr(x + s). \tag{27}$$

Let x and (y, s) be feasible for the problems (P_ν) and (D_ν), respectively, and $tr(x \circ s) = \mu r$. With (x^0, y^0, s^0) and ξ as defined above, we have $tr(x + s) \leq 2\xi r$ (Lemma 4.6 in [4]). Therefore,

$$\|d_x^f\|_F \leq 2r\theta\rho(\delta).$$

Since $\delta \leq \tau = \frac{1}{16}$ and $\rho(\delta)$ is monotonically increasing in δ, we have

$$\|d_x^f\|_F \leq 2r\theta\rho(\frac{1}{16}) = 2.129r\theta.$$

By using $\theta = \frac{\alpha}{2\sqrt{r}}$, the above inequality turns to

$$\|d_x^f\|_F \leq 1.0645\sqrt{r}\alpha. \tag{28}$$

In order to have $\delta(v^f) \leq \frac{1}{\sqrt{2}}$, by (20), we should have $\|d_x^f\|_F \leq \frac{1}{2\sqrt{2}}$. The last inequality is true if

$$1.0645\sqrt{r}\alpha \leq \frac{1}{2\sqrt{2}}.$$

This means if we take

$$\alpha = \frac{1}{3.02\sqrt{r}}, \tag{29}$$

it is guaranteed that $\delta(v^f) \leq \frac{1}{\sqrt{2}}$.

3.3. Complexity

We have seen that if at the start of an iteration the iterate satisfies $\delta(x, s; \mu) \leq \tau$, with τ as defined in (19), then after the feasibility step, with θ as defined in (19), the iterate satisfies $\delta(x^f, s^f; \mu^+) \leq \frac{1}{\sqrt{2}}$.

The centering steps serve to get iterates which satisfy $tr(x^+ \circ s^+) = \mu^+ r$ and $\delta(x^+, s^+; \mu^+) \leq \tau$, where τ is much smaller than $\frac{1}{\sqrt{2}}$. By using Lemma 2.11, the required number of centering steps can easily be obtained. This goes as follows. After the μ update, we have $\delta(x^f, s^f; \mu^+) \leq \frac{1}{\sqrt{2}}$, and hence after k centering steps, the iterate (x^+, y^+, s^+) satisfies

$$\delta(x^+, s^+; \mu^+) \leq \left(\frac{1}{\sqrt{2}}\right)^{2^k}.$$

From this, one easily deduces that $\delta(x^+, s^+; \mu^+) \leq \tau$ will hold after at most

$$\left\lceil \log_2 \left(\log_2 \frac{1}{\tau^2} \right) \right\rceil, \tag{30}$$

centering steps.

According to (30), at most

$$\log_2 \left(\log_2 \frac{1}{\tau^2} \right) = \log_2(\log_2 256) = 3$$

centering steps suffice to get the iterate (x^+, y^+, s^+) to satisfy $\delta(x^+, s^+; \mu^+) \leq \tau$. So, each iteration consists of at most four so-called 'inner' iterations, at each of which we need to compute a new search direction. In each main iteration, both the duality gap and the norms of the residual vectors are reduced by the factor $1 - \theta$. Hence, using $tr(x^0 \circ s^0) = r\xi^2$, the total number of the main iterations is bounded above by

$$\frac{1}{\theta} \log \frac{\max\{r\xi^2, \|r_p^0\|_F, \|r_d^0\|_F\}}{\epsilon}.$$

Due to (19) and (29), we have

$$\theta = \frac{1}{6.04r}.$$

Hence, the total number of inner iterations is bounded above by

$$24.16\ r \log \frac{\max\{r\xi^2, \|r_p^0\|_F, \|r_d^0\|_F\}}{\epsilon}.$$

This bound coincides with the currently best available bound for IIPMs applied to SCO. The above explanation leads to the following result which proves the polynomial iteration complexity of the algorithm.

Theorem 3.7. If $\theta = \dfrac{1}{6.04r}$, then the number of iterations of the infeasible primal-dual algorithm with full Nesterov–Todd steps does not exceed

$$24.16\ r \log \frac{\{r\xi^2, \|r_p^0\|_F, \|r_d^0\|_F\}}{\epsilon}.$$

4. CONCLUSIONS

We analyzed an infeasible interior-point algorithm with a full Nesterov–Todd step for symmetric cone optimization, modifying the feasibility step of the algorithm given in [4]. This modification tendered a simple analysis to establish the polynomial iteration complexity of the algorithm.

ACKNOWLEDGEMENT

The second author thanks Sharif University of Technology for supporting its support. The authors are also thankful to two anonymous referees for their useful comments to improve the presentation.

(Received August 22, 2012)

REFERENCES

[1] J. Faraut and A. Korányi: Analysis on Symmetric Cones. Oxford Mathematical Monographs, The Clarendon Press Oxford University Press, New York 1994.

[2] L. Faybusovich: Linear systems in Jordan algebras and primal-dual interior-point algorithms. J. Comput. Appl. Math. *86* (1997), 149–175.

[3] L. Faybusovich: A Jordan-algebraic approach to potential-reduction algorithms. Math. Z. *239* (2002), 1, 117–129.

[4] G. Gu, M. Zangiabadi, and C. Roos: Full Nesterov–Todd step infeasible interior-point method for symmetric optimization. European J. Oper. Res. *214* (2011), 3, 473–484.

[5] O. Güler: Barrier functions in interior-point methods. Math. Oper. Res. *21* (1996), 4, 860–885.

[6] M. Muramatsu: On a commutative class of search directions for linear programming over symmetric cones. J. Optim. Theory Appl. *112* (2002), 3, 595–625.

[7] Y. E. Nesterov and A. Nemirovski: Interior-point Polynomial Algorithms in Convex Programming. SIAM, Philadelphia 1994.

[8] Y. E. Nesterov and M. J. Todd: Self-scaled barriers and interior-point methods for convex programming. Math. Oper. Res. *22* (1997), 1, 1–42.

[9] B. K. Rangarajan: Polynomial convergence of infeasible-interior-point methods over symmetric cones. SIAM J. Optim. *16* (2006), 4, 1211–1229.

[10] C. Roos: A full-Newton step $O(n)$ infeasible interior-point algorithm for linear optimization. SIAM J. Optim. *16* (2006), 4, 1110–1136.

[11] C. Roos, T. Terlaky, and J.-Ph. Vial: Theory and Algorithms for Linear Optimization: An Interior-Point Approach. John Wiley and Sons, Chichester 1997, Second edition Springer 2006.

[12] S. H. Schmieta and F. Alizadeh: Extension of primal-dual interior-point algorithm to symmetric cones. Math. Program. *96* (2003), 3, 409–438.

[13] J. F. Sturm: Similarity and other spectral relations for symmetric cones. Algebra Appl. *312* (2000), 1 – 3, 135–154.

[14] M. V. C. Vieira: Jordan Algebraic Approach to Symmetric Optimization. Ph.D. Thesis, Electrical Engineering, Mathematics and Computer Science, Delft University of Technology 2007.

Behrouz Kheirfam, Department of Mathematics, Azarbaijan Shahid Madani University, Tabriz. Iran.

e-mail: b.kheirfam@azaruniv.edu

Nezam Mahdavi-Amiri, Department of Mathematical Sciences, Sharif University of Technology, Tehran. Iran.

e-mail: nezamm@sharif.edu