June 6, 1996

ON WEIGHTED CENTERS FOR SEMIDEFINITE PROGRAMMING

Jos F. Sturm[1] and Shuzhong Zhang[2]

Report 9636/A, Econometric Institute
Erasmus University Rotterdam
P.O. Box 1738, 3000 DR Rotterdam, The Netherlands

ABSTRACT

In this paper, we generalize the notion of weighted centers to semidefinite programming. Our analysis fits in the v-space framework, which is purely based on the symmetric primal-dual transformation and does not make use of barriers. Existence and scale invariance properties are proven for the weighted centers. Relations with other primal-dual maps are discussed.

Key words. semidefinite programming, symmetric primal–dual transformation, weighted center.

[1] Econometric Institute, Erasmus University Rotterdam, The Netherlands, sturm@few.eur.nl.

[2] Econometric Institute, Erasmus University Rotterdam, The Netherlands, zhang@few.eur.nl

1. Introduction

The central path plays a fundamental role in the interior point methodology, both for linear and semidefinite programming. Megiddo [10] showed some highly interesting properties of the central path for linear programming. The fact that μ-centers are the minimizers of the logarithmic barrier function with parameter μ plays a crucial role in his analysis. With the introduction of primal-dual interior point methods by Kojima, Mizuno and Yoshise [6], it became more popular to define μ-centers as the solutions of the perturbed KKT-systems

$$x_i z_i = \mu \text{ for } i = 1, 2, \ldots, n,$$

where $x \in \Re^n$ denotes the primal nonnegative decision variable and $z \in \Re^n$ is the corresponding dual slack variable.

The analysis of the central path for semidefinite programming evolves in a similar way. Namely, early definitions of the central path are based on the logarithmic barrier (Nesterov and Nemirovsky [13], Vandenberghe and Boyd [17]), whereas Kojima, Shindoh and Hara [7] proposed a general framework for primal-dual path following algorithms for semidefinite programming, in which the μ-centers are defined as solutions of the perturbed KKT-systems

$$XZ = \mu I,$$

where $X \in \Re^{n\times n}$ ($Z \in \Re^{n\times n}$) is the primal (dual) positive definite decision variable, and I is the order n identity matrix. With this definition, some basic properties of the central path were obtained by Luo, Sturm and S. Zhang [8]. Independently, the central path was investigated by Goldfarb and Scheinberg [2] based on the logarithmic barrier definition.

In linear programming, the concept of *weighted centers* has played an increasingly important role in recent years, largely due to the unifying works of Kojima et al. [5] and Jansen et al. [4]. The weighted center for linear programming can be defined as the minimizer of a W-weighted logarithmic barrier function or alternatively as the solution of the perturbed KKT system

$$x_i z_i = w_i \text{ for } i = 1, 2, \ldots, n$$

for some $w \in \Re^n_+$.

Due to the existence of the weighted centers, it has been popular to consider the iterates in a homogenized product space of x and z, viz. the v-space [4, 5]. This proves to be an elegant and convenient way of analyzing various primal-dual interior point algorithms for linear programming. However, the generalization of this concept to semidefinite programming is not straightforward. The major obstacle is the lack of a proper weighted barrier function for semidefinite programming.

In this paper, we propose to define weighted centers in terms of the eigenvalues of the matrix XZ. It is shown that if the primal-dual Slater condition holds, then any set of n positive eigenvalues can be attained by a feasible primal-dual solution pair (X, Z). Moreover, this notion of weighted center is *scale-invariant* (see Todd, Toh and Tütüncü [16] for a definition). A different generalization of weighted centers towards semidefinite programming was recently proposed by Monteiro and Pang [12]. Their weighted center has the property of uniqueness, but lacks some other desirable properties as we will discuss later.

This paper is organized as follows. In Section 2, we review some terminology and we define the concept of weighted center for semidefinite programming. We will also discuss how Newton's method can be used to derive search directions that can be used in an interior point method for semidefinite programming. The quality of these search direction is estimated in Section 3. We show in Section 4 that for any positive definite matrix W, there exists a W-weighted center under the primal-dual Slater condition. In Section 5, we discuss the new definition of weighted centers in relation to the Monteiro-Pang weighted centers [12].

Notation.
Given X and Y in $\Re^{n\times n}$, the standard inner product is defined by

$$X \bullet Y = \mathrm{tr} X^{\mathrm{T}} Y.$$

The Euclidean norm and its associated operator norm, viz. the spectral norm, are both denoted by $\|\cdot\|$. The Frobenius norm of a matrix $X \in \Re^{n\times n}$ is $\|X\|_F = \sqrt{X \bullet X}$. It is well known that if $\sigma_1, \sigma_2, \ldots, \sigma_n$ are the singular values of an $n \times n$ matrix X, then

$$\|X\|_F^2 = \sum_{i=1}^n \sigma_i^2, \quad \|X\| = \max_{1\le i\le n} \sigma_i.$$

Hence, $\|X\| \le \|X\|_F$. We will also make use of the obvious inequality for $n \times n$ matrices X, Y, Z that

$$\|XYZ\|_F \le \|X\| \, \|Y\|_F \, \|Z\| .$$

For any $X \in \Re^{n\times n}$ and invertible $Y \in \Re^{n\times n}$ there holds

$$\mathrm{eig}(X) = \mathrm{eig}(YXY^{-1}), \tag{1}$$

where $\mathrm{eig}(X)$ denotes the set of eigenvalues of X.

The space of symmetric $n \times n$ matrices is denoted by $\mathcal{S}$. Its orthoplement in $\Re^{n\times n}$, viz. the space of skew symmetric $n \times n$ matrices, is denoted by $\mathcal{S}^\perp$. For given $X \in \Re^{n\times n}$ we let

$$P_{\mathcal{S}}(X) := \frac{X + X^{\mathrm{T}}}{2}, P_{\mathcal{S}^\perp}(X) := \frac{X - X^{\mathrm{T}}}{2}$$

denote the orthogonal projections of X on $\mathcal{S}$ and $\mathcal{S}^\perp$ respectively. If $X \in \mathcal{S}$ is positive (semi-) definite, we write $X \succ 0$ ($X \succeq 0$). The cone of positive semi-definite matrices is denoted by

$\mathcal{S}_+$ and the cone of positive definite matrices is $\mathcal{S}_{++}$. For $X \in \mathcal{S}_+$ we let $\lambda_{\min}(X)$ denote the smallest eigenvalue of X. The order n identity matrix is denoted by I.

2. The semidefinite programming problem and transformations

2.1. Problem statement

Consider the primal-dual pair of semidefinite programming problems in conic formulation

$$\min\{C_2 \bullet X \mid X - C_1 \in \mathcal{A}, X \succeq 0\}, \tag{2}$$

$$\min\{C_1 \bullet Z \mid Z - C_2 \in \mathcal{A}^\perp, Z \succeq 0\}, \tag{3}$$

where $C_1 \in \mathcal{S}, C_2 \in \mathcal{S}$, $\mathcal{A}$ is a linear subspace of $\mathcal{S}$ with orthogonal complement $\mathcal{A}^\perp$. The semidefinite matrix X is the primal decision variable, and Z is the dual decision variable. The problems (2) and (3) satisfy the weak duality relation

$$0 \leq X \bullet Z = C_1 \bullet Z + C_2 \bullet X - C_1 \bullet C_2$$

for any feasible pair (X, Z). The above relation shows that solving (2) and (3) is equivalent to minimizing the duality gap $X \bullet Z$ over all feasible primal-dual pairs (X, Z).

We make the following assumption:

Assumption 1. *There exist solutions $X \succ 0$ and $Z \succ 0$ such that $X - C_1 \in \mathcal{A}$ and $Z - C_2 \in \mathcal{A}^\perp$.*

It is well known that with the above assumption (a primal-dual Slater condition), the pair (2)-(3) is equivalent to finding a complementary solution pair, which is by definition a feasible solution pair (X, Z) with $X \bullet Z = 0$ (i.e. no duality gap), see [1, 13, 9] among others. It is convenient to combine (2)-(3) into the following formulation:

$$\begin{array}{lll} (SDP) & \min & X \bullet Z \\ & s.t. & X - C_1 \in \mathcal{A} \\ & & Z - C_2 \in \mathcal{A}^\perp \\ & & X \succeq 0, \quad Z \succeq 0. \end{array}$$

Notice that the set of complementary solution pairs for (2)-(3) is the optimal solution set of (SDP).

2.2. Primal-dual transformations

Primal-dual transformations for semidefinite programming are based on the observation that given an invertible matrix $L \in \Re^{n\times n}$, we have

$$X \succeq 0 \text{ if and only if } L^{-1}XL^{-\mathrm{T}} \succeq 0. \tag{4}$$

The relation (4) implies that (SDP) is equivalent to the linearly transformed semidefinite programming problem

$$\min\{\bar{X} \bullet \bar{Z} \mid L\bar{X}L^{\mathrm{T}} - C_1 \in \mathcal{A}, L^{-\mathrm{T}}\bar{Z}L^{-1} - C_2 \in \mathcal{A}^{\perp}, \bar{X} \succeq 0, \bar{Z} \succeq 0\}. \tag{5}$$

Let

$$\mathcal{A}(L) := \{X \in \mathcal{S} \mid LXL^{\mathrm{T}} \in \mathcal{A}\}$$

and remark that the orthogonal complement of $\mathcal{A}(L)$ in $\mathcal{S}$ is

$$\mathcal{A}(L)^{\perp} = \{Z \in \mathcal{S} \mid L^{-\mathrm{T}}ZL^{-1} \in \mathcal{A}^{\perp}\}.$$

The set of feasible primal-dual pairs (X, Z) for (5) is denoted by

$$\mathcal{F}(L) := \{(X, Z) \in \mathcal{S}_{+} \times \mathcal{S}_{+} \mid X - L^{-1}C_1L^{-\mathrm{T}} \in \mathcal{A}(L), Z - L^{\mathrm{T}}C_2L \in \mathcal{A}(L)^{\perp}\}.$$

We can now formulate (SDP) and (5) as

$$\min\{X \bullet Z \mid (X, Z) \in \mathcal{F}(I)\} = \min\{\bar{X} \bullet \bar{Z} \mid (\bar{X}, \bar{Z}) \in \mathcal{F}(L)\} = 0.$$

We let $\overset{o}{\mathcal{F}}(L)$ denote the set of strictly feasible (or *interior*) solutions, i.e.

$$\overset{o}{\mathcal{F}}(L) := \mathcal{F}(L) \cap (\mathcal{S}_{++} \times \mathcal{S}_{++}).$$

Consider an interior solution pair $(X, Z) \in \overset{o}{\mathcal{F}}(I)$. It is shown in [15] that if L_d is an invertible $n \times n$ matrix such that

$$L_dL_d^{\mathrm{T}} = Z^{-1/2}(Z^{1/2}XZ^{1/2})^{1/2}Z^{-1/2}, \tag{6}$$

then

$$L_d^{-1}XL_d^{-\mathrm{T}} = L_d^{\mathrm{T}}ZL_d$$

and vice versa. The subscript 'd' in L_d is reminiscent to the standard notation used for primal-dual interior point methods in linear programming, where d denotes a primal-dual scaling vector. The transformation

$$X \to L_d^{-1}XL_d^{-\mathrm{T}}, \quad Z \to L_d^{\mathrm{T}}ZL_d$$

is known as the symmetric primal-dual transformation, since it maps both X and Z into the same positive definite matrix, viz.

$$V = L_d^{-1}XL_d^{-\mathrm{T}} = L_d^{\mathrm{T}}ZL_d \tag{7}$$

and we have

$$(V, V) \in \mathcal{F}(L_d).$$

2.3. Definition of weighted centers

Consider $(X,Z) \in \overset{o}{\mathcal{F}}(I)$ and let Λ_{XZ} be a diagonal matrix with the eigenvalues of the matrix XZ on its diagonal. Since X and Z are both positive definite, Λ_{XZ} must be a positive diagonal matrix, see (1). Moreover, the duality gap is the sum of the eigenvalues of XZ,

$$X \bullet Z = \mathrm{tr} \Lambda_{XZ}.$$

If we want to approach the optimal solution set of (SDP) from the interior solution set $\overset{o}{\mathcal{F}}(I)$, we see that the corresponding eigenvalue matrix Λ_{XZ} has to approach the origin (the matrix of all zeros) from the set of positive diagonal matrices. A μ-center [17] is a pair $(X(\mu), Z(\mu)) \in \overset{o}{\mathcal{F}}(I)$ such that $\Lambda_{XZ}(\mu) = \mu I$. It can easily be verified, e.g. by using (1), that

$$\Lambda_{XZ}(\mu) = \mu I \text{ if and only if } X(\mu)Z(\mu) = \mu I,$$

which means that $X(\mu)$ and $Z(\mu)$ commute in this case. It is known that the *central path*, defined as the set

$$\{(X(\mu), Z(\mu)) \mid \mu > 0\}$$

is a smooth curve [7]. Path-following algorithms generate a sequence of approximate μ-centers with $\mu \downarrow 0$.

A *W-weighted center* can now be defined as a pair $(X,Z) \in \overset{o}{\mathcal{F}}(I)$ such that $\Lambda_{XZ}^{1/2} = W$ for some positive diagonal matrix W. Unlike the μ-center case however, X and Z are in general *not* commutable and $\Lambda_{XZ}^{1/2} = W$ does *not* imply $XZ = W^2$. Moreover, there is no *unique* path

$$\{(X(t), Z(t)) \in \overset{o}{\mathcal{F}}(I) \mid \Lambda_{XZ}(t) = tW^2,\, t > 0\}, \tag{8}$$

as can be seen from the following example, with $n = 2$, where the set (8) is a 2-dimensional surface.

Example 2.1. *Consider the semidefinite programming problem with data*

$$C_1 = C_2 = I, \mathcal{A} = \emptyset, \mathcal{A}^{\perp} = \mathcal{S}.$$

We are interested in trajectories satisfying (8) with

$$W^2 = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}.$$

For any smooth function $\xi : \Re_+ \to [-1,1]$, the trajectory

$$X(t) = I, Z(t) = t \begin{bmatrix} 2 + \sqrt{1-\xi(t)^2} & \xi(t) \\ \xi(t) & 2 - \sqrt{1-\xi(t)^2} \end{bmatrix}$$

is such a trajectory.

However, if $W = \sqrt{\mu} I$, i.e. the μ-center case, then the pair $(X(\mu), Z(\mu))$ is unique. This fact follows from the strict convexity of the barrier function (see [2, 7]), but can also easily be shown using the symmetric primal-dual transformation:

Lemma 2.1. *Suppose* $(\sqrt{\mu} I, \sqrt{\mu} I) \in \overset{o}{\mathcal{F}} (L_d)$. *If* $(\bar{X}, \bar{Z}) \in \overset{o}{\mathcal{F}} (L_d)$ *and*

$$\bar{X}\bar{Z} = \mu I$$

then $\bar{X} = \bar{Z} = \sqrt{\mu} I$.

Proof:
Let

$$D_X := \frac{1}{2}(\bar{X} - \sqrt{\mu} I), \quad D_Z := \frac{1}{2}(\bar{Z} - \sqrt{\mu} I).$$

Then

$$\begin{aligned} \mu I = \bar{X}\bar{Z} &= (\sqrt{\mu} I + 2D_X)(\sqrt{\mu} I + 2D_Z) \\ &= (\sqrt{\mu} I + D_X + D_Z + (D_X - D_Z))(\sqrt{\mu} I + D_X + D_Z - (D_X - D_Z)) \\ &= (\sqrt{\mu} I + D_X + D_Z)^2 - (D_X - D_Z)^2 \\ &\quad + (D_X - D_Z)(\sqrt{\mu} I + D_X + D_Z) - (\sqrt{\mu} I + D_X + D_Z)(D_X - D_Z). \end{aligned}$$

Using the fact that the above matrix is symmetric, it follows that

$$\mu I = (\sqrt{\mu} I + D_X + D_Z)^2 - (D_X - D_Z)^2 \preceq (\sqrt{\mu} I + D_X + D_Z)^2$$

and hence

$$D_X + D_Z \succeq 0. \tag{9}$$

However, using $D_X \perp D_Z$ and $\bar{X}\bar{Z} = \mu I$, we have

$$n\mu = \operatorname{tr}(\bar{X}\bar{Z}) = n\mu + 2\sqrt{\mu} \operatorname{tr}(D_X + D_Z)$$

i.e. $\operatorname{tr}(D_X + D_Z) = 0$. Together with (9), this implies that $D_X + D_Z = 0$. From the orthogonality $D_X \perp D_Z$, we further obtain $D_X = D_Z = 0$. This concludes the proof.

□

2.4. The V-space

Consider an interior solution pair $(X, Z) \in \overset{o}{\mathcal{F}} (I)$. From (7) we have

$$V^2 = (L_d^{-1} X L_d^{-\mathrm{T}})(L_d^{\mathrm{T}} Z L_d) = L_d^{-1} X Z L_d,$$

which shows that the eigenvalues of V^2 are identical to the eigenvalues of XZ, see (1). In fact, we may choose L_d in such a way that $V = \Lambda_{XZ}^{1/2}$. Namely, such an L_d can be computed by the following procedure (see Todd, Toh and Tütüncü [16]):

1. Compute Cholesky factorization L_X,

$$X = L_X L_X^{\mathrm{T}}.$$

2. Compute eigenvector-eigenvalue decomposition (Q, Λ_{XZ}),

$$L_X^{\mathrm{T}} Z L_X = Q^{\mathrm{T}} \Lambda_{XZ} Q.$$

3. Let

$$L_d = L_X Q \Lambda_{XZ}^{-1/4}, \quad V = \Lambda_{XZ}^{1/2}.$$

The weighted center with respect to a positive definite matrix W can now be characterized as a pair $(X, Z) \in \overset{o}{\mathcal{F}}(I)$ with

$$V = L_d^{-1} X L_d^{-\mathrm{T}} = L_d^{\mathrm{T}} Z L_d = W. \tag{10}$$

With this characterization, we can define weighted centers for any positive definite W, not necessarily diagonal. Similarly, having Λ_{XZ} approach 0 from the set of positive diagonal matrices corresponds to letting V approach 0 from the cone of positive definite matrices.
In this paper, we are interested in computing a W-center (10) for given $W \in \mathcal{S}_{++}$. However, the Newton equation for solving (10) is underdetermined, due to the nonuniqueness of W-centers as illustrated by Example 2.1. In the sequel, we will make the Newton direction well-defined by adding restrictions on the choice of L_d.
Consider a trajectory $(\bar{X}(t), \bar{Z}(t)) \in \overset{o}{\mathcal{F}}(L_d)$, with $\bar{X}(0) = V, \bar{Z}(0) = V$ and let $G(t)$ be the positive definite matrix defined by

$$\bar{X}(t) = G(t)^2 \bar{Z}(t) G(t)^2.$$

We let

$$V(t) = G(t)\bar{Z}(t)G(t) = G(t)^{-1}\bar{X}(t)G(t)^{-1}. \tag{11}$$

Obviously, $G(0) = I$ and $V(0) = V$. Noticing that

$$(V(t), V(t)) \in \mathcal{F}(L_d G(t)),$$

we see that $V(t)$ is both primal and dual feasible if we use the transformation $L_d G(t)$. Let

$$D_X = \frac{1}{2}\frac{\mathrm{d}\bar{X}(t)}{\mathrm{d}t}\,|_{t=0}, \quad D_Z = \frac{1}{2}\frac{\mathrm{d}\bar{Z}(t)}{\mathrm{d}t}\,|_{t=0}\,.$$

Using (11), we can now linearize $V(t)$ as follows:

$$\begin{aligned} V(t) &= \frac{1}{2}\left[\; G(t)^{-1}\bar{X}(t)G(t)^{-1} + G(t)\bar{Z}(t)G(t) \;\right] \\ &= V + t(D_X + D_Z) + o(t). \end{aligned} \tag{12}$$

From the requirement $(\bar{X}(t), \bar{Z}(t)) \in \overset{o}{\mathcal{F}}(L_d)$, we further obtain

$$D_X \in \mathcal{A}(L_d), \quad D_Z \in \mathcal{A}^{\perp}(L_d). \tag{13}$$

The linearization of $V(t)$ will be denoted by $F(t)$. Letting

$$D_v = D_X + D_Z, \tag{14}$$

it follows from (12) that

$$F(t) := V + tD_v. \tag{15}$$

Remark that (D_X, D_Z) is the orthogonal decomposition of the matrix D_v onto $\mathcal{A}(L_d)$ and $\mathcal{A}(L_d)^{\perp}$ respectively. We have derived that the Newton direction for solving the nonlinear equation $V(1) = W$ is the unique solution (D_X, D_Z) of (13)-(14) with $D_v = W - V$. From now on, we let $(\bar{X}(t), \bar{Z}(t))$ denote the solution pair in $\overset{o}{\mathcal{F}}(L_d)$ that is obtained by a partial Newton step:

$$\bar{X}(t) = V + 2tD_X, \quad \bar{Z}(t) = V + 2tD_Z.$$

In the next section, we will estimate the error term $\|F(t) - V(t)\|_F$ that emerges by linearizing $V(t)$.

3. A bound on higher order terms

In this section, we will estimate the error term $\|F(t) - V(t)\|_F$ of the linearization $F(t)$ of $V(t)$.

We see that

$$\bar{X}(t) = \frac{\bar{X}(t) + \bar{Z}(t)}{2} + \frac{\bar{X}(t) - \bar{Z}(t)}{2}, \quad \bar{Z}(t) = \frac{\bar{X}(t) + \bar{Z}(t)}{2} - \frac{\bar{X}(t) - \bar{Z}(t)}{2} \tag{16}$$

and

$$\frac{\bar{X}(t) + \bar{Z}(t)}{2} = F(t), \quad \frac{\bar{X}(t) - \bar{Z}(t)}{2} = t(D_X - D_Z), \tag{17}$$

where $F(t)$ is defined by (15). Since $D_X \perp D_Z$, we have

$$\|D_X - D_Z\|_F = \|D_v\|_F. \tag{18}$$

For $t \geq 0$ with $F(t) \succ 0$, we define

$$\rho(t) := \frac{t\,\|D_v\|_F}{\lambda_{\min}(F(t))}. \tag{19}$$

The next lemma shows that $\bar{X}(t)$ and $\bar{Z}(t)$ are interior feasible solutions if $\rho(t) < 1$.

Lemma 3.1. *Let $\bar{t} \geq 0$ be such that $F(\bar{t}) \succ 0$. Then*

$$\bar{X}(t) \succeq (1-\rho(\bar{t}))\lambda_{\min}(F(\bar{t}))I$$

and

$$\bar{Z}(t) \succeq (1-\rho(\bar{t}))\lambda_{\min}(F(\bar{t}))I$$

for all $t \in [0, \bar{t}]$. If $\rho(\bar{t}) < 1$ then $V(t)$ exists and is positive definite for $t \in [0, \bar{t}]$.

Proof:
By definition (15), we have

$$F(t) = F(\bar{t}) - (\bar{t} - t)D_v$$

where $D_v = F(1) - V$. Suppose $0 \leq t \leq \bar{t}$. Using (19), it follows that

$$\lambda_{\min}(F(t)) \geq (1 - (1 - \frac{t}{\bar{t}})\rho(\bar{t}))\lambda_{\min}(F(\bar{t})).$$

Using (16)-(18), we have

$$\begin{aligned} \bar{X}(t) = F(t) + t(D_X - D_Z) &\succeq F(t) - t\left\|D_v\right\|_F I \\ &\succeq (1-\rho(\bar{t}))\lambda_{\min}(F(\bar{t}))I \end{aligned}$$

and similarly

$$\bar{Z}(t) \succeq (1-\rho(\bar{t}))\lambda_{\min}(F(\bar{t}))I.$$

This shows that if $\rho(\bar{t}) < 1$ then $\bar{X}(t)$ and $\bar{Z}(t)$ are positive definite and hence $V(t)$ is well defined.

$\square$

The above lemma provides a sufficient condition for the existence of $V(t)$. We will now examine the difference between $V(t)$ and its linearization $F(t)$.

Lemma 3.2. *Let $t \geq 0$ be such that $F(t) \succ 0$ and $\rho(t) < 1$. There holds*

$$F(t) - V(t) = \frac{1}{2}P_{\mathcal{S}}\left[\ (I - G(t))(\bar{Z}(t) - V(t))(I + G(t))\ \right].$$

Proof:
First remark using (11) that

$$\bar{Z}(t) - V(t) = \bar{Z}(t) - G(t)\bar{Z}(t)G(t) = P_{\mathcal{S}}\left[\ (I - G(t))\bar{Z}(t)(I + G(t))\ \right]$$

and

$$\bar{X}(t) - V(t) = G(t)V(t)G(t) - V(t) = -P_{\mathcal{S}}\left[\ (I - G(t))V(t)(I + G(t))\ \right].$$

Adding the above two relations yields

$$\frac{\bar{X}(t)+\bar{Z}(t)}{2} - V(t) = \frac{1}{2} P_{\mathcal{S}} \left[\ (I-G(t))(\bar{Z}(t)-V(t))(I+G(t)) \ \right].$$

$\square$

Based on the above lemma, it is natural to bound $\|F(t)-V(t)\|_F$ by deriving estimates for $\|I-G(t)\|$ and $\|\bar{Z}(t)-V(t)\|_F$. Such estimates are given by Lemma 3.3 and Lemma 3.4 below.

Lemma 3.3. *Let $t \geq 0$ be such that $F(t) \succ 0$ and $\rho(t) < 1$. There holds*

$$\frac{1-\rho(t)}{1+\rho(t)} I \preceq G(t)^4 \preceq \frac{1+\rho(t)}{1-\rho(t)} I$$

and therefore

$$\|G(t)-I\| \leq \left[\ \tfrac{1+\rho(t)}{1-\rho(t)} \ \right]^{1/4} - 1.$$

Proof:
Remark that

$$\begin{aligned} (\bar{Z}(t)^{1/2} G(t)^2 \bar{Z}(t)^{1/2})^2 &= \bar{Z}(t)^{1/2} \bar{X}(t) \bar{Z}(t)^{1/2} \\ &= \bar{Z}(t)(I + 2t\bar{Z}(t)^{-1/2}(D_X - D_Z)\bar{Z}(t)^{-1/2})\bar{Z}(t). \end{aligned} \tag{20}$$

From Lemma 3.1 and relation (18) we have

$$\left\| \bar{Z}(t)^{-1/2}(D_X - D_Z)\bar{Z}(t)^{-1/2} \right\| \leq \frac{\|D_X - D_Z\|}{(1-\rho(t))\lambda_{\min}(F(t))} \leq \frac{\|D_v\|_F}{(1-\rho(t))\lambda_{\min}(F(t))}.$$

Hence, using (20),

$$\begin{aligned} (\bar{Z}(t)^{1/2} G(t)^2 \bar{Z}(t)^{1/2})^2 &\preceq (1 + \frac{2t\,\|D_v\|_F}{(1-\rho(t))\lambda_{\min}(F(t))})\bar{Z}(t)^2 \\ &= \frac{1+\rho(t)}{1-\rho(t)} \bar{Z}(t)^2, \end{aligned}$$

which implies that

$$G(t)^4 \preceq \frac{1+\rho(t)}{1-\rho(t)} I.$$

Using the primal-dual symmetry, it follows that also

$$G(t)^{-4} \preceq \frac{1+\rho(t)}{1-\rho(t)} I.$$

$\square$

Lemma 3.4. *Let $t \geq 0$ be such that $F(t) \succ 0$ and $\rho(t) < 4/5$. There holds*

$$\|\bar{Z}(t) - V(t)\|_F \leq \frac{2\sqrt{1-\rho(t)}}{3\sqrt{1-\rho(t)} - \sqrt{1+\rho(t)}} t \|D_v\|_F .$$

Proof:
We have

$$\bar{Z}(t) - V(t) + G(t)(\bar{Z}(t) - V(t))G(t) = \bar{Z}(t) - \bar{X}(t) = 2t(D_Z - D_X),$$

from which we obtain

$$\begin{aligned} 2(\bar{Z}(t) - V(t)) &= 2t(D_Z - D_X) + 2P_{\mathcal{S}}\left[\ (G(t) - I)(V(t) - \bar{Z}(t))\ \right] \\ &\quad +(G(t) - I)(V(t) - \bar{Z}(t))(G(t) - I). \end{aligned}$$

Applying the triangle inequality, it further follows that

$$2\|\bar{Z}(t) - V(t)\|_F \leq 2t\|D_v\|_F + (2\|G(t) - I\| + \|G(t) - I\|^2)\|\bar{Z}(t) - V(t)\|_F$$

or, equivalently,

$$(3 - (1 + \|G(t) - I\|)^2)\|\bar{Z}(t) - V(t)\|_F \leq 2t\|D_v\|_F . \tag{21}$$

Remark now that Lemma 3.3 implies

$$(1 + \|G(t) - I\|)^4 \leq \frac{1+\rho(t)}{1-\rho(t)}. \tag{22}$$

Combining (21) and (22) we obtain for $\rho(t) < 4/5$ that

$$\|\bar{Z}(t) - V(t)\|_F \leq \frac{2\sqrt{1-\rho(t)}}{3\sqrt{1-\rho(t)} - \sqrt{1+\rho(t)}} t \|D_v\|_F .$$

□

Combining Lemma 3.2, 3.3 and 3.4, we obtain the following estimate for the error term $\|F(t) - V(t)\|_F$.

Lemma 3.5. *Let $t \geq 0$ be such that $F(t) \succ 0$ and $\rho(t) < 2/3$. Then*

$$\|F(t) - V(t)\|_F \leq \frac{\rho(t)t\|D_v\|_F}{2 - 3\rho(t)}.$$

Proof:
From Lemma 3.3 we have

$$\begin{aligned} \|I + G(t)\|\,\|I - G(t)\| &\leq (1 + \frac{(1+\rho(t))^{1/4}}{(1-\rho(t))^{1/4}})(\frac{(1+\rho(t))^{1/4}}{(1-\rho(t))^{1/4}} - 1) \\ &= \frac{\sqrt{1+\rho(t)}}{\sqrt{1-\rho(t)}} - 1. \end{aligned} \tag{23}$$

Combining Lemma 3.2 and Lemma 3.4 with (23) we obtain

$$\begin{aligned}
\|F(t)-V(t)\|_F &\leq \frac{1}{2}\|I+G(t)\|\,\|\bar{Z}(t)-V(t)\|_F\,\|I-G(t)\| \\
&\leq \frac{\sqrt{1+\rho(t)}-\sqrt{1-\rho(t)}}{3\sqrt{1-\rho(t)}-\sqrt{1+\rho(t)}}t\,\|D_v\|_F \\
&= \frac{\sqrt{1+2\rho(t)/(1-\rho(t))}-1}{3-\sqrt{1+2\rho(t)/(1-\rho(t))}}t\,\|D_v\|_F\,. \qquad (24)
\end{aligned}$$

Remark now that

$$\sqrt{1+2\frac{\rho(t)}{1-\rho(t)}} \leq \sqrt{1+2\frac{\rho(t)}{1-\rho(t)}+\frac{\rho(t)^2}{(1-\rho(t))^2}} = 1+\frac{\rho(t)}{1-\rho(t)},$$

which together with (24) implies for $\rho(t)<2/3$ that

$$\|F(t)-V(t)\|_F \leq \frac{\rho(t)t\,\|D_v\|_F}{2-3\rho(t)}.$$

□

As a corollary to Lemma 3.5, we obtain a quadratic convergence result:

Corollary 3.1 (q-quadratic convergence). *Suppose $F(1)\succ 0$. If $\rho(1)<2/3$ then*

$$\|F(1)-V(1)\|_F \leq \frac{\rho(1)}{2-3\rho(1)}\|F(1)-V\|_F = \frac{\|F(1)-V\|_F^2}{(2-3\rho(1))\lambda_{\min}(F(1))}.$$

4. Weighted centers

We have already seen that given $(X^{(0)},Z^{(0)})\in\overset{o}{\mathcal{F}}(I)$, one can compute $L_d^{(0)}$ such that $(V^{(0)},V^{(0)})\in\mathcal{F}(L_d^{(0)})$ for some positive definite matrix $V^{(0)}$. We will now show that for any positive definite matrix $W\in\mathcal{S}_{++}$, there exists an invertible matrix L such that $(W,W)\in\mathcal{F}(L)$. This fact will be shown by construction. To be more precise, we will show that the following algorithm produces a sequence $L_d^{(k)}$ such that $L=\lim_{k\to\infty}L_d^{(k)}$ and $W=\lim_{k\to\infty}V^{(k)}$, if it is properly initialized.

Algorithm 1 (Weighted center).
Input: weight $W \in \mathcal{S}_{++}$ *and initial invertible* $L_d^{(0)}$ *and positive definite* $V^{(0)}$ *such that* $(V^{(0)}, V^{(0)}) \in \overset{o}{\mathcal{F}} (L_d^{(0)})$ *and* $\left\| W - V^{(0)} \right\|_F \leq \frac{1}{3}\lambda_{\min}(W)$.

Step 0 *Let* $k = 0$.

Step 1 *Let*

$$D_v^{(k)} = W - V^{(k)}$$

and decompose it into $D_X^{(k)} + D_Z^{(k)} = D_v^{(k)}$ *such that* $D_X^{(k)} \in \mathcal{A}^{\perp}(L_d^{(k)})$ *and* $D_Z^{(k)} \in \mathcal{A}(L_d^{(k)})$.

Step 2 *Compute*

$$\bar{X}^{(k)}(1) = V^{(k)} + 2D_X^{(k)}, \quad \bar{Z}^{(k)}(1) = V^{(k)} + 2D_Z^{(k)}.$$

Step 3 *Let* $G^{(k)}$ *be positive definite and satisfy*

$$\bar{X}^{(k)}(1) = (G^{(k)})^2 \bar{Z}^{(k)}(1) (G^{(k)})^2.$$

Let $L_d^{(k+1)} = L_d^{(k)} G^{(k)}$. *Compute*

$$V^{(k+1)} = G^{(k)} \bar{Z}^{(k)}(1) G^{(k)}.$$

Step 4 *Set* $k = k + 1$ *and return to Step 1.*

In each iteration $k = 0, 1, 2, \ldots$ of Algorithm 1, a full Newton step is made towards W, i.e.

$$F^{(k)}(1) = V^{(k)} + D_v^{(k)} = W. \tag{25}$$

Lemma 4.1. *Consider Algorithm 1. If*

$$\left\| W - V^{(0)} \right\|_F \leq \frac{1}{3}\lambda_{\min}(W) \tag{26}$$

then for any $k \in \{0, 1, 2, \ldots\}$ *there holds*

$$\left\| W - V^{(k+1)} \right\|_F \leq \frac{1}{\lambda_{\min}(W)} \left\| W - V^{(k)} \right\|_F^2 \leq \frac{1}{3} \left\| W - V^{(k)} \right\|_F \tag{27}$$

and hence

$$\lim_{k \to \infty} V^{(k)} = W.$$

Proof:
By definition (19), we have

$$\rho^{(k)}(1) = \frac{\left\| D_v^{(k)} \right\|_F}{\lambda_{\min}(F^{(k)}(1))} = \frac{\left\| W - V^{(k)} \right\|_F}{\lambda_{\min}(W)}.$$

Relation (26) is therefore equivalent to $\rho^{(0)}(1) \leq 1/3$. Applying Corollary 3.1, we obtain (27).

□

We can now prove the existence of all W-weighted centers in a neighborhood of $V^{(0)}$.

Lemma 4.2. *Let $W \in \mathcal{S}_{++}$ and $(V^{(0)}, V^{(0)}) \in \mathcal{F}(L_d^{(0)})$ for some invertible matrix $L_d^{(0)} \in \Re^{n\times n}$. If*

$$\left\| W - V^{(0)} \right\|_F \leq \frac{1}{3}\lambda_{\min}(W)$$

then there exists an invertible matrix $L \in \Re^{n\times n}$ such that

$$(W, W) \in \mathcal{F}(L).$$

Proof:
Initialize Algorithm 1 with $L_d^{(0)}$ and $V^{(0)}$. We like to show that $\lim_{k\to\infty} L_d^{(k)}$ exists. To this end, we notice that

$$L_d^{(k+1)} = L_d^{(k)} G^{(k)}$$

for all $k \geq 0$. Now we have from Lemma 3.3,

$$\|G^{(k)} - I\| \leq [\frac{1+\rho^{(k)}}{1-\rho^{(k)}}]^{1/4} - 1$$

and

$$\rho^{(k+1)} \leq \frac{1}{3}\rho^{(k)}$$

for $k = 0, 1, ...$ This shows that

$$\sum_{k=0}^{\infty} \rho^{(k)} < \infty. \tag{28}$$

Hence,

$$\begin{aligned}
\|L_d^{(k+1)}\| &= \|L_d^{(k)} G^{(k)}\| = \|L_d^{(0)} G^{(0)} \cdots G^{(k)}\| \\
&\leq \|L_d^{(0)}\|(1 + \|G^{(0)} - I\|) \cdots (1 + \|G^{(k)} - I\|) \\
&\leq \|L_d^{(0)}\| \prod_{i=0}^{k} [\frac{1+\rho^{(i)}}{1-\rho^{(i)}}]^{1/4} \qquad (29)
\end{aligned}$$

for all $k \geq 0$. Now, let

$$B := \|L_d^{(0)}\| \prod_{i=0}^{\infty} [\frac{1+\rho^{(i)}}{1-\rho^{(i)}}]^{1/4}.$$

From (28) we have $B < \infty$ and from (29)

$$\|L_d^{(k+1)}\| \leq B$$

for all $k \geq 0$. This implies that

$$\|L_d^{(k+1)} - L_d^{(k)}\| = \|L_d^{(k)}(G^{(k)} - I)\| \leq B\|G^{(k)} - I\|.$$

Since

$$\sum_{k=0}^{\infty} \|G^{(k)} - I\| < \infty$$

it follows that

$$\{L_d^{(k)} \mid k = 0, 1, ...\}$$

is a convergent sequence. Let

$$L = \lim_{k\to\infty} L_d^{(k)}.$$

Similarly, we can show that

$$\{(L_d^{(k)})^{-1} \mid k = 0, 1, ..\}$$

is a convergent sequence and

$$L^{-1} = \lim_{k\to\infty} (L_d^{(k)})^{-1}.$$

Moreover, by Lemma 4.1 we have

$$W = \lim_{k\to\infty} V^{(k)}.$$

Now consider $(X^{(k)}, Z^{(k)}) \in \mathcal{F}(I)$. We have

$$\lim_{k\to\infty} X^{(k)} = \lim_{k\to\infty} L_d^{(k)} V^{(k)} (L_d^{(k)})^{\mathrm{T}} = LWL^{\mathrm{T}}$$

and similarly,

$$\lim_{k\to\infty} Z^{(k)} = L^{-\mathrm{T}} W L^{-1}.$$

Since $\mathcal{F}(I)$ is closed, it follows that

$$(LWL^{\mathrm{T}}, L^{-\mathrm{T}}WL^{-1}) \in \mathcal{F}(I)$$

and hence

$$(W, W) \in \mathcal{F}(L).$$

$\square$

With an inductive argument, we can now prove that there exists a W-center for any $W \in \mathcal{S}_{++}$.

Theorem 4.1. *For any $W \in \mathcal{S}_{++}$ there exists an invertible matrix $L \in \Re^{n\times n}$ such that*

$$(W, W) \in \mathcal{F}(L)$$

and

$$(LWL^T, L^{-T}WL^{-1}) \in \mathcal{F}(I).$$

Proof:
Due to Assumption 1, we know that there exists a pair $(X, Z) \in \overset{o}{\mathcal{F}}(I)$ and hence

$$(V, V) \in \overset{o}{\mathcal{F}}(L_d),$$

where V and L_d are defined by (7) and (6) respectively. The lemma is proved for the case $W = V$. Now suppose $V \neq W$ and let

$$\theta := \min(\lambda_{\min}(V), \lambda_{\min}(W)).$$

Since $V, W \in \mathcal{S}_{++}$, we have $\theta > 0$. Now consider

$$W(\alpha) := (1-\alpha)V + \alpha W$$

for $\alpha \in [0, 1]$. Remark that for any $\alpha \in [0, 1]$ we have

$$\lambda_{\min}(W(\alpha)) \geq (1-\alpha)\lambda_{\min}(V) + \alpha\lambda_{\min}(W) \geq \theta.$$

Now define $\alpha_0, \alpha_1, \alpha_2, \ldots$ by

$$\alpha_j := \min(1, \frac{j\theta}{3\|W - V\|_F}) \text{ for } j = 0, 1, 2, \ldots.$$

Remark that

$$W = W(1) = W(\alpha_j) \text{ for } j \geq \frac{3}{\theta}\|W - V\|_F.$$

We will prove by induction that for any j,

$$\text{there exists } L^{(j)} \text{ such that } (W(\alpha_j), W(\alpha_j)) \in \mathcal{F}(L^{(j)}). \tag{30}$$

Since $W(\alpha_0) = V$, the statement (30) holds for $j = 0$ with $L^{(0)} = L_d$. We have

$$\|W(\alpha_{j+1}) - W(\alpha_j)\|_F = (\alpha_{j+1} - \alpha_j)\|W - V\|_F \leq \frac{\theta}{3} \leq \frac{\lambda_{\min}(W(\alpha_{j+1}))}{3}.$$

Together with Lemma 4.2, this implies that if the statement (30) holds for $W(\alpha_j)$ then there exists $L^{(j+1)}$ such that

$$(W(\alpha_{j+1}), W(\alpha_{j+1})) \in \mathcal{F}(L^{(j+1)}).$$

Thus the theorem is proved by induction.

□

5. Discussion

In this paper, we studied a possible way of generalizing the concept of weighted centers from linear towards semidefinite programming. A different approach was recently proposed by Monteiro and Pang [12]. In this section, we will compare the properties of the Monteiro-Pang weighted center with the weighted center that we proposed in this paper.

Monteiro and Pang [12] studied the map

$$\frac{1}{2}(XZ + ZX)$$

based on the theory of local homeomorphic maps. One of theire results states that for any positive definite matrix W there exists a *unique* matrix pair (X, Z) such that X is primal feasible and Z is dual feasible and the symmetric part of XZ equals W, i.e.

$$P_{\mathcal{S}}(XZ) = \frac{1}{2}(XZ + ZX) = W$$

However, the converse is not true: given an interior feasible pair $(X, Z) \in \overset{o}{\mathcal{F}}(I)$, the matrix $P_{\mathcal{S}}(XZ)$ is in general *not* positive definite. The following example illustrates this fact. (See also page 464 of Horn and Johnson [3]).

Example 5.1. *Let*

$$X := \begin{bmatrix} 5 & 2 \\ 2 & 1 \end{bmatrix}, Z := \begin{bmatrix} 5 & -2 \\ -2 & 1 \end{bmatrix}$$

then

$$P_{\mathcal{S}}(XZ) = \begin{bmatrix} 21 & 0 \\ 0 & -3 \end{bmatrix}.$$

For the weighted center that we proposed in this paper, we know from Theorem 4.1 that for any positive definite matrix W, there exists a, possibly not unique, pair $(X, Z) \in \overset{o}{\mathcal{F}}(I)$ and an invertible matrix L such that

$$L^{-1}XL^{-\mathrm{T}} = L^{\mathrm{T}}ZL = W. \tag{31}$$

Conversely, we know from (7) that for any pair $(X, Z) \in \overset{o}{\mathcal{F}}(I)$ there exists a matrix L and a positive definite matrix W such that (31) holds.

Another issue is *scale-invariance*, see Todd, Toh and Tütüncü [16]. The weighted center proposed in this paper is scale invariant in the sense that if the pair (X, Z) is a W-weighted center and L is an invertible $n \times n$ matrix, then $(L^{-1}XL^{-\mathrm{T}}, L^{\mathrm{T}}ZL)$ is a weighted center for the transformed problem (5). However, the next example shows that the Monteiro-Pang weighted center is *not* scale invariant.

Example 5.2. *Let*

$$X := \begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix}, \quad Z := \begin{bmatrix} 3 & -1 \\ -1 & 2 \end{bmatrix}, \quad L := \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$$

then

$$P_{\mathcal{S}}(XZ) = \begin{bmatrix} 8 & 0 \\ 0 & 3 \end{bmatrix}$$

but

$$P_{\mathcal{S}}(L^{-1}XZL) = \begin{bmatrix} 8 & 3/4 \\ 3/4 & 3 \end{bmatrix}.$$

As a consequence, we see that the eigenvalues of the symmetric part $P_{\mathcal{S}}(XZ)$ are in general different from the eigenvalues of the matrix product XZ, that defines the weight used in this paper.

References

[1] Alizadeh, F., "Interior point methods in semidefinite programming with applications to combinatorial optimization," *SIAM Journal on Optimization* 5 (1995) 13-51.

[2] Goldfarb, D. and Scheinberg, K., "Interior point trajectories in semidefintie programming," manuscript, Department of Industrial Engineering and Operations Research, Columbia University, New York, 1996.

[3] Horn, R. and Johnson, C., *Matrix Analysis*, Cambridge University Press, Cambridge, 1985.

[4] Jansen, B., Roos, C., Terlaky, T. and Vial, J.-Ph., "Primal-dual target-following algorithms for linear programming," Report 93-107, Faculty of Technical Mathematics and Informatics, Technical University Delft, Delft, The Netherlands, 1993.

[5] Kojima, M., Megiddo, N., Noma, T. and Yoshise, A., *A Unified Approach to Interior Point Algorithms for Linear Complementarity Problems*, Springer-Verlag, Berlin, 1991.

[6] Kojima, M., Mizuno, S. and Yoshise, A., "A primal-dual interior point algorithm for linear programming," in *Progress in Mathematical Programming: Interior point and related methods* pp. 29-37, (ed. Megiddo, N.), Springer Verlag, New York, 1989.

[7] Kojima, M., Shindoh, S. and Hara, S., "Interior-point methods for the monotone linear complementarity problem in symmetric matrices," Research Reports on Information Sciences, B-282, Dept. of Information Sciences, Tokyo Institute of Technology, 2-12-1 Oh-Okayama, Meguro-ku, Tokyo 152, Japan, 1994.

[8] Luo, Z.-Q., Sturm, J.F. and Zhang, S., "Superlinear convergence of a symmetric primal-dual path following algorithm for semidefinite programming," Report 9607/A, Econometric Institute, Erasmus University Rotterdam, Rotterdam, The Netherlands, 1996.

[9] Luo, Z.-Q., Sturm, J.F. and Zhang, S., "Duality and self-duality for conic convex programming," Report 9620/A, Econometric Institute, Erasmus University Rotterdam, Rotterdam, The Netherlands, 1996.

[10] Megiddo, N., "Pathways to the optimal set in linear programming," in *Progress in Mathematical Programming: Interior point and related methods* pp. 131-158, (ed. Megiddo, N.), Springer Verlag, New York, 1989.

[11] Monteiro, R.D.C., "Primal-dual path following algorithms for semidefinite programming," Technical Report, School of Industrial and Systems Engineering, Georgia Tech, Atlanta, Georgia, U.S.A., 1995.

[12] Monteiro, R.D.C. and Pang, J.-S., "On two interior point mappings for nonlinear semidefinite complementarity problems," Technical Report, School of Industrial and Systems Engineering, Georgia Tech, Atlanta, Georgia, U.S.A., 1996.

[13] Nesterov, Y. and Nemirovsky, A., "Interior point polynomial methods in convex programming," Studies in Applied Mathematics 13 (SIAM, Philadelphia, PA, 1994).

[14] Nesterov, Y. and Todd, M.J., "Primal-dual interior-point methods for self-scaled cones," Technical Report 1125, School of Operations Research and Industrial Engineering, Cornell University, Ithaca, New York, 1995.

[15] Sturm, J.F. and Zhang, S., "Symmetric primal-dual path following algorithms for semidefinite programming," Report 9554/A, Econometric Institute, Erasmus University Rotterdam, Rotterdam, The Netherlands, 1995.

[16] Todd, M.J., Toh, K.C. and Tütüncü, R.H., "On the Nesterov-Todd direction in semidefinite programming," Technical Report 1154, School of Operations Research and Industrial Engineering, Cornell University, Ithaca, New York, 1996.

[17] Vandenberghe, L. and Boyd, S., "Semidefinite programming," *SIAM Review* 38 (1996) 1, 49-95.

[18] Zhang, Y., 'On extending primal-dual interior-point algorithms from linear programming to semidefinite programming," Technical Report, Dept. of Mathematics and Statistics, University of Maryland Baltimore County, Baltimore, Maryland, U.S.A., 1995.