The Q Method for Symmetric Cone Programming

Yu Xia* Farid Alizadeh†

March 14, 2005

Abstract

We extend the Q method of semidefinte programming, developed by Alizadeh, Haeberly and Overton, to optimization problems over symmetric cones. An infeasible interior point algorithm and a Newton-type algorithm are given. We give convergence results of the interior point algorithm and prove that the Newton-type algorithm is good for "warm starting".

Key words. Symmetric cone programming, infeasible interior point method, polar decomposition, Jordan algebras, complementarity, Newton's method.

1 Introduction

In this paper we generalize the Q method of [3, 5] for semidefinite programming (SDP) to the class of optimization problems over symmetric cones. Semidefinite programming in standard form can be formulated by a pair of dual optimization problems

$$\begin{array}{llll} \min & C \bullet X & \max & \mathbf{b}^\top \mathbf{y} \\ \text{s.t.} & \mathcal{A}X = \mathbf{b} & \text{s.t.} & \mathcal{A}^\top \mathbf{y} + Z = C \\ & X \succcurlyeq 0 & & Z \succcurlyeq 0 \end{array} \tag{1}$$

Here X, Z and C are in $\mathbb{S}_n$, the set of $n \times n$ symmetric matrices, $\mathcal{A}$ is a linear mapping from $\mathbb{S}_n$ to $\mathbb{R}^m$ and $\mathcal{A}^\top$ its dual mapping from $\mathbb{R}^m$ to $\mathbb{S}_n$, and $X \succcurlyeq 0$ means that X is a positive semidefinite matrix. It is well-known that under certain constraint qualifications the pair of dual problems above have equal optimal values that is at the optimum $C \bullet X = \sum_{ij} C_{ij} X_{ij} = \mathbf{b}^\top \mathbf{y}$. Furthermore, this implies that at the optimum $X \bullet Z = 0$ which together with the fact that $X \succcurlyeq 0$ and $Z \succcurlyeq 0$ implies the stronger $XZ = 0$. Thus, in principle the set of equations

$$\begin{array}{l} \mathcal{A}X = \mathbf{b} \\ \mathcal{A}^\top \mathbf{y} + Z = C \\ XZ = 0 \end{array} \tag{2}$$

forms a square system. The relationship $XZ = 0$ is the *complementary relation* for SDP. For technical reasons which will become clear shortly it is preferable to replace this relation by its equivalent $XZ + ZX = 0$.

To motivate the rationale for the Q method we present a brief review of the primal-dual interior point methods for linear programming and their evolution to semidefinite programs.

*The Institute of Statistical Mathematics, 4-6-7 Minami-Azabu, Minato-ku, Tokyo 106-8569, Japan. yuxia@cas.mcmaster.ca

†Management Science and Information Sstems and Rutgers Center for Operations Research, Rutgers, the State University of New Jersey, U.S.A. alizadeh@rutcor.rutgers.edu research supported in part by the U.S. National Science Foundation.

The square system (2) is quite similar to the one derived from standard form linear programming (LP):

$$
\begin{array}{lll} \min & \mathbf{c}^\top \mathbf{x} & \max \quad \mathbf{b}^\top \mathbf{y} \\ \text{s.t.} & \mathcal{A}\mathbf{x} = \mathbf{b} & \qquad \mathcal{A}^\top \mathbf{y} + \mathbf{z} = \mathbf{c} \\ & \mathbf{x} \geq \mathbf{0} & \qquad \mathbf{z} \geq \mathbf{0} \end{array} \tag{3}
$$

The difference is that the complementary slackness conditions in LP are $x_i z_i = 0$ for $i = 1, \ldots, n$.

Primal-dual interior point methods for linear programming were proposed by Megiddo [16] and polynomial time complexity analysis for them was developed by Kojima, Mizuno and Yoshisi [14] and expressed in a very simple way by Monteiro and Adler [18]. In practice these methods have been reported to have the most favorable numerical properties and many interior point LP solvers are primarily based on them.

The primal-dual interior point methods for LP can be derived by applying the logarithmic barrier function $-\sum_i \ln(x_i)$. After routine manipulation the algorithm boils down to applying Newton's method to a system of equations that contains primal and dual feasibility and a relaxed version of complementarity conditions, that is $x_i z_i = \mu$. The Newton direction from this system can be obtained from the linear system of equations in the form of

$$
\begin{pmatrix} \mathcal{A} & 0 & 0 \\ 0 & \mathcal{A}^\top & I \\ \mathcal{Z} & 0 & \mathcal{X} \end{pmatrix} \begin{pmatrix} \Delta\mathbf{x} \\ \Delta\mathbf{y} \\ \Delta\mathbf{z} \end{pmatrix} = \begin{pmatrix} \mathbf{b} - \mathcal{A}\mathbf{x} \\ \mathbf{c} - \mathcal{A}^\top \mathbf{y} \\ \mu\mathbf{1} - \mathcal{X}\mathcal{Z}\mathbf{1} \end{pmatrix}. \tag{4}
$$

where $\mathcal{X} = \mathrm{Diag}(\mathbf{x})$ and $\mathcal{Z} = \mathrm{Diag}(\mathbf{z})$ are diagonal matrices and $\mathbf{1}$ is the vector of all ones. The algorithm will start with an initial estimate of the optimal solution $(\mathbf{x}, \mathbf{y}, \mathbf{z})$ and in each subsequent iteration replaces it with $(\mathbf{x} + \Delta\mathbf{x}, \mathbf{y} + \Delta\mathbf{y}, \mathbf{z} + \Delta\mathbf{z})$. Solving (4) using Schur complement requires solving a system of equations involving the matrix $\mathcal{A}\mathcal{Z}^{-1}\mathcal{X}\mathcal{A}^\top$. Since $\mathcal{X}$ and $\mathcal{Z}$ are diagonal matrices forming and factoring this matrix is not harder than forming and factoring $\mathcal{A}\mathcal{A}^\top$.

In semidefinite programming the situation is similar to LP in that by using the logarithmic barrier function $-\ln \mathrm{Det}\, X$ and following similar procedure we get the relaxed complementarity condition $\frac{XZ+ZX}{2} = \mu I$. After applying Newton's method we arrive at a system which is quite similar to (4) with the following notable difference:

- For the third set of equations the right hand side is $\mu I - \frac{XZ+ZX}{2}$,
- $\mathcal{X}$ and $\mathcal{Z}$ are not diagonal matrices any more, instead $\mathcal{X} = \frac{I \otimes X + X \otimes I}{2}$ and $\mathcal{Z} = \frac{I \otimes Z + Z \otimes I}{2}$

The resulting Newton direction is called the $XZ + ZX$ method or the AHO method. In this method the Schur complement matrix $\mathcal{A}\mathcal{Z}^{-1}\mathcal{X}\mathcal{A}^\top$ is much harder to form (it requires solving Lyapunov equations of the form $AY + YA = B$). Furthermore if matrices X and Z do not satisfy $XZ + ZX = 2\mu I$ for some μ they do not generally commute and therefore $\mathcal{X}$ and $\mathcal{Z}$ would not commute either. However, this method has an order of convergence of two asymptotically.

Many researchers have proposed transformations of the original SDP problem to an equivalent one that results in $\mathcal{X}$ and $\mathcal{Z}$ matrices that do commute. In this way both computational complexity analysis and numerical calculations will be immensely simplified. Monteiro [17] and Zhang [23] have observed that many of these different transformations are special case of the following: Let P be a positive definite matrix, Replace X with $\tilde{X} \leftarrow PXP$ and Z with $\tilde{Z} \leftarrow P^{-1}ZP^{-1}$. Then, after appropriate transformation of $\mathcal{A}$ and C we derive an equivalent problem. With judicious choice of P we can make $\tilde{X}$ and $\tilde{Z}$ commute. In such cases the Schur complement $\tilde{\mathcal{A}}\tilde{\mathcal{Z}}^{-1}\tilde{X}\tilde{A}^\top$ is somewhat easier to form (it requires solving systems of equation involving $\tilde{Z}$ rather than Lyapunov equations). Also it is somewhat easier to analyze computational complexity of resulting interior point methods. The class of Newton directions obtained from applying those matrices P that result in commuting $\tilde{X}$ an $\tilde{Z}$ is referred to as the *commutative class* of directions.

Applying Monteiro-Zhang type of transformations almost always results in a matrix P that depends on the current estimate of X and Z. Thus, strictly speaking the Newton method is applied to a different system at each iteration and therefore second order order asymptotic rate of convergence may be lost.

The Q method was proposed in [3, 5]. It is an algorithm which remains as close to the Newton method as possible while also retaining commutativity of X and Z estimates at each iteration. To achieve this we first observe that if X and Z commute then they share a common system of eigenvectors; and since they are symmetric these eigenvectors can be chosen to be orthonormal. In other words there is an orthogonal matrix Q such that $X = Q\Lambda Q^\top$ and $Z = Q\Omega Q^\top$. The next step is to replace X and Z as unknowns of the optimization problem with matrix Q and vector of eigenvalues λ and ω. In other words by forcing X and Z to have the same set of eigenvectors, we ensure that they commute. The price we pay in this case is that the linear parts of the system (4) become nonlinear: Instead of $\mathcal{A}X = \mathbf{b}$ we have to deal with $\mathcal{A}Q\mathbf{diag}(\lambda)Q^\top = \mathbf{b}$. Furthermore, since the variable Q is an orthogonal matrix, it is awkward to *add* a Newton correction ΔQ and expect the result to remain orthogonal. Instead, We observe that for a skew symmetric matrix S, $\exp(S)$ is orthogonal. Therefore, ΔQ is represented by $\exp(S)$ and is *multiplied* to the current estimated Q, that is $Q \leftarrow Q\exp(S)$. The next step is to linearize $\exp(S) = I + S + \cdots$ where "$\cdots$" indicates terms involving nonlinear terms in S. This procedure will result in a linear system of equations in $\Delta\lambda, \Delta\omega, \Delta\mathbf{y}$, and S.

Several researcher have observed that the "conventional" interior point algorithms, in particular the primal-dual ones, can be extended to optimization problems over symmetric cones. These cones are self dual and their automorphism groups act transitively on their interiors; this simply means that for any two points in the interior of such a cone, there is a linear transformation that maps one point to the other, and maps the interior of the cone onto itself. It turns out that both the nonnegative orthant and the positive semidefinite matrices are symmetric cones. Also the second order cone (the Lorentz cone) is symmetric. In these cases one needs to replace the ordinary algebra of symmetric matrices with the Euclidean Jordan algebras. The papers by Güler [10], Faybusovich [7, 8] for example, Alizadeh and Schmieta [20, 21], Schmieta [19], Tsuchiya [22] develop these extensions. Our goal in this paper is to extend the Q method to symmetric cones as well.

We present a version of the Q method which for symmetric cones and show that the Jacobian of the method converge to a nonsingular matrix at the optimum in the absence of degeneracy. We will also show the infeasible system converges in finite number of steps.

The rest of the paper is organized as follows. In § 2 we review the basic properties of Jordan algebra and its associated symmetric cone. In § 3, we develop the Q method and show that each iteration is well-defined. We give an infeasible interior point algorithm and its convergence proof in § 4.

Finally some numerical results on the Q method for the Second Cone Programming (SOCP) problem are presented in § 5.

2 A Review of Euclidean Jordan algebras

In this section we outline a minimal foundation of the theory of Euclidean Jordan algebras. This theory serves as our basic toolbox for the analysis of the Q method. Our presentation mostly follows Faraut and Korányi [6], however we also draw from [13] and [21].

2.1 Definitions and Examples

Let $\mathcal{J}$ be an n-dimensional vector space over the field of real numbers with a multiplication "$\circ$" where the map $(\mathbf{x}, \mathbf{y}) \rightarrow \mathbf{x} \circ \mathbf{y}$ is bilinear. Then $(\mathcal{J}, \circ)$ is a Jordan algebra if for all $\mathbf{x}, \mathbf{y} \in \mathcal{J}$

1. $\mathbf{x} \circ \mathbf{y} = \mathbf{y} \circ \mathbf{x}$,

2. $\mathbf{x} \circ (\mathbf{x}^2 \circ \mathbf{y}) = \mathbf{x}^2 \circ (\mathbf{x} \circ \mathbf{y})$ where $\mathbf{x}^2 = \mathbf{x} \circ \mathbf{x}$.

A Jordan algebra $\mathcal{J}$ is called *Euclidean* if there exists a symmetric, positive definite quadratic form Q on $\mathcal{J}$ which is also associative that is

$$Q(\mathbf{x} \circ \mathbf{y}, \mathbf{z}) = Q(\mathbf{x}, \mathbf{y} \circ \mathbf{z}).$$

To make our presentation more concrete, throughout this section we examine the following two examples of Euclidean Jordan algebras.

Example 1 (The Jordan algebras of matrices $\mathbb{M}_n^+$ and $\mathbb{S}_n^+$) The set $\mathbb{M}_n$ of $n \times n$ real matrices with the multiplication $X \circ Y \stackrel{\text{def}}{=} (XY + YX)/2$ forms a Jordan algebra which will be denoted by $\mathbb{M}_n^+$. It is not a Euclidean Jordan algebra, though. The subspace $\mathbb{S}_n$ of real symmetric matrices also forms a Jordan algebra under the "$\circ$" operation; in fact it is Jordan subalgebra of $\mathbb{M}_n^+$. $(\mathbb{S}_n, \circ)$ is Euclidean since if we define $Q(X,Y) = \operatorname{Trace}(X \circ Y) = \operatorname{Trace}(XY)$, then clearly Trace is positive definite, since $\operatorname{Trace}(X \circ X) > 0$ for $X \neq 0$. Its associativity is easy to prove by using the fact that $\operatorname{Trace}(XY) = \operatorname{Trace}(YX)$. We write $\mathbb{S}_n^+$ for this algebra. ■

Example 2 (The quadratic forms algebra $\mathcal{E}_{n+1}^+$) Let $\mathcal{E}_{n+1}$ be the $(n+1)$-dimensional real vector space whose elements $\mathbf{x}$ are indexed from zero. Define the product

$$\mathbf{x} \circ \mathbf{y} \stackrel{\text{def}}{=} \begin{pmatrix} \mathbf{x}^\top \mathbf{y} \\ x_0 y_1 + x_1 y_0 \\ \vdots \\ x_0 y_n + x_n y_0 \end{pmatrix}.$$

Then it is easily verified that $\mathcal{E}_{n+1}^+ = (\mathcal{E}_{n+1}, \circ)$ is a Jordan algebra. Furthermore, $Q(\mathbf{x}, \mathbf{y}) \stackrel{\text{def}}{=} \mathbf{x}^\top \mathbf{y}$ is both associative and positive definite. Thus, $(\mathcal{E}_{n+1}, \circ)$ is Euclidean. ■

Definition 2.1 *If $\mathcal{J}$ is a Euclidean Jordan Algebra then its cone of squares is the set*

$$\mathcal{K}(\mathcal{J}) \stackrel{\text{def}}{=} \{\mathbf{x}^2 : \mathbf{x} \in \mathcal{J}\}.$$

Recall that symmetric cones are closed, pointed, convex cones that are self-dual and their automorphism group acts transitively on their interior. The relevance of the theory of Euclidean Jordan algebras for $\mathcal{K}$-LP optimization problem stems from the following theorem, which can be found in [6].

Theorem 2.1 (Jordan algebraic characterization of symmetric cones) *A cone is symmetric iff it is the cone of squares of some Euclidean Jordan algebra.*

It is well-known and easy to verify that all closed, pointed and convex cones $\mathcal{K}$ induce a partial order "$\succcurlyeq_{\mathcal{K}}$", where $\mathbf{x} \succcurlyeq_{\mathcal{K}} \mathbf{y}$ when $\mathbf{x} - \mathbf{y} \in \mathcal{K}$. We also write $\mathbf{x} \succ_{\mathcal{K}} \mathbf{y}$ when $\mathbf{x} - \mathbf{y} \in \operatorname{Int} \mathcal{K}$. Often we simply write $\succcurlyeq$ and $\succ$ when $\mathcal{K}$ is understood from the context.

Example 3 (Cone of Squares of $\mathbb{S}_n^+$) A symmetric matrix is square of of another symmetric matrix iff it is positive semidefinite. Thus the cone of squares of $\mathbb{S}_n^+$ is the cone of positive semidefinite matrices. We write $X \succcurlyeq 0$ if X is positive semidefinite. We will write $\mathcal{P}_n$ for the cone of positive semidefinite matrices of order n. ■

Example 4 (Cone of squares of $\mathcal{E}_{n+1}^+$) It is straightforward to show that the cone of squares of $\mathcal{E}_{n+1}^+$ is $\mathcal{Q} \stackrel{\text{def}}{=} \{x \in \mathbb{R}^{n+1} : x_0 \geq \|\bar{x}\|\}$ where $\bar{x} = (x_1, \dots, x_n)^\top$, and $\|\cdot\|$ indicates the Euclidean norm, see for example [4, 1]. $\mathcal{Q}$ is called the Lorentz cone, the Second Order Cone and the circular cone. ■

A Jordan algebra $\mathcal{J}$ has an identity, if there exists a (necessarily unique) element $\mathbf{e}$ such that $\mathbf{x} \circ \mathbf{e} = \mathbf{e} \circ \mathbf{x} = \mathbf{x}$ for all $\mathbf{x} \in \mathcal{J}$. Jordan algebras are not necessarily associative, but they are power associative, i.e. the algebra generated by an element $\mathbf{x} \in \mathcal{J}$ is associative.

In the subsequent development we deal exclusively with Euclidean Jordan algebras with identity.

Since "$\circ$" is bilinear for every $\mathbf{x} \in \mathcal{J}$, there exists a matrix $L(\mathbf{x})$ such that for every $\mathbf{y}$, $\mathbf{x} \circ \mathbf{y} = L(\mathbf{x})\mathbf{y}$. In particular, $L(\mathbf{x})\mathbf{e} = \mathbf{x}$ and $L(\mathbf{x})\mathbf{x} = \mathbf{x}^2$.

Example 5 (Identity, and L operators for $\mathbb{S}_n^+$) Clearly the identity element for $\mathbb{S}_n^+$ is the usual identity I for square matrices. Applying the vec operator to a $n \times n$ matrix to turn it into a vector, we get,

$$\operatorname{vec}\big(X \circ Y\big) = \operatorname{vec}\left(\frac{XY + YX}{2}\right) = \frac{1}{2}\Big(I \otimes X + X \otimes I\Big) \operatorname{vec}(Y).$$

Thus, for the $\mathbb{S}_n^+$ algebra, $L(X) = \frac{1}{2}\big(X \otimes I + I \otimes X\big)$, which is also known as Kronocker sum of X and X, see, for example [11] for properties of this operator. ■

Example 6 (Identity, and L operators for $\mathcal{E}_{n+1}^+$) The identity element is the vector $e = (1, 0, \dots, 0)^\top$. From the definition of Jordan multiplication for $\mathcal{E}_{n+1}^+$ it is seen that

$$L(x) = \operatorname{Arw}(\mathbf{x}) \stackrel{\text{def}}{=} \begin{pmatrix} x_0 & \bar{x}^\top \\ \bar{x} & x_0 I \end{pmatrix}$$

This matrix is an arrow-shaped matrix which is related to Lorentz transformations. ■

Since a Jordan algebra $\mathcal{J}$ is power associative we can define rank, minimum and characteristic polynomials, eigenvalues, trace, and determinant for it in the following way.

For each $\mathbf{x} \in \mathcal{J}$ let r be the smallest integer such that the set $\{e, \mathbf{x}, \mathbf{x}^2, \dots, \mathbf{x}^r\}$ is linearly dependent. Then r is the *degree* of $\mathbf{x}$ which we denote as $\deg(\mathbf{x})$. The *algebra rank* of $\mathcal{J}$, $\operatorname{rk}(\mathcal{J})$, is the largest $\deg(\mathbf{x})$ of any member $\mathbf{x} \in \mathcal{J}$. An element $\mathbf{x}$ is called *regular* if its degree equals the algebra rank of the Jordan algebra.

For an element $\mathbf{x}$ of degree d in a rank-r algebra $\mathcal{J}$, since $\{\mathbf{e}, \mathbf{x}, \mathbf{x}^2, \dots, \mathbf{x}^d\}$ is linearly dependent, there are real numbers $a_1(\mathbf{x}), \dots, a_d(\mathbf{x})$ such that

$$\mathbf{x}^d - a_1(\mathbf{x})\mathbf{x}^{d-1} + \cdots + (-1)^d a_d(\mathbf{x})\mathbf{e} = \mathbf{0}. \quad (\mathbf{0} \text{ is the zero vector})$$

The polynomial $\lambda^d - a_1(\mathbf{x})\lambda^{d-1} + \cdots + (-1)^d a_d(\mathbf{x})$ is the *minimum polynomial* of $\mathbf{x}$.

Now, as shown in Faraut and Koranyi, [6], each coefficient $a_i(\mathbf{x})$ of the minimum polynomial is a homogeneous polynomial of degree i, thus in particular it is a continuous function of $\mathbf{x}$. We can now define the notion of characteristic polynomials as follows: If $\mathbf{x}$ is a regular element of the algebra, then we define its characteristic polynomial to be equal to its minimum polynomial. Next, since the set of regular elements are dense in $\mathcal{J}$ (see [6]) we can continuously extend the characteristic polynomials to all elements of $\mathbf{x}$ of $\mathcal{J}$. Therefore, the characteristic polynomial is a degree r polynomial in λ.

Definition 2.2 *The roots, $\lambda_1, \cdots, \lambda_r$ of the characteristic polynomial of* $\mathbf{x}$ *are the* eigenvalues *of* $\mathbf{x}$.

It is possible, in fact certain in the case of nonregular elements, that the characteristic polynomial have multiple roots. However, the minimum polynomial has only simple roots. Indeed, the characteristic and minimum polynomials have the same set of roots, except for their multiplicities. Thus the minimum polynomial of $\mathbf{x}$ always divides its characteristic polynomial.

Definition 2.3 *Let* $\mathbf{x} \in \mathcal{J}$ *and* $\lambda_1, \ldots, \lambda_r$ *be the roots of its characteristic polynomial* $p(\lambda) = \lambda^r - p_1(\mathbf{x})\lambda^{r-1} + \cdots + (-1)^r p_r(\mathbf{x})$*. Then*

1. $\mathrm{tr}(\mathbf{x}) \stackrel{\text{def}}{=} \lambda_1 + \cdots + \lambda_r = p_1(\mathbf{x})$ *is the trace of* $\mathbf{x}$ *in* $\mathcal{J}$*;*

2. $\det(\mathbf{x}) \stackrel{\text{def}}{=} \lambda_1 \cdots \lambda_r = p_r(\mathbf{x})$ *is the determinant of* $\mathbf{x}$ *in* $\mathcal{J}$*.*

Note that trace is a linear function of $\mathbf{x}$.

Example 7 (Characteristic polynomials, eigenvalues, trace and determinant in $\mathbb{S}_n^+$) These notions coincide with the familiar ones in symmetric matrices. Note that $\deg(X)$ is the number of distinct eigenvalues of X and thus is at most n for an $n \times n$ symmetric matrix, in other words the the algebra rank of $\mathbb{S}_n^+$ is n. ■

Example 8 (Characteristic polynomials, eigenvalues, trace and determinant in $\mathcal{E}_{n+1}^+$) Every vector $\mathbf{x} \in \mathcal{E}_{n+1}^+$ satisfies the quadratic equation

$$\mathbf{x}^2 - 2x_0\mathbf{x} + (x_0^2 - \|\overline{\mathbf{x}}\|^2)\mathbf{e} = \mathbf{0} \tag{5}$$

Thus $\mathrm{rank}(\mathcal{E}_{n+1}^+) = 2$ independent of the dimension of its underlying vector space. Furthermore, each element $\mathbf{x}$ has two eigenvalues, $x_0 \pm \|\overline{\mathbf{x}}\|$; $\mathrm{tr}(\mathbf{x}) = 2x_0$ and $\det(\mathbf{x}) = x_0^2 - \|\overline{\mathbf{x}}\|^2$. Except for multiples of identity, every element has degree 2. ■

Together with the eigenvalues comes a decomposition of $\mathbf{x}$ into idempotents, its *spectral decomposition*. Recall that an *idempotent* $\mathbf{c}$ is a nonzero element of $\mathcal{J}$ where $\mathbf{c}^2 = \mathbf{c}$.

1. A *complete system of orthogonal idempotents* is a set $\{\mathbf{c}_1, \ldots, \mathbf{c}_k\}$ of idempotents where $\mathbf{c}_i \circ \mathbf{c}_j = \mathbf{0}$ for all $i \neq j$, and $\mathbf{c}_1 + \cdots + \mathbf{c}_k = \mathbf{e}$.

2. An idempotent is *primitive* if it is not sum of two other idempotents.

3. A complete system of orthogonal primitive idempotents is called a *Jordan frame*.

Note that when the algebra rank is r, then Jordan frames always have r primitive idempotents in them.

Theorem 2.2 (Spectral decomposition, type I) *Let* $\mathcal{J}$ *be a Euclidean Jordan algebra. Then for* $\mathbf{x} \in \mathcal{J}$ *there exist unique real numbers* $\lambda_1, \ldots, \lambda_k$*, all distinct, and a unique complete system of orthogonal idempotents* $\mathbf{c}_1, \ldots, \mathbf{c}_k$ *such that*

$$\mathbf{x} = \lambda_1\mathbf{c}_1 + \cdots + \lambda_k\mathbf{c}_k. \tag{6}$$

See [6].

Theorem 2.3 (Spectral decomposition, type II) *Let* $\mathcal{J}$ *be a Euclidean Jordan algebra with rank* r*. Then for* $\mathbf{x} \in \mathcal{J}$ *there exists a Jordan frame* $\mathbf{c}_1, \ldots, \mathbf{c}_r$ *and real numbers* $\lambda_1, \ldots, \lambda_r$ *such that*

$$\mathbf{x} = \lambda_1\mathbf{c}_1 + \cdots + \lambda_r\mathbf{c}_r \tag{7}$$

and the λ_i *are the eigenvalues of* $\mathbf{x}$*.*

A direct consequence of these facts is that eigenvalues of elements of Euclidean Jordan algebras are always real; this is not the case for arbitrary power associative algebras or even non-Euclidean Jordan algebras.

Example 9 (Spectral decomposition in $\mathbb{S}_n^+$) Every symmetric matrix can be diagonalized by an orthogonal matrix: $X = Q\Lambda Q^\top$. This relation may be written as $X = \lambda_1 \mathbf{q}_1\mathbf{q}_1^\top + \cdots + \lambda_n \mathbf{q}_n\mathbf{q}_n^\top$, where the λ_i are the eigenvalues of X and $\mathbf{q}_i$, columns of Q, their corresponding eigenvectors. Since the $\mathbf{q}_i$ form an orthonormal set, it follows that the set of rank one matrices $\mathbf{q}_i\mathbf{q}_i^\top$ form a Jordan frame: $\left(\mathbf{q}_i\mathbf{q}_i^\top\right)^2 = \mathbf{q}_i\mathbf{q}_i^\top$ and $\left(\mathbf{q}_i\mathbf{q}_i^\top\right)\left(\mathbf{q}_j\mathbf{q}_j^\top\right) = 0$ for $i \neq j$; finally $\sum_i \mathbf{q}_i\mathbf{q}_i^\top = I$. This gives the type II spectral decomposition of X. For type I, let $\lambda_1 > \cdots > \lambda_k$ be distinct eigenvalues of X, where each λ_i has multiplicity m_i. Suppose that $\mathbf{q}_{i_1}, \ldots, \mathbf{q}_{i_{m_i}}$ are a set of orthogonal eigenvectors of λ_i. Define $P_i = \mathbf{q}_{i_1}\mathbf{q}_{i_1}^\top + \cdots + \mathbf{q}_{i_{m_i}}\mathbf{q}_{i_{m_i}}^\top$. Then the P_i form an orthogonal system of idempotents, which add up to I. Also, note that even though for a given eigenvalue λ_i, the corresponding eigenvectors $\mathbf{q}_{i_r}$ may not be unique, the P_i *are* unique for each λ_i. Thus, the identity $X = \lambda_1 P_1 + \cdots + \lambda_k P_k$ is the type I spectral decomposition of X. ■

Example 10 (Spectral decomposition in $\mathcal{E}_{n+1}^+$) Consider the following identity

$$\mathbf{x} = \frac{1}{2}(x_0 + \|\bar{\mathbf{x}}\|)\begin{pmatrix} 1 \\ \frac{\bar{\mathbf{x}}}{\|\bar{\mathbf{x}}\|} \end{pmatrix} + \frac{1}{2}(x_0 - \|\bar{\mathbf{x}}\|)\begin{pmatrix} 1 \\ -\frac{\bar{\mathbf{x}}}{\|\bar{\mathbf{x}}\|} \end{pmatrix} \tag{8}$$

We have already mentioned that $\lambda_{1,2} = x_0 \pm \|\bar{\mathbf{x}}\|$. Let us define $\mathbf{c}_1 = \frac{1}{2}\left(1, \frac{\bar{\mathbf{x}}}{\|\bar{\mathbf{x}}\|}\right)^\top$, and $\mathbf{c}_2 = \frac{1}{2}\left(1, -\frac{\bar{\mathbf{x}}}{\|\bar{\mathbf{x}}\|}\right)^\top$, and observe that $\mathbf{c}_2 = R\mathbf{c}_1$, with R as defined earlier. Also, $\mathbf{c}_i^2 = \mathbf{c}_i$ for $i = 1, 2$, and $\mathbf{c}_1 \circ \mathbf{c}_2 = \mathbf{0}$. Thus, (8) is the type II spectral decomposition of $\mathbf{x}$ which can be alternatively written as $\mathbf{x} = \lambda_1\mathbf{c}_1 + \lambda_2\mathbf{c}_2$. Since only multiples of identity $\alpha\mathbf{e}$ have multiple eigenvalues, their type I spectral decomposition is simply $\alpha\mathbf{e}$, with $\mathbf{e}$ the singleton system of orthonormal idempotents. ■

Now it is possible to extend the definition of any real valued continuous function $f(\cdot)$ to elements of Jordan algebras using their eigenvalues:

$$f(\mathbf{x}) \stackrel{\text{def}}{=} f(\lambda_1)\mathbf{c}_1 + \cdots + f(\lambda_k)\mathbf{c}_k.$$

We are in particular interested in the following functions:

1. The square root: $\mathbf{x}^{1/2} \stackrel{\text{def}}{=} \lambda_1^{1/2}\mathbf{c}_1 + \cdots + \lambda_k^{1/2}\mathbf{c}_k$, whenever all $\lambda_i \geq 0$, and undefined otherwise.

2. The inverse: $\mathbf{x}^{-1} \stackrel{\text{def}}{=} \lambda_1^{-1}\mathbf{c}_1 + \cdots + \lambda_k^{-1}\mathbf{c}_k$ whenever all $\lambda_i \neq 0$ and undefined otherwise.

Note that $\left(\mathbf{x}^{1/2}\right)^2 = \mathbf{x}$, and $\mathbf{x}^{-1} \circ \mathbf{x} = \mathbf{e}$. If $\mathbf{x}^{-1}$ is defined, we call $\mathbf{x}$ *invertible* We call $\mathbf{x} \in \mathcal{J}$ *positive semidefinite* if all its eigenvalues are nonnegative, and *positive definite* if all its eigenvalues are positive. We write $\mathbf{x} \succcurlyeq 0$ (respectively $\mathbf{x} \succ 0$) if $\mathbf{x}$ is positive semidefinite (respectively positive definite.) It is clear that an element is positive semidefinite if, and only if belongs to the cone of squares; it is positive definite if, and only if it belongs to the interior of the cone of squares.

We may also define various norms on $\mathcal{J}$ as functions of eigenvalues much the same way that unitarily invariant norms are defined on square matrices:

$$\|\mathbf{x}\|_{\mathrm{F}} \stackrel{\text{def}}{=} \left(\sum \lambda_i^2\right)^{1/2} = \sqrt{\operatorname{tr}(\mathbf{x}^2)}, \qquad \|\mathbf{x}\|_2 = \max_i |\lambda_i| \tag{9}$$

Observe that $\|e\|_{\mathrm{F}} = \sqrt{r}$. Finally since "$\circ$" is bilinear and trace is a symmetric positive definite quadratic form which is associative, $\operatorname{tr}(\mathbf{x} \circ (\mathbf{y} \circ \mathbf{z})) = \operatorname{tr}((\mathbf{x} \circ \mathbf{y}) \circ \mathbf{z})$, we define the inner product:

$$\langle \mathbf{x}, \mathbf{y} \rangle \stackrel{\text{def}}{=} \operatorname{tr}(\mathbf{x} \circ \mathbf{y})$$

Example 11 (Inverse, square root, and norms in $\mathbb{S}_n^+$) Again, here these notions coincide with the familiar ones. $\|X\|_F = \left(\sum_{ij} X_{ij}^2\right)^{1/2} = \left(\sum_i \lambda_i^2\right)^{1/2}$ is the Frobenius norm of X and $\|X\|_2 = \max_i |\lambda_i(X)|$ is the familiar spectral norm. The inner product $\mathrm{Trace}(X \circ Y) = \mathrm{Trace}(XY)$, is denoted by $X \bullet Y$. ■

Example 12 (Inverse, square root, and norms in $\mathcal{E}_{n+1}^+$) Let $\mathbf{x} = \lambda_1 \mathbf{c}_1 + \lambda_2 \mathbf{c}_2$ with $\{\mathbf{c}_1, \mathbf{c}_2\}$ its Jordan frame. Then,

$$\begin{aligned}
&\mathbf{x}^{-1} = \tfrac{1}{\lambda_1}\mathbf{c}_1 + \tfrac{1}{\lambda_2}\mathbf{c}_2 = \tfrac{R\mathbf{x}}{\det(x)} \qquad \text{when } \det \mathbf{x} \neq 0,\\
&\mathbf{x}^{1/2} = \sqrt{\lambda_1}\mathbf{c}_1 + \sqrt{\lambda_2}\mathbf{c}_2,\\
&\|\mathbf{x}\|_F^2 = \lambda_1^2 + \lambda_2^2 = x_0^2 + \|\overline{\mathbf{x}}\|^2 = \|\mathbf{x}\|^2,\\
&\|\mathbf{x}\|_2 = \max\{\lambda_1, \lambda_2\} = |x_0| + \|\overline{\mathbf{x}}\|,\\
&\langle \mathbf{x}, \mathbf{y} \rangle = \mathrm{tr}(x \circ y) = 2\mathbf{x}^\top \mathbf{y}.
\end{aligned}$$

■

Note that since the inner product $\langle \mathbf{x}, \mathbf{y} \rangle$ is associative, it follows that $L(\mathbf{x})$ is symmetric with respect to $\langle \cdot, \cdot \rangle$, because, $\langle L(\mathbf{x})\mathbf{y}, \mathbf{z} \rangle = \langle \mathbf{x} \circ \mathbf{y}, \mathbf{z} \rangle = \langle \mathbf{y}, \mathbf{x} \circ \mathbf{z} \rangle = \langle \mathbf{y}, L(x)\mathbf{z} \rangle$. From definition of $\mathbf{Q}_{\mathbf{x}}$, it follows that it too is symmetric.

2.2 Peirce decomposition

An important concept in the theory of Jordan algebras is the *Peirce decomposition*. For an idempotent $\mathbf{c}$, since $\mathbf{c}^2 = \mathbf{c}$, one can show that $2L^3(\mathbf{c}) - 3L^2(\mathbf{c}) + L(\mathbf{c}) = \mathbf{0}$, see [6], Proposition III.1.3. Therefore, any idempotent $\mathbf{c}$ has minimum polynomial $2\lambda^3 - 3\lambda^2 + \lambda$, and therefore eigenvalues of $L(\mathbf{c})$ are 0, $\frac{1}{2}$ and 1. Furthermore, the eigenspace corresponding to each eigenvalue of $L(\mathbf{c})$ is the set of $\mathbf{x}$ such that $L(c)\mathbf{x} = i\mathbf{x}$ or equivalently $\mathbf{c} \circ \mathbf{x} = i\mathbf{x}$, for $i = 0, \frac{1}{2}, 1$. Since $L(\mathbf{x})$ is symmetric, these eigenspaces are mutually orthogonal. Therefore, [12].

Theorem 2.4 (Peirce decomposition, type I) *Let $\mathcal{J}$ be a Jordan algebra and $\mathbf{c}$ an idempotent. Then $\mathcal{J}$, as a vector space, can be decomposed as*

$$\mathcal{J} = \mathcal{J}_1(\mathbf{c}) \oplus \mathcal{J}_0(\mathbf{c}) \oplus \mathcal{J}_{\frac{1}{2}}(\mathbf{c}) \tag{10}$$

where

$$\mathcal{J}_i(\mathbf{c}) = \{\mathbf{x} \mid \mathbf{x} \circ \mathbf{c} = i\mathbf{x}\}. \tag{11}$$

The three eigenspaces $\mathcal{J}_i(\mathbf{c})$, are called *Peirce spaces* with respect to $\mathbf{c}$.

Now let $\{\mathbf{c}_1, \ldots, \mathbf{c}_k\}$ be an orthonormal system of idempotents. Each $\mathbf{c}_i$ has it own set of three Peirce spaces $\mathcal{J}_0(\mathbf{c}_i)$, $\mathcal{J}_{\frac{1}{2}}(\mathbf{c}_i)$, and $\mathcal{J}_1(\mathbf{c}_i)$. It can be shown that $L(\mathbf{c}_i)$ all commute and thus share a common system of eigenvectors, [6] Lemma IV.1.3. In fact, the common eigenspaces are of two types ([6] Theorem IV.2.1):

i. $\mathcal{J}_{ii} \stackrel{\text{def}}{=} \mathcal{J}_1(\mathbf{c}_i)$ for all $j \neq i$, and

ii. $\mathcal{J}_{ij} \stackrel{\text{def}}{=} \mathcal{J}_{\frac{1}{2}}(\mathbf{c}_i) \cap \mathcal{J}_{\frac{1}{2}}(\mathbf{c}_j)$.

Thus, with respect to an orthonormal system of idempotents $\{\mathbf{c}_1, \ldots, \mathbf{c}_k\}$, one can give a finer decomposition:

Theorem 2.5 (Peirce decomposition, type II ([6] Theorem IV.2.1)) *Let $\mathcal{J}$ be a Jordan algebra with identity and $\mathbf{c}_i$ a system of orthogonal idempotents such that $\mathbf{e} = \sum_i \mathbf{c}_i$. Then we have the Peirce decomposition $\mathcal{J} = \bigoplus_{i \le j} \mathcal{J}_{ij}$ where*

$$\mathcal{J}_{ii} = \mathcal{J}_1(\mathbf{c}_i) = \{\mathbf{x} \mid \mathbf{x} \circ \mathbf{c}_i = \mathbf{x}\} \tag{12}$$

$$\mathcal{J}_{ij} = \mathcal{J}_{\frac{1}{2}}(\mathbf{c}_i) \cap \mathcal{J}_{\frac{1}{2}}(\mathbf{c}_j) = \left\{\mathbf{x} \mid \mathbf{x} \circ \mathbf{c}_i = \frac{1}{2}\mathbf{x} = \mathbf{x} \circ \mathbf{c}_j\right\}. \tag{13}$$

The Peirce spaces $\mathcal{J}_{ij}$ are orthogonal with respect to any symmetric bilinear form.

Lemma 2.1 (Properties of Peirce spaces) *Let $\mathcal{J}$ be a Jordan algebra with $\mathbf{e} = \sum_i \mathbf{c}_i$ where the $\mathbf{c}_i$ are orthogonal idempotents and let $\mathcal{J}_{ij}$ be the Peirce decomposition relative to the $\mathbf{c}_i$. Then if i, j, k, and l are distinct,*

1. $\mathcal{J}_{ii} \circ \mathcal{J}_{ii} \subseteq \mathcal{J}_{ii}$, $\mathcal{J}_{ii} \circ \mathcal{J}_{ij} \subseteq \mathcal{J}_{ij}$, $\mathcal{J}_{ij} \circ \mathcal{J}_{ij} \subseteq \mathcal{J}_{ii} + \mathcal{J}_{jj}$
2. $\mathcal{J}_{ii} \circ \mathcal{J}_{jj} = \{\mathbf{0}\}$*, if* $i \ne j$
3. $\mathcal{J}_{ij} \circ \mathcal{J}_{jk} \subseteq \mathcal{J}_{ik}$, $\mathcal{J}_{ij} \circ \mathcal{J}_{kk} = \{\mathbf{0}\}$
4. $\mathcal{J}_{ij} \circ \mathcal{J}_{kl} = \{\mathbf{0}\}$ *if* $\{i, j\} \cap \{k, l\} = \emptyset$
5. $\mathcal{J}_0(\mathbf{c}_i) = \bigoplus_{j,k \ne i} \mathcal{J}_{jk}$.

In a Euclidean Jordan algebra, the Peirce decomposition is closely related to the orthogonal decomposition of the vector space with respect to $L(\mathbf{x})$. If $\mathbf{x} = \sum_{i=1}^k \lambda_i \mathbf{c}_i$ then the Peirce spaces $\mathcal{J}_{ij}$ corresponding to the system of orthogonal idempotents $\{\mathbf{c}_1, \ldots \mathbf{c}_k\}$ are eigenspaces of $L(\mathbf{x})$. An immediate consequence of this observation is the following:

Lemma 2.2 *Let $\mathbf{x} \in \mathcal{J}$ with spectral decomposition type I: $\mathbf{x} = \lambda_1 \mathbf{c}_1 + \cdots + \lambda_k \mathbf{c}_k$. Then the eigenvalues of $L(\mathbf{x})$ have the form*

$$\frac{\lambda_i + \lambda_j}{2} \qquad 1 \le i \le j \le k,$$

in particular, all λ_i are eigenvalues of $L(\mathbf{x})$, and $\mathbf{x}$ is positive (semi-definite) definite iff $L(\mathbf{x})$ is positive (semi-definite) definite, moreover, for $\mathbf{x} \succcurlyeq \mathbf{0}$, $L(\mathbf{x})$ is invertible if $\mathbf{x}$ is positive definite.

Another useful representation is as follows:

Proposition 2.1 *Let $\mathbf{x} \in \mathcal{J}$ and $\{\mathbf{f}_1 \ldots, \mathbf{f}_r\}$ some Jordan frame. Then with respect to this frame*

$$\mathbf{x} = \sum_{i=1}^r \mathbf{x}_i + \sum_{i<j} \mathbf{x}_{ij} \tag{14}$$

where $\mathbf{x}_i = x_i \mathbf{c}_i$ and $\mathbf{x}_{ij} \in \mathcal{J}_{ij}$

Thus, as Faraut and Korány state, Peirce decomposition can be interpreted as writing a vector$\mathbf{x}$ of the Euclidean Jordan algebra as an $r \times r$ matrix where each entry is a vector in $\mathcal{J}$. The diagonal entries are multiples of corresponding primitive idempotents. We will also refer to this representation as the Peirce decomposition.

Example 13 (Peirce decomposition in $\mathbb{S}_n^+$) Let $E = \{\mathbf{q}_1\mathbf{q}^\top, \ldots, \mathbf{q}_n\mathbf{q}_n^\top\}$ be a Jordan frame, where the $\mathbf{q}_i$ are orthonormal set of vectors in $\mathbb{R}^n$. To see the Peirce spaces associated with the idempotent $C = \mathbf{q}_1\mathbf{q}_1^\top + \cdots + \mathbf{q}_m\mathbf{q}_m^\top$, first observe that C has eigenvalues 1 (with multiplicity m) and 0 (with multiplicity $n-m$). Next recall that $L(C) = \frac{1}{2}\big(C \otimes I + I \otimes C\big)$. In general, if A and B are square matrices with eigenvalues λ_i and ω_j, respectively, and with corresponding eigenvectors $\mathbf{u}_i$ and $\mathbf{v}_j$ then $A \otimes I + I \otimes B$ has eigenvalues $\lambda_i + \omega_j$, and corresponding eigenvectors $\mathbf{u}_i \otimes \mathbf{v}_j = \mathrm{vec}(\mathbf{v}_j\mathbf{u}_i^\top)$, for $i, j = 1, \ldots, n$; Thus, the eigenvalues of $L(C)$ are $1, \frac{1}{2}, 0$. Therefore, the Peirce space $\mathcal{J}_i(C)$ consists of those matrices $A \in \mathbb{S}_n$ where $iA = A \circ (\sum_{i=1}^m \mathbf{q}_i\mathbf{q}_i^\top) = \frac{1}{2}\sum_{i=1}^m (A\mathbf{q}_i\mathbf{q}_i^\top + \mathbf{q}_i\mathbf{q}_i^\top A)$. It follows that,

$$
\begin{aligned}
\mathcal{J}_1(C) &= \big\{A \in \mathbb{S}_n \mid \mathbf{q}_i^\top A\mathbf{q}_j = 0 \quad \text{if } m+1 \le i \le n \quad \text{or} \quad m+1 \le j \le n\big\},\\
\mathcal{J}_0(C) &= \big\{A \in \mathbb{S}_n \mid \mathbf{q}_i^\top A\mathbf{q}_j = 0 \quad \text{if } 1 \le i \le m \quad \text{or} \quad 1 \le i \le m\big\},\\
\mathcal{J}_{\frac{1}{2}}(C) &= \big\{A \in \mathbb{S}_n \mid \mathbf{q}_i^\top A\mathbf{q}_j = 0 \quad \text{if } 1 \le i, j \le m \quad \text{or} \quad m+1 \le i, j \le n\big\}.
\end{aligned}
$$

Also it is clear that $\dim(\mathcal{J}_0(C)) = \frac{m(m+1)}{2}$, $\dim(\mathcal{J}_{\frac{1}{2}}(C)) = nm$, and $\dim(\mathcal{J}_1(C)) = \frac{(n-m)(n-m+1)}{2}$.
■

Example 14 (Peirce decomposition in $\mathcal{E}_{n+1}^+$) Let $\mathbf{x} = \lambda_1\mathbf{c}_1 + \lambda_2\mathbf{c}_2$ be the spectral decomposition of $\mathbf{x}$ with $\lambda_1 \neq \lambda_2$. First, it can be verified by inspection that the matrix $\mathrm{Arw}(\mathbf{x})$ has eigenvalues $\lambda_{1,2} = x_0 \pm \|\overline{\mathbf{x}}\|$, each with multiplicity 1 and corresponding eigenvectors $\mathbf{c}_{1,2}$; and $\lambda_3 = x_0$ with multiplicity $n-1$. An idempotent $\mathbf{c}$ which is not equal to identity element $\mathbf{e}$ is of the form $\mathbf{c} = \frac{1}{2}(1, \mathbf{q})$ where $\mathbf{q}$ is a unit length vector. Thus $\mathrm{Arw}(\mathbf{c})$ has one eigenvalue equal to 1, another equal to 0 and the remaining $n-1$ eigenvalues equal to $\frac{1}{2}$. It is easy to verify that

$$
\begin{aligned}
\mathcal{J}_1(\mathbf{c}) &= \{\alpha\mathbf{c} \mid \alpha \in \mathbb{R}\},\\
\mathcal{J}_0(\mathbf{c}) &= \{\alpha R\mathbf{c} \mid \alpha \in \mathbb{R}\}, \quad R = \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & -I \end{pmatrix},\\
\mathcal{J}_{\frac{1}{2}}(\mathbf{c}) &= \{(0, \mathbf{p})^\top \mid \mathbf{p}^\top\mathbf{q} = 0\}.
\end{aligned}
$$

■

We need the notion of a simple algebra in the subsequent discussion.

Definition 2.4 *An (Euclidean Jordan) algebra $\mathcal{A}$ is called* simple *iff is not the direct sum of two (Euclidean Jordan) subalgebras.*

Proposition 2.2 ([6]) *If $\mathcal{J}$ is a Euclidean Jordan algebra, then it is, in a unique way, a direct sum of simple Euclidean Jordan algebras.*

Lemma 2.3 ([6]) *Let $\mathcal{J}$ be an n-dimensional simple Euclidean Jordan algebra and $(\mathbf{a}, \mathbf{b})$, $(\mathbf{a}_1, \mathbf{b}_1)$ two pairs of orthogonal primitive idempotents. Then*

$$
\dim\Big(\mathcal{J}_{\frac{1}{2}}(\mathbf{a}) \cap \mathcal{J}_{\frac{1}{2}}(\mathbf{b})\Big) = \dim\Big(\mathcal{J}_{\frac{1}{2}}(\mathbf{a}_1) \cap \mathcal{J}_{\frac{1}{2}}(\mathbf{b}_1)\Big) = d, \tag{15}
$$

and, if $\mathrm{rk}(\mathcal{J}) = r$,

$$
n = r + \frac{d}{2}r(r-1).
$$

Example 15 (The parameter d for $\mathbb{S}_n^+$ and $\mathcal{E}_{n+1}^+$) For $\mathbb{S}_n^+$, $d = 1$. For $\mathcal{E}_{n+1}^+$, $d = n-1$. ■

Proposition 2.3 *Two elements* $\mathbf{x}$ *and* $\mathbf{y}$ *of a simple Euclidean Jordan algebra have the same spectrum if, and only if* $L(\mathbf{x})$ *and* $L(\mathbf{y})$ *have the same spectrum.*

We now state a central theorem which specifies the classification of Euclidean Jordan algebras:

Theorem 2.6 ([6] Chapter V.) *Let* $\mathcal{J}$ *be a simple Euclidean Jordan algebra. Then* $\mathcal{J}$ *is isomorphic to one of the following algebras*

1. *the algebra* $\mathcal{E}_{n+1}^+$,
2. *the algebra* $\mathbb{S}_n^+$ *of* $n \times n$ *symmetric matrices,*
3. *the algebra* $(\mathbb{H}_n, \circ)$ *of* $n\times n$ *complex Hermitian matrices under the operation* $X \circ Y = \frac{1}{2}\big(XY + YX\big)$,
4. *the algebra* $(\mathbb{Q}_n, \circ)$ *of* $n \times n$ *quaternion Hermitian matrices under the operation,* $X \circ Y = \frac{1}{2}\big(XY + YX\big)$,
5. *the exceptional Albert algebra, that is the algebra* $(\mathbb{O}_3, \circ)$ *of* 3×3 *octonion Hermitian matrices under the operation,* $X \circ Y = \frac{1}{2}\big(XY + YX\big)$.

Since octonion multiplication is not associative, the 27-dimensional Albert algebra is not induced by an associative operation, as the other four are. That is why it is called *exceptional.*

2.3 Operator commutativity

We say two elements $\mathbf{x}, \mathbf{y}$ of a Jordan algebra $\mathcal{J}$ *operator commute* if $L(\mathbf{x})L(\mathbf{y}) = L(\mathbf{y})L(\mathbf{x})$. In other words, $\mathbf{x}$ and $\mathbf{y}$ operator commute if for all $\mathbf{z}$, $\mathbf{x} \circ (\mathbf{y} \circ \mathbf{z}) = \mathbf{y} \circ (\mathbf{x} \circ \mathbf{z})$. If $A \subseteq \mathcal{J}$ we denote the set of elements in $\mathcal{J}$ that operator commute with all $\mathbf{a} \in A$ by $C_{\mathcal{J}}(A)$.

Lemma 2.4 ([12]) *If* $\mathbf{c}$ *is an idempotent in* $\mathcal{J}$ *then* $C_{\mathcal{J}}(\{\mathbf{c}\}) = \mathcal{J}_0(\mathbf{c}) \oplus \mathcal{J}_1(\mathbf{c})$. *So* $C_{\mathcal{J}}(\mathbf{c})$ *is a subalgebra of* $\mathcal{J}$.

Theorem 2.7 (Operator commutativity [12]) *Let* $\mathcal{J}$ *be an arbitrary finite-dimensional Jordan algebra,* B *a subalgebra. Then* $C_{\mathcal{J}}(B)$ *is a subalgebra and*

$$C_{\mathcal{J}}(B) = \bigcap_{i=1}^{k} C_{\mathcal{J}}(\{\mathbf{c}_i\}) = \bigcap_{i=1}^{k} \left(\mathcal{J}_0(\mathbf{c}_i) \oplus \mathcal{J}_1(\mathbf{c}_i)\right), \tag{16}$$

where $\{\mathbf{c}_1, \ldots, \mathbf{c}_k\}$ *are idempotents that form a basis of* B.

Lemma 2.5 *If* $\{\mathbf{c}_1, \ldots, \mathbf{c}_k\}$ *is a complete system of orthogonal idempotents in* $\mathcal{J}$ *and* B *is the subalgebra generated by them, then*

$$C_{\mathcal{J}}(B) = \bigoplus_{i=1}^{k} \mathcal{J}_1(\mathbf{c}_i). \tag{17}$$

Lemma 2.6 *Let* $\mathcal{J}$ *be a Euclidean Jordan algebra and* $\mathbf{x} \in \mathcal{J}$. *Let also that* $\mathbf{x}$ *have a type I spectral decomposition with idempotents* $\mathbf{c}_1, \ldots, \mathbf{c}_k$. *Denote by* B *the subalgebra of* $\mathcal{J}$ *generated by the* $\mathbf{c}_i$. *Then* $C_{\mathcal{J}}(\{\mathbf{x}\}) = C_{\mathcal{J}}(B)$.

Theorem 2.8 *Let* $\mathbf{x}$ *and* $\mathbf{y}$ *be two elements of a Euclidean Jordan algebra* $\mathcal{J}$*. Then* $\mathbf{x}$ *and* $\mathbf{y}$ *operator commute if, and only if there is a Jordan frame* $\mathbf{c}_1, \ldots, \mathbf{c}_r$ *such that* $\mathbf{x} = \sum_{i=1}^r \lambda_i \mathbf{c}_i$ *and* $\mathbf{y} = \sum_{i=1}^r \mu_i \mathbf{c}_i$*.*

Example 16 (Operator commutativity in $\mathbb{S}_n^+$) The operator commutativity is a generalization of the notion of commutativity of the associative matrix product. If $X, Y \in \mathbb{S}_n$ commute, they share a common system of orthonormal eigenvectors, which means they have a common Jordan frame. From the eigenstructure of Kronocker sums it is clear that X and Y commute iff $X \otimes I + I \otimes X$ and $Y \otimes I + I \otimes Y$ commute. ■

Example 17 (Operator commutativity in $\mathcal{E}_{n+1}^+$) Here from the eigenstructure of $\operatorname{Arw}(\cdot)$ described in Example 14 it is easily verified that if $\operatorname{Arw}(\mathbf{x})$ and $\operatorname{Arw}(\mathbf{y})$ commute then there is a Jordan frame $\{\mathbf{c}_1, \mathbf{c}_2\}$ such that $\mathbf{x} = \lambda_1 \mathbf{c}_1 + \lambda_2 \mathbf{c}_2$, and $\mathbf{y} = \omega_1 \mathbf{c}_1 + \omega_2 \mathbf{c}_2$. ■

Proposition 2.4 *[6, p. 65, proposition IV.1.4] Let* $\mathbf{a}$ *and* $\mathbf{b}$ *be two orthogonal primitive idempotents. If* $\mathbf{x} \in \mathcal{J}_{\frac{1}{2}}(\mathbf{a}) \cap \mathcal{J}_{\frac{1}{2}}(\mathbf{b})$*, then*

$$\mathbf{x}^2 = \frac{1}{2}\|\mathbf{x}\|^2(\mathbf{a} + \mathbf{b}).$$

We now prove a statemet which was stated as an exercise in [6, p. 79, exercise IV.7], and will be needed later.

Proposition 2.5 *Let*

$$\mathbf{x} = \sum_{i=1}^r x_i \mathbf{f}_i + \sum_{i<j} \mathbf{x}_{ij}$$

be as in the Peirce decomposition 14 of $\mathbf{x} = \lambda_1 \mathbf{f}_1 + \cdots + \lambda_r \mathbf{f}_r$ *with* $\mathbf{f}_i$ *its Jordan frame. Then*

$$x_i \geq 0, \quad \|\mathbf{x}_{ij}\|^2 \leq 2 x_i x_j,$$

with both inequalities strict when $\mathbf{x} \in \operatorname{Int} \mathcal{K}$*.*

Proof: For any $\mathbf{u}$, λ and μ satisfying

$$\mathbf{u} \in \mathcal{J}_{\frac{1}{2}}(\mathbf{f}_i) \cap \mathcal{J}_{\frac{1}{2}}(\mathbf{f}_j), \quad \|\mathbf{u}\|^2 = 2, \quad \lambda^2 + \mu^2 = 1,$$

set

$$\mathbf{w} = \lambda^2 \mathbf{f}_i + \mu^2 \mathbf{f}_j + \lambda\mu \mathbf{u}.$$

By Proposition 2.4,

$$\mathbf{u}^2 = \frac{1}{2}\|\mathbf{u}\|^2 (\mathbf{f}_i + \mathbf{f}_j) = \mathbf{f}_i + \mathbf{f}_j.$$

Then

$$\begin{aligned} \mathbf{w}^2 &= \lambda^4 \mathbf{f}_i + \mu^4 \mathbf{f}_j + 2\lambda^3\mu \mathbf{f}_i \mathbf{u} + 2\lambda\mu^3 \mathbf{f}_j \mathbf{u} + \lambda^2\mu^2 \mathbf{u}^2 \\ &= \lambda^4 \mathbf{f}_i + \mu^4 \mathbf{f}_j + \lambda^3\mu \mathbf{u} + \lambda\mu^3 \mathbf{u} + \lambda^2\mu^2 (\mathbf{f}_i + \mathbf{f}_j) \\ &= \lambda^2 \mathbf{f}_i + \mu^2 \mathbf{f}_j + \lambda\mu \mathbf{u}. \end{aligned}$$

Therefore, $\mathbf{w}$ is an idempotent. Since $L(\mathbf{x})$ is positive semidefinite,

$$\langle \mathbf{x}, \mathbf{w} \rangle = \langle \mathbf{w}, L(\mathbf{x})\mathbf{w} \rangle \geq 0.$$

The Peirce decomposition is orthogonal with respect to the inner product. Therefore,

$$\langle \mathbf{x}, \mathbf{w} \rangle = \lambda^2 x_i + \mu^2 x_j + \lambda\mu \langle \mathbf{u}, \mathbf{x}_{ij} \rangle.$$

Thus we have

$$\lambda^2 x_i + \mu^2 x_j + \lambda\mu \langle \mathbf{u}, \mathbf{x}_{ij} \rangle \geq 0. \tag{18}$$

Setting $\lambda = 1$ and $\mu = 0$ in (18) we get

$$x_i \geq 0, \quad (i = 1, \dots, r).$$

If $\mathbf{x}_{ij} = \mathbf{0}$ then $\|\mathbf{x}_{ij}\|^2 \leq 2x_i x_j$ is alreaddy satisfied. So in the following we assume $\mathbf{x}_{ij} \neq \mathbf{0}$, and in (18), set

$$\mathbf{u} = \frac{\sqrt{2}}{\|\mathbf{x}_{ij}\|} \mathbf{x}_{ij}. \tag{19}$$

Then

$$\langle \mathbf{u}, \mathbf{x}_{ij} \rangle = \sqrt{2} \|\mathbf{x}_{ij}\|.$$

If neither x_i nor x_j is zero, $\|\mathbf{x}_{ij}\|^2 \leq 2x_i x_j$ is proved by setting

$$\lambda = \sqrt{\frac{x_j}{x_i + x_j}}, \quad \mu = -\sqrt{\frac{x_i}{x_i + x_j}}$$

in (18).

Next we show that if either x_i or x_j is 0, x_{ij} must also be 0. Assume $x_i = 0$ and $x_{ij} \neq 0$. let $\{(\lambda_k, \mu_k)\}$ be a sequence such that $\forall\, k > 0$:

$$\lambda_k < 0, \quad \mu_k > 0, \quad \lambda_k^2 + \mu_k^2 = 1, \quad \lambda_k \to -1, \quad \mu_k \to 0.$$

Then (18), with $\mathbf{u}$ chosen as in (19), reduces to:

$$\mu_k x_j + \sqrt{2} \lambda_k \|x_{ij}\| \geq 0.$$

Hence $\|\mathbf{x}_{ij}\| = 0$ by (18) with λ and μ being the above sequence.
Analogously, when $x_j = 0$, $x_{ij} = \mathbf{0}$ is proved by exchanging λ_k and μ_k in the above sequence.

When $\mathbf{x} \in \operatorname{Int} \mathcal{K}$, $L(\mathbf{x})$ is positive definite. Therefore, λ and μ can not both be zero, and $\mathcal{J}_{ii} + \mathcal{J}_{jj} + \mathcal{J}_{ij}$ is a direct sum. So $\mathbf{w} \neq \mathbf{0}$. Hence $\langle \mathbf{x}, \mathbf{w} \rangle > 0$. and in (18) and all inequalities following it $\geq$ can be replaced by $>$. ■

2.4 Automorphisms of symmetric cones

Let $\mathcal{K}$ be a symmetric cone which is the cone of squares of a simple Euclidean Jordan Algebra $\mathcal{J}$, with identity element $\mathbf{e}$. The automorphism group of $\mathcal{K}$, $G(\mathcal{K})$, plays a central role in design of interior point methods for both the Q method and conventional methods. We will not distinguish between the linear transformations in $G(\mathcal{K})$ and the nonsingular matrices representing these transformations.

It can be shown that the group $G(\mathcal{K})$ is a continuous–indeed a Lie–group and that a subgroup that fixes any point $\mathbf{a} \in \mathcal{K}$ in $G(\mathcal{K})$, denoted by $G_{\mathbf{a}}$, is a compact Lie group. Furthermore, from the fact that $\mathcal{K}$ is self-dual, it follows that if $A \in G(\mathcal{K})$ then $A^\top \in G(\mathcal{K})$.

The full automorphism group of a symmetric cone is too large for our purposes. In the Q method we need automorphisms that map Jordan frames to other Jordan frames. Therefore we focus our attention on some essential subgroups of $G(\mathcal{K})$. Specifically, $G(\mathcal{K})$ may consist of several connected components; the component that contains the identity matrix is a subgroup which we represent by $G_{\mathcal{K}}$. It can be shown that $G_{\mathcal{K}}$ already acts transitively on the interior of $\mathcal{K}$. Furthermore, the group of dilations of $\mathcal{K}$, that is$\{\alpha I \mid \alpha > 0\}$ (which is isomorphic to $\mathbb{R}_+$, the set of positive real numbers, under multiplication) is a subgroup of automorphism group of $\mathcal{K}$. Since, dilations don't preserve Jordan frames, we wish to exclude them from consideration. To do so we define a subgroup $K_{\mathcal{K}} = G_{\mathcal{K}} \cap O_n$ where O_n is the orthogonal group of order n (represented by the set of all $n \times n$ orthogonal matrices). It can be shown [6] that $K_{\mathcal{K}}$ is the subgroup of $G(\mathcal{K})$ that fixes the identity element $\mathbf{e}$. The group $K_{\mathcal{K}}$ and particular subgroups of it are the focus of our attention.

Example 18 (Automorphism group of positive semidefinite real matrices) For the cone of positive semidefinite matrices $\mathcal{P}_n$, observe that X is positive definite if and only if $PXP^\top$ is positive definite for all nonsingular matrices P. Thus, $GL_n \otimes GL_n$ which is isomorphic to GL_n, the group of nonsingular $n \times n$ matrices, is the automorphism group of $\mathcal{P}$. Recall that O_n is the group of orthogonal matrices; and SO_n, the group of orthogonal matrices with determinant 1, is the connected component of O_n that contains the identity matrix. Let $Q \in SO_n$. The way Q acts on a positive semidefinite matrix X is by the operation $X \to QXQ^\top$. In other words in the space of symmetric matrices, $K_{\mathcal{P}} = \{Q \otimes Q \mid Q \in SO_n\}$, which is isomorphic to SO_n. Thus $K_{\mathcal{P}} = SO_n$. ■

Example 19 (Automorphism group of $\mathcal{Q}$) For the Lorentz cone $\mathcal{Q} \subseteq \mathbb{R}^{n+1}$, the automorphism group is $SO_{1,n} \oplus \mathbb{R}_+$, and $K = SO_n$. The group $SO_{1,n}$ consists of matrices that preserves the bilinear form $x_0y_0 - x_1y_1 - \cdots - x_nv_n$; it includes hyperbolic rotations in any two dimensional plane spanned by a Jordan frame $\{\mathbf{c}_1, \mathbf{c}_2\}$. The set of transformation fixing the identity element $\mathbf{e} = (1; \mathbf{0})$, that is $K_{\mathcal{Q}}$, is all orthogonal transformations of $\mathbb{R}^n$ with determinant 1. ■

2.5 Polar decomposition

The following theorem is essential for the development of the Q method:

Theorem 2.9 *Let $\mathcal{J}$ be a simple Euclidean Jordan algebra and $\mathcal{K}$ its cone of squares. The group $K_{\mathcal{K}}$ acts transitively on the set of Jordan frames of $\mathcal{J}$. Thus given any fixed Jordan frame $\{\mathbf{d}_1, \ldots, \mathbf{d}_r\}$, and any element $\mathbf{x} \in \mathcal{J}$ with spectral decomposition $\mathbf{x} = \lambda_1\mathbf{c}_1 + \cdots + \lambda_r\mathbf{c}_r$ there is an orthogonal matrix Q in $K_{\mathcal{K}}$ such that $Q\mathbf{c}_i = \mathbf{d}_i$ for $i = 1, \ldots, r$. Therefore, the vector $\mathbf{x}$ can be turned by this element of $K_{\mathcal{K}}$ to a vector $\mathbf{a} = \lambda_1\mathbf{d}_1 + \ldots + \lambda_r\mathbf{d}_r$ such that $\mathbf{x} = Q\mathbf{a}$.*

The decomposition $\mathbf{x} = Q\mathbf{a}$ is referred to as the *polar decomposition* and is a generalization of the diagonalizability of symmetric matrices by orthogonal matrices. A key component of this concept is that we need to fix a particular Jordan frame $\mathbf{d}_i$, which we call *the standard frame*. Then the polar decomposition is with respect to this frame. Notice that the standrd frame is completely arbitrary, but fixed. In practice it may be convenient to choose a frame that makes subsequent computations as efficient as possible.

Example 20 (Polar decomposition and diagonalization in $\mathbb{S}_n$) Consider the Jordan frame $E_i = \mathbf{e}_i\mathbf{e}_i^\top$ where $\mathbf{e}_i$ is the vector with all zero entries, except the i^{th} entry which is one. Then the theorem above states that for any symmetric matrix $X = \lambda_1\mathbf{q}_1\mathbf{q}_1^\top + \cdots + \lambda_n\mathbf{q}_n\mathbf{q}_n^\top$ there is an orthogonal Q with determinant 1 such that $QXQ^\top = \lambda_1E_1 + \cdots + \lambda_nE_n$ which is a diagonal matrix of eigenvalues of X. Thus diagonalization of a symmetric matrix X is essentially finding an element of $K_{\mathcal{P}}$ (that is an $n \times n$ orthogonal matrix with determinant 1) such that the set of eigenvectors of X is mapped to the standard basis $\mathbf{e}_i$. The standard frame in this context is simply the E_i. ■

In the case of $\mathbb{S}_n$ if a matrix X has distinct eigenvalues then there is a unique, up to reordering of columns, orthogonal matrix Q that diagonalized X. In fact, if $F_{\mathbf{p}} = \{\mathbf{p}_1\mathbf{p}_1^\top, \ldots, \mathbf{p}_n\mathbf{p}_n^\top\}$ and $F_{\mathbf{q}} = \{\mathbf{q}_1\mathbf{q}_1^\top, \ldots, \mathbf{q}_n\mathbf{q}_n^\top\}$ are Jordan frames, then the we can construct orthogonal matrices $P = (\mathbf{p}_1, \ldots, \pm\mathbf{p}_n)$ and $Q = (\mathbf{q}_1, \ldots, \pm\mathbf{q}_n)$ (with sign of ± 1 in the last columns chosen so that $\operatorname{Det} P =$

$\operatorname{Det} Q = 1$). Then the system of equations in U: $U\mathbf{p}_i = \mathbf{q}_i$ for $i = 1, \ldots, n$ is equivalent to $UP = Q$ which yields $U = QP^{-1}$ and U is orthogonal and has determinant 1 so it is in $K_{\mathcal{P}}$.

In other Euclidean Jordan algebras there may be many orthogonal transformations in $K_{\mathcal{K}}$ that map a given Jordan frame to another. For instance for $\mathcal{E}_{n+1}$ and the associated symmetric cone $\mathcal{Q}$ let $\{\mathbf{d}_1 = \frac{1}{2}(1; \mathbf{p}), \mathbf{d}_2 = \frac{1}{2}(1; -\mathbf{p})\}$ be the standard frame and $\{\mathbf{c}_1 = \frac{1}{2}(1; \mathbf{q}), \mathbf{c}_2 = \frac{1}{2}(1; -\mathbf{q})\}$ any other Jordan frame. Then for dimensions four and more, there are many matrices $Q \in K_{\mathcal{Q}}$ where $Q\mathbf{q}_1 = \mathbf{p}_1$ and $Q\mathbf{q}_2 = \mathbf{p}_2$.

In such cases, we can further narrow our focus to a subgroup of K determined by the standard frame $\mathbf{d}_i$. Specifically let $M_{\mathcal{K}}$ be the subgroup of $K_{\mathcal{K}}$ that fixes every $\mathbf{d}_i$ in the standard frame, and thus every point in the linear space $R_{\mathbf{d}} = \{a_1\mathbf{d}_1 + \cdots + a_r\mathbf{d}_r \mid a_i \in \mathbb{R}^r\}$. Then the quotient group $L = K/M$ composed of the left cosets of M in K is also a subgroup of K. The polar decomposition now can be further refined as follows:

Proposition 2.6 (Polar Decomposition in simple Euclidean Jordan algebra) *Given Jordan frame $\{\mathbf{c}_1, \ldots, \mathbf{c}_r\}$ and the standard frame $\{\mathbf{d}_1, \ldots, \mathbf{d}_r\}$, there is a unique orthogonal matrix Q in $L_{\mathcal{K}} = K_{\mathcal{K}}/M_{\mathcal{K}}$ such that $\mathbf{d}_i = Q\mathbf{c}_i$ for $i = 1, \ldots, r$.*

Corollary 2.1 *If $\mathbf{x}$ is regular (that is has r distinct eigenvalues) then there is a unique (up to reordering of columns) orthogonal matrix Q in $L_{\mathcal{K}}$ that maps $\mathbf{x}$ to $R_{\mathbf{d}} = \{a_1\mathbf{d}_1 + \cdots + a_r\mathbf{d}_r\}$.*

Example 21 (Polar decomposition in $\mathcal{E}_{n+1}$.) As was mentioned earlier in general for $n \geq 4$ there are many orthogonal matrices Q in $K_{\mathcal{Q}} = SO_n$ that map the Jordan frame $\{\mathbf{c}_1, \mathbf{c}_2\}$ to the standard frame $\{\mathbf{d}_1, \mathbf{d}_2\}$. Since $K_{\mathcal{Q}} = SO_n$, the group $M_{\mathcal{Q}}$ fixing $\mathbf{d}_1$ and $\mathbf{d}_2$ is SO_{n-2}. Therefore, $L_{\mathcal{Q}} = SO_n/SO_{n-2} = SO_2$. The unique element of SO_2 that maps $\mathbf{c}_1$ to $\mathbf{d}_1$ and $\mathbf{c}_2$ to $\mathbf{d}_2$ is the rotation map in the plane spanned by $\mathbf{c}_1$ and $\mathbf{d}_1$. ∎

2.6 The Exponential Map

We shall see later in our development of the Q method that we need to search $L_{\mathcal{K}}$ for improvement of the current iterate's Jordan frame. However it is awkward to do this in $L_{\mathcal{K}}$ which is a smooth manifold. It is more convenient if the search is carried out in a linear space. The mechanism by which $L_{\mathcal{K}}$ is searched indirectly through a linear space involves the *exponential map*

First we note that for skew symmetric matries $S = -S^\top$, the matrix exponential $\exp(S) = \sum_k \frac{S^k}{k!}$ is an orthogonal matrix. Thus, $\exp(\cdot)$ maps the linear space of skew symmetric matrices onto the special orthogonal group SO_n.

For $\mathbb{S}_n^+$ the groups $K_{\mathcal{P}}$ and $L_{\mathcal{P}}$ are both equal to SO_n. However for other kinds of Euclidean Jordan algebra (for instance $\mathcal{E}_{n+1}^+$) we need to find a linear subspace $\mathfrak{l}_{\mathcal{K}}$ of skew symmetric matrices which, under the exponential function, is mapped onto $L_{\mathcal{K}}$. Given a standard frame $\{\mathbf{d}_1, \ldots, \mathbf{d}_r\}$ the space $\mathfrak{l}_{\mathcal{K}}$ is constructed as follows ([6] Chapter VI):

1. For each pair of primitive idempotents $\mathbf{d}_i, \mathbf{d}_j$, $i > j$, in the standard frame define

$$\mathfrak{l}_{ij} = \big\{[L(\mathbf{d}_i), L(\mathbf{p})] \mid \mathbf{p} \in \mathcal{J}_{ij}\big\}$$

   where $[L(\mathbf{d}_i), L(\mathbf{p})] = L(\mathbf{d}_i)L(\mathbf{p}) - L(\mathbf{p})L(\mathbf{d}_i)$ is the Jacobi *bracket* operation. Also recall that $\mathcal{J}_{ij} = \mathcal{J}_{\frac{1}{2}}(\mathbf{d}_i) \cap \mathcal{J}_{\frac{1}{2}}(\mathbf{d}_j)$ are subspaces involved in the Peirce decomposition.

2. Set

$$\mathfrak{l}_{\mathcal{K}} = \sum_{i<j} \mathfrak{l}_{ij}$$

It can be shown that this sum is in fact a direct sum as the spaces $\mathfrak{l}_{ij}$ are mutually orthogonal.

Proposition 2.7 *The map* $\exp : \mathfrak{l}_{\mathcal{K}} \to L_{\mathcal{K}}$ *is onto.*

Example 22 ($\mathfrak{l}$ with respect to $\mathbb{S}_n$ and the standard basis) As before let $E_i = \mathbf{e}_i\mathbf{e}_i^\top$, and consider the Jordan frame $\{E_1, \ldots, E_n\}$. Then the space $\mathcal{J}_{\frac{1}{2}}(E_i)$ consists of matrices of the form

$$\mathbf{u}\mathbf{e}_i^\top - \mathbf{e}_i\mathbf{u}^\top = \begin{pmatrix} & & & u_1 & & & \\ & & & \vdots & & & \\ & & & u_{i-1} & & & \\ u_1 & \ldots & u_{i-1} & 0 & u_{i+1} & \ldots & u_m \\ & & & u_{i+1} & & & \\ & & & \vdots & & & \\ & & & u_m & & & \end{pmatrix}.$$

Therefore,

$$\mathcal{J}_{ij} = \{a(E_{ij} + E_{ji}) \mid a \in \mathbb{R}\}.$$

Thus for any $U_{ij} \in \mathcal{J}_{ij}$, using standard properties of Kronocker product, we have

$$[L(E_{ii}), L(U_{ij})] = \frac{1}{4} I \otimes (u_{ij}E_{ij} - u_{ij}E_{ji}) + \frac{1}{4}(u_{ij}E_{ij} - u_{ij}E_{ji}) \otimes I.$$

Hence

$$\mathfrak{l} = \left\{ \left(I \otimes S - S^T \otimes I\right) : S^T = -S \right\},$$

which is as expected since $\mathfrak{l}$ is isomorphic to $n \times n$ skew symmetric matrices. ■

Example 23 ($\mathfrak{l}$ with respect to $\mathcal{E}_{n+1}$) Unlike the symmetric matrices there is no agreed upon standard frame for $\mathcal{E}_{n+1}^+$. Therefore, in analogy to symmetric matrices, we elect to set our standard frame to one that is as sparse as possible:

$$\mathbf{d}_1 = \frac{1}{2}\begin{pmatrix}1\\1\\\mathbf{0}\end{pmatrix} \quad \mathbf{d}_2 = \mathbf{e} - \mathbf{d}_1 = \frac{1}{2}\begin{pmatrix}1\\-1\\\mathbf{0}\end{pmatrix}$$

Therefore,

$$\mathcal{J}_{12} = \{(0; 0; \mathbf{s}) \mid \mathbf{s} \in \mathbb{R}^{n-1}\}.$$

After routine calculation we find that $\mathfrak{l}$ consists of matrices of the form

$$S_\mathbf{s} = \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & s_2 & \cdots & s_n \\ 0 & -s_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & -s_n & 0 & \cdots & 0 \end{pmatrix}, \quad \text{and} \quad S_\mathbf{e} = \begin{pmatrix} 1 & 0 & 0 & \mathbf{0}^\top \\ 0 & -1 & 0 & \mathbf{0}^\top \\ 0 & 0 & -1 & \mathbf{0}^\top \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & I \end{pmatrix}$$

In this case it is straightforward to calculate $\exp(S_\mathbf{s})$. First note that

$$S_\mathbf{s}^2 = \begin{pmatrix} 0 & 0 & \mathbf{0}^\top \\ 0 & -\mathbf{s}^\top\mathbf{s} & \mathbf{0}^\top \\ \mathbf{0} & \mathbf{0} & -\mathbf{s}\mathbf{s}^\top \end{pmatrix}$$

Now,

$$S_\mathbf{s}^{2k+2} = (\mathbf{s}^\top\mathbf{s})^k S_\mathbf{s}^2 \quad \text{and} \quad S_\mathbf{s}^{2k+1} = -(\mathbf{s}^\top\mathbf{s})^k S_\mathbf{s}$$

Thus, when $\|\mathbf{s}\| \neq 0$,

$$\begin{aligned}\exp(S_{\mathbf{s}}) &= I + \frac{S_{\mathbf{s}}^2}{\|\mathbf{s}\|^2}\Big(\sum_{i=1}^{\infty}(-1)^{i+1}\frac{\|\mathbf{s}\|^2 i}{(2i)!}\Big) + \frac{S_{\mathbf{s}}}{\|\mathbf{s}\|^2}\Big(\sum_{i=0}^{\infty}(-1)^{i}\frac{\|\mathbf{s}\|^{2i+1}}{(2i+1)!}\Big)\\ &= I + \frac{1-\cos(\|\mathbf{s}\|)}{\|\mathbf{s}\|^2}S_{\mathbf{s}}^2 + \frac{\sin(\|\mathbf{s}\|)}{\|\mathbf{s}\|}S_{\mathbf{s}}.\end{aligned}$$

■

2.7 The case of non-simple Eculidean Jordan algebra

The preceding section was primalrly focused on the simple Jordan algebra and its cone of squares. In optimization, the symmetric cones arise as direct sum of many simple cones, and thus the underlying Jordan algebra $\mathcal{J}$ is a direct sum of simple algebras $\mathcal{J}_i$. Let $\mathcal{K} = \mathcal{K}_1 \oplus \cdots \oplus \mathcal{K}_n$, each $\mathcal{K}_k \subseteq \mathbb{R}^{n_k}$. Then the automorphim groups discussed in the previous section also extend to the direct sum algebra through direct sums. Thus, $K_{\mathcal{K}} = \bigoplus_k K_{\mathcal{K}_k}$, $L_{\mathcal{K}} = \bigoplus_k L_{\mathcal{K}_k}$. For each $\mathcal{J}_k$ let $F_k = \{\mathbf{d}_1^{(k)}, \ldots, \mathbf{d}_{r_i}^{(k)}\}$ be the standard frame for $\mathcal{J}_k$. Then the $F = \bigcup_k F_k$ will be the standard frame for $\mathcal{J}$, and $\mathfrak{l}$ is defined as $\bigoplus_k \mathfrak{l}_k$. Once more the map $\exp : \mathfrak{l} \to L_{\mathcal{K}}$ is onto.

3 Symmetric Cone Programming

3.1 Barrier function and the central path

Optimization over symmetric cone in standard form is defined by the pair of primal and dual programs:

$$(20)\qquad \begin{array}{llll} & \underline{\textbf{Primal}} & & \underline{\textbf{Dual}} \\ \min & \mathbf{c}_1^\top \mathbf{x}_1 + \cdots + \mathbf{c}_n^\top \mathbf{x}_n & \max & \mathbf{b}^\top \mathbf{y} \\ \text{s.t.} & A_1\mathbf{x}_1 + \cdots + \mathcal{A}_n \mathbf{x}_n = \mathbf{b} & \text{s.t.} & \mathcal{A}_1^\top \mathbf{y} + \mathbf{z}_k = \mathbf{c}_k \quad k = 1, \ldots, n \\ & \mathbf{x}_k \in \mathcal{K}_k & & \mathbf{z}_k \in \mathcal{K}_k \end{array}$$

Here each $\mathcal{K}_k$ is a symmetric cone that is the cone of squares of a *simple* Jordan algebra $\mathcal{J}_k$. We also define $\mathbf{x} = (\mathbf{x}_1; \ldots; \mathbf{x}_n)$, $\mathbf{z} = (\mathbf{z}_1; \ldots; \mathbf{z}_n)$, $\mathbf{c} = (\mathbf{c}_1; \ldots; \mathbf{c}_n)$, $A = (A_1, \ldots, A_n)$, $\mathcal{K} = \bigoplus_k \mathcal{K}_k$, and $\mathcal{J} = \bigoplus_k \mathcal{J}_k$.

First we make a comment about notation. Since all vectors are column vectors, we use $(\mathbf{x}; \mathbf{y})$ to concatenate vectors columnwise, and use $(\mathbf{x}^\top, \mathbf{y}^\top)$ to concatenate them rowwise. We use the same notation for matrices as well. We assume that both the primal and the dual problems are feasible and in fact **Primal** has a feasible point $\mathbf{x} \in \operatorname{Int}\mathcal{K}$ and **Dual** has feasible point where $\mathbf{z} \in \operatorname{Int}\mathcal{K}$. Then the optimal values of **Primal** and **Dual** coincide. Furthermore, we assume that A has full row rank.

We now state a theorem that is the basis of complementary slackness theorem in symmetric cone programming.

Lemma 3.1 *Suppose $\mathbf{x}$ and $\mathbf{z}$ belong to $\mathcal{K}$. Then the following statements are equivalent.*

1. $\langle \mathbf{x}, \mathbf{z} \rangle = 0$.
2. $\mathbf{x} \circ \mathbf{z} = \mathbf{0}$.
3. *There is a Jordan frame $\mathbf{f}_1, \ldots, \mathbf{f}_r$ where*

$$\mathbf{x} = \sum_{i=1}^{r} \lambda_i \mathbf{f}_i, \quad \mathbf{z} = \sum_{i=1}^{r} \omega_i \mathbf{f}_i, \quad \lambda_i \geq 0, \omega_i \geq 0 \ (i = 1, \ldots, n).$$

Furthermore,

$$\lambda_i \omega_i = 0 \quad (i = 1, \dots, r).$$

Proof: (3) $\Rightarrow$ (2) $\Rightarrow$ (1) is obvious. The relation (1) $\Rightarrow$ (2) is shown in [8] and [4]. To show that $\mathbf{x} \circ \mathbf{z} = \mathbf{0}$ implies that $\mathbf{x}$ and $\mathbf{z}$ share a Jordan frame, assume that $\mathbf{x} = \lambda_1 \mathbf{c}_1 + \cdots + \lambda_l \mathbf{c}_l$, with $\lambda_1 \geq \cdots \geq \lambda_l > 0$ and $\{\mathbf{c}_1, \dots, \mathbf{c}_l\}$ a subset of the Jordan frame of $\mathbf{x}$ giving rise to this spectral decomposition. It is clear from $\mathbf{x} \circ \mathbf{c}_i = \lambda_i \mathbf{c}_i$ that $L(\mathbf{x})\mathbf{c}_i = \lambda_i \mathbf{c}_i$, and thus the $\mathbf{c}_i$'s form a partial set of orthogonal eigenvectors of $L(\mathbf{x})$. Since $L(\mathbf{x})\mathbf{z} = 0$, $\mathbf{z}$ is in the null space of $L(\mathbf{x})$, in particular, $\langle \mathbf{c}_i, \mathbf{z} \rangle = 0$, and since both $\mathbf{c}_i$ and $\mathbf{z}$ are $\in \mathcal{K}$, this implies that $\mathbf{c}_i \circ \mathbf{z} = \mathbf{0}$. Also let $\mathbf{z} = \omega_1 \mathbf{d}_1 + \cdots + \omega_m \mathbf{d}_m$ with $\omega_1 \geq \cdots \geq \omega_m > 0$, and the $\mathbf{d}_i$ form a subset of the Jordan frame of $\mathbf{z}$. By symmetry we must have $\mathbf{x} \circ \mathbf{d}_i = \mathbf{0}$. Indeed, $\mathbf{c}_i \circ \mathbf{d}_j = \mathbf{0}$, and thus $\{\mathbf{c}_1, \dots, \mathbf{c}_l, \mathbf{d}_1, \dots, \mathbf{d}_m\}$ is a set of mutually orthogonal primitive idempotents that, if $l + m < r$, can be completed to a Jordan frame. This Jordan frame gives the spectral decomposition of both $\mathbf{x}$ and $\mathbf{z}$ proving that they operator commute. ■

It has been shown elsewhere (see for instance, [21] and references therein) that the function $-\ln \det \mathbf{x}$ is a barrier function for $\mathcal{K}$. Following standard practice, if we drop $\mathbf{x} \in \mathcal{K}$ and add the term $-\ln \det \mathbf{x}$ to the objective function in **Primal**, and apply the Karush-Kuhn-Tucker conditions on the resulting relaxed problem we get

$$\begin{aligned} A\mathbf{x} &= \mathbf{b} \\ A^\top \mathbf{y} + \mathbf{z} &= \mathbf{c} \\ \mathbf{x} \circ \mathbf{z} &= \mu \mathbf{e} \end{aligned} \tag{21}$$

where $\mathbf{e}$ is the identity element of $\mathcal{J}$. The set of points $(\mathbf{x}; \mathbf{y}; \mathbf{z})$ that satsify (21) is the *primal-dual central path* or simply the central path.

It is clear that for $\mu > 0$, if $(\mathbf{x}; \mathbf{y}; \mathbf{z})$ is on the central path then $\mathbf{x}$ and $\mathbf{z}$ operator commute, since $\mathbf{x}^{-1} = \mu \mathbf{z}$ and by definition of the inverse they share a common Jordan frame. We have also seen that for $\mu = 0$, that is at the optimal point, $\mathbf{x}$ and $\mathbf{z}$ operator commute.

3.2 Outline of the Q method

In order to describe the Q method, we have to rewrite (21) by replacing $\mathbf{x}$ and $\mathbf{z}$ with their spectral decomposition. Since, both $\mathbf{x}$ and $\mathbf{z}$ share a Jordan frame on the central path, we elect to enforce this condition *everywhere*, that is we restrict our search only to those pairs of $\mathbf{x}$ and $\mathbf{z}$ that operator commute. To this end let's write $\mathbf{x} = \lambda_1 \mathbf{c}_1 + \cdots + \lambda_r \mathbf{c}_r$ and $\mathbf{z} = \omega_1 \mathbf{c}_1 + \cdots + \omega_r \mathbf{c}_r$. Then the system (21) can be written as

$$\begin{aligned} &A\big(\lambda_1 \mathbf{c}_1 + \cdots + \lambda_r \mathbf{c}_r\big) = \mathbf{b} \\ &A^\top \mathbf{y} + \omega_1 \mathbf{c}_1 + \cdots + \omega_r \mathbf{c}_r = \mathbf{c} \\ &\lambda_i \omega_i = \mu \quad \text{for } i = 1, \dots, r \end{aligned} \tag{22}$$

The outcome of this transformation is that the first two sets of equalities, the primal and dual feasibility conditions, are now nonlinear. On the other hand the third set of equalities, the relaxed complementarity conditions, are now simpler and resemble the analogous conditions in linear programming. Our new set of variables are now λ_i, ω_i and $\mathbf{c}_i$.

As it stands the system (22) is not convenient for Newton method or any other conventional iterative process. The reason is that we need to ensure that $\mathbf{c}_i$ in each iteration form a legitimate Jordan frame. To ensure this, we set a fixed standard Jordan frame $\mathbf{d}_i$. For instance for the $\mathbb{S}_n^+$, the standard frame could be $\mathbf{e}_i \mathbf{e}_i^\top$, and for $\mathcal{E}_{n+1}^+$, $\mathbf{d}_i$ could be as described in Example (23). Then we can write $\mathbf{c}_i = Q\mathbf{d}_i$, where $Q \in L_{\mathcal{K}}$, which by Corollary 2.1 is unique up to reordering of its columns if λ_i and ω_i are distinct. Thus, now, Q replaces $\mathbf{c}_i$ as a variable and the search space for Q is the group $L_{\mathcal{K}}$ which is a smooth manifold. However, $L_{\mathcal{K}}$ is not suitable as a search space for Newton method either. Instead we use the exponential map and write $Q = \exp(S)$ where $S \in \mathfrak{l}_{\mathcal{K}}$. Thus the system of equations, after decomposing $\mathbf{x}$ and $\mathbf{z}$ into $\mathbf{x}_i$ and $\mathbf{z}_i$ where each block belongs to a simple Jordan algebra, is further refined into further refined into:

$$
\begin{aligned}
&\sum_{i=1}^{n} A_i\Big(\exp(S_i)\big(\sum_{j=1}^{n_i}(\lambda_i)_j(\mathbf{d}_i)_j\big)\Big) \\
&A_i^{\top}\mathbf{y} + \exp(S_i)\sum_{j=1}^{n_i}(\omega_i)_j(\mathbf{d}_i)_j = \mathbf{c}_i \quad \text{for } i = 1,\dots,n \\
&(\lambda_i)_j(\omega_i)_j = \mu \quad \text{for } i = 1,\dots,n,\ j = 1,\dots,r_i
\end{aligned}
\tag{23}
$$

The unknowns in this system are $(\lambda_i)_j$, $(\omega_i)_j$ and S_i each of which belong to particular linear spaces. Thus, Newton's method, or other conventional methods can be applied to this system and, in fact techniques such as line search can be carried out without difficulty.

3.3 The Newton System and its properties

Applying Newton's method to (23) can be carried out as follows. Assuming that current estimate of the solution for **Primal** and **Dual** problems are $(\lambda_i^k)_j, (\omega_i^k)_j$, and Q_i, we replace

$$
\begin{aligned}
(\lambda_i^k)_j &\leftarrow (\lambda_i^k)_j + (\Delta\lambda_i)_j \\
(\omega_i^k)_j &\leftarrow (\omega_i^k)_j + (\Delta\omega_i)_j \\
Q_i^k &\leftarrow Q_i \exp(S_i)
\end{aligned}
$$

Replacing these new values in (23), writing $\exp(S) = I + S + \cdots$, and dropping all nonlinear terms in Δ's and in S yields

$$B^k S\sum_{i=1}^{n}\sum_{j=1}^{r_i}(\lambda_i^k)_j(\mathbf{d}_i)_j + B^k\sum_{i=1}^{n}\sum_{j=1}^{r_i}\Delta(\lambda_i)_j(\mathbf{d}_i)_j = \mathbf{r}_p^k, \tag{24a}$$

$$(B^k)^{\top}\Delta\mathbf{y} + S\sum_{i=1}^{n}\sum_{j=1}^{r_i}(\omega_i^k)_j(\mathbf{d}_i)_j + \sum_{i=1}^{n}\sum_{j=1}^{r_i}\Delta(\omega_i)_j(\mathbf{d}_i)_j = \mathbf{r}_d^k, \tag{24b}$$

$$\Lambda^k\Delta\omega + \Omega^k\Delta\lambda = \mathbf{r}_c^k, \tag{24c}$$

where

$$
\begin{aligned}
&B^k = AQ^k \\
&\mathbf{r}_p^k \stackrel{\text{def}}{=} \mathbf{b} - A\mathbf{x}^k, \quad \mathbf{r}_d^k \stackrel{\text{def}}{=} (Q^k)^{\top}\left(\mathbf{c} - \mathbf{z}^k - A^{\top}\mathbf{y}^k\right), \quad (r_{ci}^k)_j \stackrel{\text{def}}{=} \mu^k - (\lambda_i^k)_j(\omega_i^k)_j \\
&\Lambda^k = \operatorname{Diag}\Big(\big(\lambda_i^k\big)_j\Big) \quad \text{and} \quad \Omega^k = \operatorname{Diag}\Big(\big(\omega_i^k\big)_j\Big)
\end{aligned}
$$

The following lemma shows that these iterates are well defined.

Lemma 3.2 *The linear System in* (24) *has a unique solution if the iterate* $\mathbf{x}^k$ *and* $\mathbf{z}^k$ *are regular.*

Proof: Since $S_i \in \mathfrak{l}_i$, it can be represented as

$$S_i = \sum_{j<l}\Big[L\left((\mathbf{d}_i)_j\right), L(\xi_{i(jl)}^k)\Big], \quad \xi_{i(jl)}^k \in (\mathcal{J}_i)_{jl}.$$

Then

$$S_i \sum_{j=1}^{r_i} (\omega_i^k)_j (\mathbf{d}_i)_j = \frac{1}{4} \sum_{j<l} \Big[(\omega_i^k)_l - (\omega_i^k)_j \Big] \xi_{i(jl)}^k, \quad S_i \sum_{j=1}^{r_i} (\lambda_i^k)_j (\mathbf{d}_i)_j = \frac{1}{4} \sum_{j<l} \Big[(\lambda_i^k)_l - (\lambda_i^k)_j \Big] \xi_{i(jl)}^k.$$

From (24c) we have

$$\Delta\lambda = \left(\Omega^k\right)^{-1} \left(\mathbf{r}_c^k - \Lambda^k \Delta\omega\right). \tag{25}$$

which we substitute for $\Delta(\lambda_i)_j$ in (24a). Note that since $\mathbf{x}^k$ and $\mathbf{z}^k$ are regular, we have

$$(\lambda_i^k)_l - (\lambda_i^k)_j \neq 0, \quad (\omega_i^k)_l - (\omega_i^k)_j \neq 0 \quad (\text{for } l \neq j).$$

Also $\mathbf{x}^k, \mathbf{z}^k \in \operatorname{Int} \mathcal{K}$ implies $(\lambda_i^k)_j / (\omega_i^k)_j > 0$. The Peirce decomposition is a direct sum, therefore there is a one-to-one linear transformation P that maps

$$S \sum_{i=1}^{n} \sum_{j=1}^{r_i} (\omega_i^k)_j (\mathbf{d}_i)_j + \sum_{i=1}^{n} \sum_{j=1}^{r_i} \Delta(\omega_i)_j (\mathbf{d}_i)_j$$

to

$$S \sum_{i=1}^{n} \sum_{j=1}^{r_i} (\lambda_i^k)_j (\mathbf{d}_i)_j - \sum_{i=1}^{n} \sum_{j=1}^{r_i} \frac{(\lambda_i^k)_j}{(\omega_i^k)_j} \Delta(\omega_i)_j (\mathbf{d}_i)_j. \tag{26}$$

Next, we apply $B^k P$ to both sides of (24b) and subtract (26) from it. Since A is surjective, $B^k P (B^k)^\top$ is a bijection. Therefore, $\Delta\mathbf{y}$ is uniquely determined by system (24). Then, by the fact that Peirce decomposition is a direct sum, $\Delta\omega$ and S can be obtained from (24b). Finally, $\Delta\lambda$ is obtained from (24c). ■

Remark 3.1 *We can always ensure that λ^{k+1} and ω^{k+1} are regular by careful choice of step sizes. For example, assume ω^k is regular. Only when $\Delta(\omega_i)_1 \neq \Delta(\omega_i)_2$ and $\beta = \frac{(\omega_i^k)_2 - (\omega_i^k)_1}{\Delta(\omega_i)_1 - \Delta(\omega_i)_2}$, is it possible that $(\omega_i^k)_1 + \beta\Delta(\omega_i)_1 = (\omega_i^k)_2 + \beta\Delta(\omega_i)_2$. In this case, we can use a smaller step size β_i'. It is obvious that β' can be at least as large as $\frac{\beta}{2}$. And β_i' are not necessarily the same for all i.*

3.4 Nonsingularity of the Newton system at the optimum

Let $\dim \mathcal{J}_i = n_i$, $r = \sum_{i=1}^n r_i$ be the rank of the Jordan frame underlying the problem, and $N = \sum_{i=1}^n n_i$. Also let F^k represent the linear transformation from $\mathfrak{l} \times \mathbb{R}^m \times \mathbb{R}^r \times \mathbb{R}^r$ into $\mathbb{R}^m \times \mathbb{R}^N \times \mathbb{R}^r$ defined by the left hand side of (24) at k^{th} iteration. Similarly, let F be the same linear transformation evaluated at the optimum. From the properties of the Peirce decomposition it is easy to see that the dimension of $\mathfrak{l}_i$ is $n_i - r_i$. Therefore, F^k maps a linear space into a space of the same dimension. In this section, we present conditions under which F is one-to-one.

3.4.1 Definitions

Faybusovich in [8] has proved that the Jacobian of the Newton system arising from the method described in [2] is nonsingular for symmetric cone programming. Here we present analogous result for the Q method. To prove that F is one-to-one, we first give some definitions and cite related results from Faybusovich here.

Let $(\mathbf{x}, \mathbf{z})$ be the primal-dual solutions of the symmetric cone program, λ_i and ω_i denote their eigenvalues respectively.

Definition 3.1 *[8, definition 3.4] The pair* $(\mathbf{x}, \mathbf{z})$ *is* strictly complementary *if* $\mathbf{x} + \mathbf{z} \in \operatorname{Int} \mathcal{K}$.

Through rearrangement if necessary, we can assume for $i = 1, \ldots, n$:

$$(\lambda_i)_j \neq 0 \quad (j = 1, \ldots, t_i),$$
$$(\lambda_i)_j = 0 \quad (j = t_i + 1, \ldots, r_i).$$

Since $\langle \mathbf{x}, \mathbf{z} \rangle = 0$, by Lemma 3.1, $\mathbf{x}$ and $\mathbf{z}$ are strictly complementary if and only if for $i = 1, \ldots, n$:

$$(\omega_i)_j = 0 \quad (j = 1, \ldots, t_i),$$
$$(\omega_i)_j \neq 0 \quad (j = t_i + 1, \ldots, r_i).$$

For $i = 1, \ldots, n$, define

$$e(\mathbf{x}_i) \stackrel{\text{def}}{=} \sum_{j=1}^{t_i} (\mathbf{d}_i)_j.$$

and similarly define $e(\mathbf{z}_i)$. Note that $e(\mathbf{x}_i)$ is an idempotent in $\mathcal{J}_i$. Then the Peirce decomposition of $\mathcal{J}_i$ with respect to $e(\mathbf{x}_i)$ is given by

$$\mathcal{J}_i = (\mathcal{J}_i)_1 \Big(e\big(\mathbf{x}_i\big) \Big) \bigoplus (\mathcal{J}_i)_{\frac{1}{2}} \Big(e\big(\mathbf{x}_i\big) \Big) \bigoplus (\mathcal{J}_i)_0 \Big(e\big(\mathbf{x}_i\big) \Big).$$

Following [8], we partition indices into three blocks:

$$\Gamma_0 \stackrel{\text{def}}{=} \{(j, l) \colon t_i + 1 \leq j \leq l \leq r_i\}$$
$$\Gamma_{\frac{1}{2}} \stackrel{\text{def}}{=} \{(j, l) \colon 1 \leq j \leq t_i,\ t_i + 1 \leq l \leq r_i\}$$
$$\Gamma_1 \stackrel{\text{def}}{=} \{(j, l) \colon 1 \leq j \leq l \leq t_i\}.$$

Then for $k = 0, \frac{1}{2}, 1$,

$$\mathcal{J}_i \left(e(\mathbf{x}_i), k \right) = \bigoplus_{j,l \in \Gamma_k} \mathcal{J}_{i(jl)} \left(e(\mathbf{x}_i) \right).$$

Definition 3.2 *The solution* $\mathbf{x}$ *is primal nondegenerate if*

$$\left(\sum_{i=1}^{n} \mathcal{J}_i \left(e(\mathbf{x}_i), 0 \right) \right) \cap \operatorname{Range}(B^\top) = \{\mathbf{0}\}$$

Definition 3.3 *The solution* $\mathbf{z}$ *is dual nondegenerate if*

$$\left(\sum_{i=1}^{n} \mathcal{J}_i \left(e(\mathbf{z}_i), 0 \right) \right) \cap null(B) = \{\mathbf{0}\},$$

If **Primal** is nondegenerate then **Dual** must have unique solution. Conversely, if **Dual** is nondegenerate then **Primal** must have a unique solution. Both these conditions are necessary and sufficient if strict complementarity is satisfied, [8].

3.4.2 F is One-to-One

We assume that the primal and dual nondegeneracy and strict complementarity are satisfied. Furthermore we assume all $\mathbf{x}_i$ and $\mathbf{z}_i$ are regular. Then the linear system of equation derived from applying Newton's method, and represented by $F(\lambda_i, \omega_i, S) = 0$, is one-to-one at the solution.

Lemma 3.3 *Let*

$$(\mathbf{x}, \mathbf{y}, \mathbf{z}) = \left(Q \sum_{i=1}^{n} \sum_{j=1}^{r_i} (\lambda_i)_j (\mathbf{d}_i)_j, \mathbf{y}, Q \sum_{i=1}^{n} \sum_{j=1}^{r_i} (\omega_i)_j (\mathbf{d}_i)_j \right)$$

be an optimal solution of the symmetric cone program satisfying strict complementarity, primal-dual nondegeneracy and for $(\mathbf{x}, \mathbf{z})$ *regularity. Then* F *is one-to-one at* $(\lambda, \omega, Q, \mathbf{y})$.

Proof: We need to show that kernel of (24) at $(\lambda, \omega, Q, \mathbf{y})$ is $\{0\}$. For $i = 1, \ldots, n$, assume

$$\begin{aligned} &(\lambda_i)_j \neq 0 \quad (\omega_i)_j = 0 \quad (j = 1, \ldots, t_i), \\ &(\lambda_i)_j = 0 \quad (\omega_i)_j \neq 0 \quad (j = t_i + 1, \ldots, r_i). \end{aligned}$$

Then by (24c),

$$\Delta(\omega_i)_j = 0 \quad (j = 1, \ldots, t_i), \tag{27}$$

$$\Delta(\lambda_i)_j = 0 \quad (j = t_i + 1, \ldots, r_i). \tag{28}$$

For $i = 1, \ldots, n$, we write $S_i \in \mathfrak{l}_i$ as

$$S_i = \sum_{j<l} \left[L\left((\mathbf{d}_i)_j \right), L(\xi_{i(jl)}) \right], \quad \xi_{i(jl)} \in \left(\mathcal{J}_i \right)_{jl}.$$

Then

$$\begin{aligned} S_i \sum_{j=1}^{r_i} (\lambda_i)_j (\mathbf{d}_i)_j &= \frac{1}{4} \sum_{j,l \in \Gamma_1} \left[(\lambda_i)_l - (\lambda_i)_j \right] \xi_{i(jl)} - \frac{1}{4} \sum_{j,l \in \Gamma_{\frac{1}{2}}} (\lambda_i)_j \xi_{i(jl)}, \\ S_i \sum_{j=1}^{r_i} (\omega_i)_j (\mathbf{d}_i)_j &= \frac{1}{4} \sum_{j,l \in \Gamma_{\frac{1}{2}}} (\omega_i)_l \xi_{i(jl)} + \frac{1}{4} \sum_{j,l \in \Gamma_0} \left[(\omega_i)_l - (\omega_i)_j \right] \xi_{i(jl)}. \end{aligned}$$

We now calculate the inner product of (24b) and $\sum_{i=1}^{n} \sum_{j=1}^{r_i} \Delta(\lambda_i)_j (\mathbf{d}_i)_j$. Since the Peirce decomposition is orthogonal with respect to the inner product,

$$\begin{aligned} &\left\langle S_i \sum_{j=1}^{r_i} (\omega_i)_j (\mathbf{d}_i)_j, \sum_{j=1}^{r_i} \Delta(\lambda_i)_j (\mathbf{d}_i)_j \right\rangle = 0, \\ &\left\langle \sum_{j=1}^{r_i} \Delta(\omega_i)_j (\mathbf{d}_i)_j, \sum_{j=1}^{r_i} \Delta(\lambda_i)_j (\mathbf{d}_i)_j \right\rangle = 0. \end{aligned}$$

Therefore, the inner product reduces to

$$0 = \left\langle B^\top \Delta \mathbf{y}, \sum_{i=1}^{n} \sum_{j=1}^{r_i} \Delta(\lambda_i)_j (\mathbf{d}_i)_j \right\rangle = \left\langle \Delta \mathbf{y}, B \sum_{i=1}^{n} \sum_{j=1}^{r_i} \Delta(\lambda_i)_j (\mathbf{d}_i)_j \right\rangle = \left\langle \Delta \mathbf{y}, BS \sum_{i=1}^{n} \sum_{j=1}^{r_i} (\lambda_i)_j (\mathbf{d}_i)_j \right\rangle. \tag{29}$$

The last equality is from (24a). Next we evaluate the inner product of (24b) and $S \sum_{i=1}^{n} \sum_{j=1}^{r_i} (\lambda_i)_j (\mathbf{d}_i)_j$. By orthogonality with respect to the inner product of the decomposition,

$$\left\langle \sum_{j=1}^{r_i} \Delta(\omega_i)_j (\mathbf{d}_i)_j, S_i \sum_{j=1}^{r_i} (\lambda_i)_j (\mathbf{d}_i)_j \right\rangle = 0.$$

Along with (29):

$$0 = \left\langle S\sum_{i=1}^{n}\sum_{j=1}^{r_i}(\omega_i)_j(\mathbf{d}_i)_j, S\sum_{i=1}^{n}\sum_{j=1}^{r_i}(\lambda_i)_j(\mathbf{d}_i)_j\right\rangle = -\frac{1}{16}\sum_{i=1}^{n}\sum_{j,l\in\Gamma_{\frac{1}{2}}}\left\langle(\omega_i)_l\xi_{i(jl)}(\lambda_i)_j\xi_{i(jl)}\right\rangle.$$

For $i = 1,\dots,n$ and $j,l \in \Gamma_{\frac{1}{2}}$: $(\lambda_i)_j(\omega_i)_l > 0$. Therefore,

$$\xi_{i(jl)} = 0, \quad j,l \in \Gamma_{\frac{1}{2}}. \tag{30}$$

Hence

$$S_i\sum_{j=1}^{r_i}(\lambda_i)_j(\mathbf{d}_i)_j = \frac{1}{4}\sum_{j,l\in\Gamma_1}\left[(\lambda_i)_l - (\lambda_i)_j\right]\xi_{i(jl)}, \tag{31}$$

$$S_i\sum_{j=1}^{r_i}(\omega_i)_j(\mathbf{d}_i)_j = \frac{1}{4}\sum_{j,l\in\Gamma_0}\left[(\omega_i)_l - (\omega_i)_j\right]\xi_{i(jl)}. \tag{32}$$

(31) and (28) imply

$$\sum_{i=1}^{n}\sum_{j=1}^{r_i}\left[S(\lambda_i)_j(\mathbf{d}_i)_j + \Delta(\lambda_i)_j(\mathbf{d}_i)_j\right] \in \sum_{i=1}^{n}\mathcal{J}_0\Big(e\big(\mathbf{z}_i\big)\Big),$$

which along with dual nondegeneracy and (24a) implies

$$\sum_{i=1}^{n}\sum_{j=1}^{r_i}\left[S(\lambda_i)_j(\mathbf{d}_i)_j + \Delta(\lambda_i)_j(\mathbf{d}_i)_j\right] = 0.$$

Since $(\lambda_i)_l - (\lambda_i)_j \neq 0$, and the decomposition is a direct sum, (31) yields

$$\xi_{i(jl)} = 0 \quad (j,l\in\Gamma_1), \tag{33}$$

$$\Delta\lambda = 0. \tag{34}$$

Similarly, primal nondegeneracy, (27), (32), and (24b) implies

$$B^\top\Delta\mathbf{y} = 0, \tag{35}$$

$$S_i\sum_{j=1}^{r_i}(\omega_i)_j(\mathbf{d}_i)_j + \sum_{i=1}^{r_i}\Delta(\omega_i)_j(\mathbf{d}_i)_j = 0. \tag{36}$$

Since A is onto, and Q is one-to-one, (35) implies $\Delta\mathbf{y} = 0$. Similarly (36) implies

$$\xi_{i(jl)} = 0 \quad (j,l\in\Gamma_0), \tag{37}$$

$$\Delta\omega = 0. \tag{38}$$

Combining (30), (33), (37) results in

$$\xi_{i(jl)} = 0, \quad (j<l).$$

Finally from

$$S = \sum_{i=1}^{n}\sum_{j=1}^{r_i}\left[L\left((\mathbf{d}_i)_j\right), L\left(\xi_{i(jl)}\right)\right],$$

we get $S = 0$. Thus, We have proved that the only solution of (24) is zero. ■

The preceding theorem implies that the Q method, under nondegeneracy and strict complementary and regularity conditions will be numerically stable near the optimal solution.

4 An Infeasible Interior Point Algorithm

In this section, we present an infeasible interior point algorithm which is a based on the work of Kojima, Megiddo and Mizuno, and Freund and Jarre,[15, 9]. These authors develop their algorithms for linear programming, and they are based on methods that do not require primal or dual feasibility. This feature turns out to be essential for the Q method. Recall that in the Q method the primal and dual feasibility conditions are nonlinear. Therefore, even if the current iterate satisfy these relations, there is no guarantee that subsequent iterations will. All we can hope for is the iterates converge to feasible solutions. Below we adopt and extend the results of the two references mentioned above to the Q method. We describe the main algorithm in §§ 4.1. The convergence proof is presented in §§ 4.2. And to ensure global convergence, the algorithm is modified in §§ 4.3.

We use the following definition of norm of linear transformation defined on a linear space E:

$$\|L\| \stackrel{\text{def}}{=} \sup_{\substack{\mathbf{x}\neq\mathbf{0},\\ \mathbf{x}\in E}} \frac{\|L\mathbf{x}\|}{\|\mathbf{x}\|}.$$

4.1 Extension of Kojima-Megiddo-Mizuno (KMM) Algorithm to the Q method

The algorithm of [15] is designed for linear programming starting from possibly infeasible points, but with exact search directions. As we mentioned earlier the Newton system for the Q method is nonlinear and thus the search directions are not exact. This algorithm can start from an arbitrary infeasible interior point. Therefore it does not require a first phase; nor does it need to transform the problem into artificial format such as the one employed by the so called "Big M" method. As a result we expect better numerical behavior.

Accuracy measures for primal and dual infeasibility and complementarity can be chosen separately; primal and dual step sizes can be different. Given $\epsilon_p > 0$, $\epsilon_d > 0$, $\epsilon_c > 0$, we want to find an approximate solution of the symmetric cone program (20) such that

$$\|A\mathbf{x} - b\| \leq \epsilon_p, \quad \|A^\top \mathbf{y} + \mathbf{z} - \mathbf{c}\| \leq \epsilon_d, \quad \langle \mathbf{x}, \mathbf{z}\rangle \leq \epsilon_c.$$

Note that $\langle \mathbf{x}, \mathbf{z}\rangle = \lambda^T \omega$, where λ and ω are vectors of eigenvalues of $\mathbf{x}$ and $\mathbf{z}$ ordered by their shared Jordan frame.

We define a proximity neighborhood around the central path:

$$\begin{aligned}\mathcal{N}(\gamma_c, \gamma_p, \gamma_d) \stackrel{\text{def}}{=} \Big\{ &(\lambda, \omega, \mathbf{y}, Q) \colon \lambda \in \mathbb{R}^r, \omega \in \mathbb{R}^r, \mathbf{y} \in Y, Q \in L_{\mathcal{K}},\ \lambda > 0,\ \omega > 0, \\ &(\lambda_i)_j (\omega_i)_j \geq \gamma_c \frac{\lambda^T \omega}{r}\ (j = 1, \ldots, r_i; i = 1, \ldots, n), \\ &\lambda^T \omega \geq \gamma_p \|A\mathbf{x} - \mathbf{b}\| \text{ or } \|A\mathbf{x} - \mathbf{b}\| \leq \epsilon_p, \\ &\lambda^T \omega \geq \gamma_d \left\|A^\top \mathbf{y} + \mathbf{z} - \mathbf{c}\right\| \text{ or } \left\|A^\top \mathbf{y} + \mathbf{z} - \mathbf{c}\right\| \leq \epsilon_d. \Big\}\end{aligned}$$

The first inequality is the centrality condition (recall that $r = \mathrm{rk}(\mathcal{J})$). It prevents the iterates from hitting the boundary of $\mathcal{K}$ before reaching the optimum. The second and third inequalities guarantee that the complementarity not be achieved before the primal or the dual feasibility. Obviously, when $(\gamma_c', \gamma_p', \gamma_d') \leq (\gamma_c, \gamma_p, \gamma_d)$,

$$\mathcal{N}(\gamma_c, \gamma_p, \gamma_d) \subseteq \mathcal{N}(\gamma_c', \gamma_p', \gamma_d').$$

And

$$\bigcup_{(\gamma_c, \gamma_p, \gamma_d) > 0} \mathcal{N}(\gamma_c, \gamma_p, \gamma_d) = \{(\lambda, \omega, \mathbf{y}, Q) \colon \lambda > 0, \omega > 0\}.$$

Clearly, for a sequence of points in $\mathcal{N}(\gamma_c, \gamma_p, \gamma_d)$ approach the optimum when $\lambda^T\omega$ approaches 0. We now present the adoption of Kojima, Megiddo and Mizuno algorithm to the Q method.

The extended KMM Algorithm for the Q method

Choose $0 < \sigma_1 < \sigma_2 < \sigma_3 < 1$ and $\Upsilon > 0$. To start from an arbitrary point $(\lambda^0, \omega^0, \mathbf{y}^0, Q^0)$, one may select $0 < \gamma_c < 1, \gamma_p > 0, \gamma_d > 0$, so that $(\lambda^0, \omega^0, \mathbf{y}^0, Q^0) \in \mathcal{N}(\gamma_c, \gamma_p, \gamma_d)$.
Do until (1) $\|\mathbf{r}_p^k\| < \epsilon_p$, $\|\mathbf{r}_d^k\| < \epsilon_d$, and ${\lambda^k}^T\omega^k < \epsilon_c$; or (2) $\left\|(\lambda^k, \omega^k)\right\|_1 > \Upsilon$.

1. Set $\mu = \sigma_1 \frac{{\lambda^k}^T\omega^k}{r}$.
2. Compute the search direction $(\Delta\lambda, \Delta\omega, \Delta\mathbf{y}, S)$ from (24).
3. Choose step sizes α, β, θ, γ, set
$$\begin{aligned}\lambda^{k+1} &= \lambda^k + \alpha\Delta\lambda,\\ \mathbf{y}^{k+1} &= \mathbf{y}^k + \theta\Delta\mathbf{y},\\ \omega^{k+1} &= \omega^k + \beta\Delta\omega,\\ Q^{k+1} &= Q^k \exp(\gamma S).\end{aligned}$$
4. $k \leftarrow k + 1$.

End
The step sizes are chosen in the following manner. Let $\hat{\alpha}^k$ be the maximum of $\tilde{\alpha} \in [0, 1]$, so that for any $\alpha \in [0, \tilde{\alpha}]$:
$$\begin{gathered}\left(\lambda^k + \alpha\Delta\lambda, \omega^k + \alpha\Delta\omega, \mathbf{y}^k + \alpha\Delta\mathbf{y}, Q^k \exp(\alpha S)\right) \in \mathcal{N},\\ (\lambda^k + \alpha\Delta\lambda)^T(\omega^k + \alpha\Delta\omega) \le [1 - \alpha(1 - \sigma_2)]\, {\lambda^k}^T\omega^k.\end{gathered}$$
Choose $\alpha \in (0, 1]$, $\theta \in (0, 1]$, $\beta \in (0, 1]$, $\gamma \in (0, 1]$ so that
$$\begin{gathered}(\lambda^{k+1}, \omega^{k+1}, \mathbf{y}^{k+1}, Q^{k+1}) \in \mathcal{N}(\gamma_c, \gamma_p, \gamma_d),\\ {\lambda^{k+1}}^T\omega^{k+1} \le \left[1 - \hat{\alpha}^k(1 - \sigma_3)\right] {\lambda^k}^T\omega^k.\end{gathered}$$
Because $\sigma_1 < \sigma_2 < \sigma_3$, the step sizes for λ, ω, $\mathbf{y}$, and Q may not necessarily be the same.

4.2 Convergence Results

To prove the global convergence of the preceding algorithm, as in [15], we need to show the boundedness of the step sizes.

For an operator T on $\mathcal{J}$,

$$\|\exp(T) - I - T\| \le \sum_{j=2}^{\infty} \frac{\|T\|^j}{j!} \le \sum_{j=1}^{\infty} \frac{\|T\|^j}{j!}\|T\| \le \exp(\|T\|)\|T\|^2. \tag{39}$$

Following the notation used in [15], for each k, define

$$f_{ij} \stackrel{\text{def}}{=} \left[(\lambda_i^k)_j + \alpha\Delta(\lambda_i)_j\right]\left[(\omega_i^k)_j + \alpha\Delta(\omega_i)_j\right] - \frac{\gamma_c}{r}(\lambda^k + \alpha\Delta\lambda)^T(\omega^k + \alpha\Delta\omega),$$

$$g_p(\alpha) \stackrel{\text{def}}{=} (\lambda^k + \alpha\Delta\lambda)^T(\omega^k + \alpha\Delta\omega) - \gamma_p\left\|B^k\exp(\alpha S)\sum_{i=1}^{n}\sum_{j=1}^{r_i}\left[(\lambda_i^k)_j + \alpha\Delta(\lambda_i)_j\right](\mathbf{d}_i)_j - \mathbf{b}\right\|,$$

$$g_d(\alpha) \stackrel{\text{def}}{=} (\lambda^k + \alpha\Delta\lambda)^T(\omega^k + \alpha\Delta\omega)$$
$$-\gamma_d\left\|(B^k)^\top\left(\mathbf{y}^k + \alpha\Delta\mathbf{y}\right) + \exp(\alpha S)\sum_{i=1}^{n}\sum_{j=1}^{r_i}\left[(\omega_i^k)_j + \alpha\Delta(\omega_i)_j\right](\mathbf{d}_i)_j - (Q^k)^\top\mathbf{c}\right\|,$$

$$h(\alpha) \stackrel{\text{def}}{=} \left[1 - \alpha(1-\sigma_2)\right]{\lambda^k}^T\omega^k - (\lambda^k + \alpha\Delta\lambda)^T(\omega^k + \alpha\Delta\omega).$$

Therefore, $\hat{\alpha}^k$ is determined by the following inequalities:

$$\begin{aligned} &f_{ij}(\alpha) \geq 0 \quad (j = 1, \ldots, r_i;\ i = 1, \ldots, n),\\ &g_p(\alpha) \geq 0 \text{ or } \|\mathbf{r}_p^k\| \leq \epsilon_p,\\ &g_d(\alpha) \geq 0 \text{ or } \|\mathbf{r}_d^k\| \leq \epsilon_p,\\ &h(\alpha) \geq 0. \end{aligned}$$

Let $\epsilon^* \stackrel{\text{def}}{=} \min(\epsilon_c, \gamma_p\epsilon_p, \gamma_d\epsilon_d)$. Then for each intermediate iterate:

$${\lambda^k}^T\omega^k \geq \epsilon^*, \quad \left\|(\lambda^k, \omega^k)\right\|_1 \leq \Upsilon.$$

Suppose the solutions to (24) are uniformly upper bounded for all the iterations. Then there exists $\eta \geq 0$ such that

$$\left|\Delta(\lambda_i)_j\Delta(\omega_i)_j - \frac{\gamma_c}{r}\Delta\lambda^T\Delta\omega\right| \leq \eta, \quad |\Delta\lambda^T\Delta\omega| \leq \eta,$$

$$\|A\|\left\|S\sum_{i=1}^{n}\sum_{j=1}^{r_i}\Delta(\lambda_i)_j(\mathbf{d}_i)_j\right\| \leq \eta,$$

$$\|A\|\exp(\|\alpha S\|)\|S\|^2\left\|\sum_{i=1}^{n}\sum_{j=1}^{r_i}\left[(\lambda_i^k)_j + \alpha\Delta(\lambda_i)_j\right](\mathbf{d}_i)_j\right\| \leq \eta\ (0 \leq \alpha \leq 1),$$

$$\left\|S\sum_{i=1}^{n}\sum_{j=1}^{r_i}\Delta(\omega_i)_j(\mathbf{d}_i)_j\right\| \leq \eta,$$

$$\exp(\|\alpha S\|)\|S\|^2\left\|\sum_{i=1}^{n}\sum_{j=1}^{r_i}\left[(\omega_i^k)_j + \alpha\Delta(\omega_i)_j\right](\mathbf{d}_i)_j\right\| \leq \eta\ (0 \leq \alpha \leq 1).$$

We first determine a lower bound for α from $g_p(\alpha)$.

When $\|\mathbf{r}_p^k\| > \epsilon_p$:

$$\begin{aligned} g_p(\alpha) \geq (1-\alpha){\lambda^k}^T\omega^k + \alpha\sigma_1{\lambda^k}^T\omega^k + \alpha^2\Delta\lambda^T\Delta\omega \\ - \gamma_p(1-\alpha)\left\|\mathbf{b} - B^k\sum_{i=1}^{n}\sum_{j=1}^{r_i}(\lambda_i^k)_j(\mathbf{d}_i)_j\right\| - \alpha^2\gamma_p\left\|B^kS\sum_{i=1}^{n}\sum_{j=1}^{r_i}\Delta(\lambda_i)_j(\mathbf{d}_i)_j\right\| \\ - \alpha^2\gamma_p\|B^k\|\exp(\alpha\|S\|)\|S\|^2\left\|\sum_{i=1}^{n}\sum_{j=1}^{r_i}\left[(\lambda_i^k)_j + \alpha\Delta(\lambda_i)_j\right](\mathbf{d}_i)_j\right\| \geq \alpha\sigma_1\epsilon^* - \alpha^2\eta - 2\alpha^2\gamma_p\eta. \end{aligned}$$

The first inequality is from (39), (24c), and (24a); the second inequality results from the fact that the k^{th} iterate is in $\mathcal{N}$, Q is orthogonal and does not change the Euclidean norm, and from the definitions of ϵ^* and η. When $\|\mathbf{r}_p^k\| \leq \epsilon_p$:

$$\begin{aligned} \left\|AQ^k\exp(\alpha S)\sum_{i=1}^{n}\sum_{j=1}^{r_i}\left[(\lambda_i^k)_j + \alpha\Delta(\lambda_i)_j\right](\mathbf{d}_i)_j - \mathbf{b}\right\| \\ \leq (1-\alpha)\left\|AQ^k\sum_{i=1}^{n}\sum_{j=1}^{r_i}(\lambda_i^k)_j(\mathbf{d}_i)_j - \mathbf{b}\right\| + \alpha^2\left\|AQ^kS\sum_{i=1}^{n}\sum_{j=1}^{r_i}\Delta(\lambda_i)_j(\mathbf{d}_i)_j\right\| \\ + \alpha^2\|AQ^k\|\exp(\alpha\|S\|)\|S\|^2\left\|\sum_{i=1}^{n}\sum_{j=1}^{r_i}\left[(\lambda_i^k)_j + \alpha\Delta(\lambda_i)_j\right](\mathbf{d}_i)_j\right\| \leq (1-\alpha)\epsilon_p + 2\alpha^2\eta. \end{aligned}$$

Again, the first inequality is due to (24a) and (39), the last one because of the definition of η, orthogonality of Q and $\|\mathbf{r}_p^k\| \leq \epsilon_p$. Therefore,

$$\alpha \leq \min\left\{\frac{\sigma_1\epsilon^*}{\eta + 2\gamma_p\eta}, \frac{\epsilon_p}{2\eta}\right\}.$$

Now consider $g_d(\alpha)$.
When $\|\mathbf{r}_d^k\| > \epsilon_d$:

$$\begin{aligned} g_d(\alpha) \geq (1-\alpha){\lambda^k}^T\omega^k + \alpha\sigma_1{\lambda^k}^T\omega^k + \alpha^2\Delta\lambda^T\Delta\omega - \gamma_d(1-\alpha)\|\mathbf{r}_d^k\| - \alpha^2\gamma_d\left\|S\sum_{i=1}^{n}\sum_{j=1}^{r_i}\Delta(\omega_i)_j(\mathbf{d}_i)_j\right\| \\ - \alpha^2\gamma_d\exp(\alpha\|S\|)\|S\|^2\left\|\sum_{i=1}^{n}\sum_{j=1}^{r_i}\left[(\omega_i^k)_j + \alpha\Delta(\omega_i)_j\right](\mathbf{d}_i)_j\right\| \geq \alpha\sigma_1\epsilon^* - \alpha^2\eta - 2\alpha^2\gamma_d\eta. \end{aligned}$$

The first inequality is due to (39), (24c), and (24b); the second inequality is from the fact that the k^{th} iterate is in the neighborhood $\mathcal{N}$ and the definitions of ϵ^* and η.

When $\|\mathbf{r}_d^k\| \le \epsilon_d$:

$$\left\| A^\top(\mathbf{y}^k + \alpha\Delta\mathbf{y}) + Q^k \exp(\alpha S) \sum_{i=1}^{n}\sum_{j=1}^{r_i} \left[(\omega_i^k)_j + \alpha\Delta(\omega_i)_j\right] (\mathbf{d}_i)_j - \mathbf{c} \right\|$$
$$\le (1-\alpha)\left\| A^\top\mathbf{y}^k + Q^k \sum_{i=1}^{n}\sum_{j=1}^{r_i} (\omega_i^k)_j(\mathbf{d}_i)_j - \mathbf{c}\right\| + \alpha^2 \left\| Q^k S \sum_{i=1}^{n}\sum_{j=1}^{r_i} \Delta(\omega_i)_j(\mathbf{d}_i)_j \right\|$$
$$+ \alpha^2 \exp(\alpha\|S\|)\|S\|^2 \left\| \sum_{i=1}^{n}\sum_{j=1}^{r_i} \left[(\omega_i^k)_j + \alpha\Delta(\omega_i)_j\right] (\mathbf{d}_i)_j \right\| \le (1-\alpha)\epsilon_d + 2\alpha^2\eta.$$

The first inequality is due to (24b) and (39); the last one is because of the definition of η and $\|\mathbf{r}_d^k\| \le \epsilon_d$. Therefore,

$$\alpha \le \min\left\{ \frac{\sigma_1\epsilon^*}{\eta + 2\gamma_d\eta}, \frac{\epsilon_d}{2\eta} \right\}.$$

By the same arguments as those in [15]:

$$f_{ij}(\alpha) \ge \sigma_1 \frac{\epsilon^*}{r}(1-\gamma_c)\alpha - \eta\alpha^2,$$
$$h(\alpha) \ge (\sigma_2 - \sigma_1)\epsilon^*\alpha - \eta\alpha^2.$$

So there is a lower bound for the step size.

$$\hat{\alpha}^k = \min\left\{ 1, \frac{\sigma_1\epsilon^*}{\eta + 2\gamma_p\eta}, \frac{\epsilon_p}{2\eta}, \frac{\sigma_1\epsilon^*}{\eta + 2\gamma_d\eta}, \frac{\epsilon_d}{2\eta}, \frac{\sigma_1\epsilon^*}{r\eta}(1-\gamma_c), (\sigma_2-\sigma_1)\frac{\epsilon^*}{\eta} \right\}.$$

By Lemma 3.2, F^k is a bijection for regular iterates. Assume the initial point is regular. Then the subsequent iterates can always be made regular by perturbation of step sizes. Furthermore, the step sizes can always be at least half of the original ones.

Now we are ready to prove the main convergence result.

Theorem 4.1 *Suppose there exists $d > 0$ such that $\forall\, N > 0$, $\exists\, n \ge N$ so that for all unit length vectors $\mathbf{w}$, $\|F^n\mathbf{w}\| \ge d$. Then algorithm KMM must stop after finite steps.*

Proof: Assume the algorithm doesn't stop after finite iterations. Then for each $k > 0$, we have

$${\lambda^k}^T\omega^k \ge \epsilon^* \quad \text{and} \quad \left\|(\lambda^k, \omega^k)\right\|_1 \le \Upsilon,$$

because otherwise, the iteration will terminate due to the stopping criteria. Boundedness of $\mathbf{y}^k$ is due to the dual feasible constraint. Also observe that Q^k is orthogonal; so its norm is 1. If the conditions of the theorem are satisfied, there must exist a subsequence $\{(\lambda^{m_i}, \omega^{m_i}, \mathbf{y}^{m_i}, Q^{m_i})\}_{i=1}^{\infty}$ such that for all m_i, $\|(F^{m_i})^{-1}\|$ is upper bounded. The right-hand side of (24) depends continuously on $(\lambda^{m_i}, \omega^{m_i}, \mathbf{y}^{m_i}, Q^{m_i})$. So the norm of the right-hand side of (24) is upper bounded. Therefore, a uniform upper bound on the solutions to (24) for the subsequence $\{m_i\}$ exists.

By the analysis above this theorem, there's a lower bound α^* for $\hat{\alpha}^{m_i}$. After the perturbations of step sizes to ensure the regularity of $\mathbf{x}$ and $\mathbf{z}$, the lower bound on $\hat{\alpha}^k$ is at least $\frac{\alpha^*}{2}$. The algorithm imposes the decrease of the sequence $\left\{(\lambda^j)^\top\omega^j\right\}_{j=1}^{\infty}$. So for each m_i in the subsequence, by $h(\alpha) \ge 0$, we get

$$\lambda^{m_i+1^T}\omega^{m_i+1} \le \left[1-\frac{\alpha^*}{2}(1-\sigma_3)\right]\lambda^{m_i T}\omega^{m_i} \le \left[1-\frac{\alpha^*}{2}(1-\sigma_3)\right]\lambda^{m_{i-1}+1^T}\omega^{m_{i-1}+1}$$
$$\le \left[1-\frac{\alpha^*}{2}(1-\sigma_3)\right]^2\lambda^{m_{i-1}T}\omega^{m_{i-1}} \le \cdots \le \left[1-\frac{\alpha^*}{2}(1-\sigma_3)\right]^i\lambda^{m_1 T}\omega^{m_1}.$$

This means that the sequence $\{(\lambda^j)^\top\omega^j\}_{j=1}^\infty$ converges to 0, contradicting the assumption that the algorithm doesn't stop after finite steps. ■

4.3 Boundedness of Iterates

In the analysis above, our KMM based algorithm may abort due to unboundedness of variables. To ensure that each iterate is bounded, we modify the algorithm of the previous section by extending the ideas of [9] from linear programming. First we describe the algorithm and then give the convergence analysis.

Extension of Algorithm of Freund and Jarre to the Q Method

Suppose that $(\hat{\mathbf{x}}, \hat{\mathbf{z}}, \hat{\mathbf{y}})$ is an interior feasible solution of the (20) and that the eigenvalues of $\hat{\mathbf{x}}$ and $\hat{\mathbf{z}}$ satisfy $\delta_p\mathbf{1} \le \hat{\lambda} \le \chi_p\mathbf{1}$, $\delta_d\mathbf{1} \le \hat{\omega} \le \chi_d\mathbf{1}$. Since A is continuous and surjective, there exist $\zeta_p > 0$ and $\zeta_d > 0$, such that $\forall \|\tilde{\mathbf{b}}\| \le \zeta_p$, $\|\tilde{\mathbf{c}}\| \le \zeta_d$, the system

$$\begin{aligned} &A\mathbf{x} = \tilde{\mathbf{b}} \\ &A^\top\mathbf{y} + \mathbf{z} = \tilde{\mathbf{c}} \end{aligned} \tag{40}$$

has a solution $(\tilde{\mathbf{x}}, \tilde{\mathbf{z}}, \tilde{\mathbf{y}})$ with its primal and dual eigenvalues satisfying $\left\|\tilde{\lambda}\right\|_\infty \le \frac{1}{2}\delta_p$ and $\|\tilde{\omega}\|_\infty \le \frac{1}{2}\delta_d$. For instance, let A^+ denote the Moore-Penrose generalized inverse of A. Let $\mathbf{h} \stackrel{\text{def}}{=} A^+\mathbf{b}$. Let λ_h be the eigenvalues of $\mathbf{h}$. Note that $\|\mathbf{h}\| = \sqrt{\operatorname{tr}(\mathbf{h}^2)} = \|\mathbf{h}\|_2$. Then

$$\|\lambda_h\|_\infty \le \|\lambda_h\|_2 = \|\mathbf{h}\| \le \|A^+\|\|\tilde{\mathbf{b}}\|.$$

Obviously, $(\mathbf{h}, \mathbf{0}, \tilde{\mathbf{c}})$ is a solution to (40). So one can set $\zeta_p = \frac{1}{2}\frac{\delta_p}{\|A^+\|}$, $\zeta_d = \frac{1}{2}\delta_d$.

If $\epsilon_p > \zeta_p$, we replace ϵ_p with ζ_p; if $\epsilon_d > \zeta_d$, we replace ϵ_d with ζ_d. The neighborhood $\tilde{\mathcal{N}}$ is defined as follows:

$$\begin{aligned} \tilde{\mathcal{N}} \stackrel{\text{def}}{=} \Big\{ &(\lambda, \omega, \mathbf{y}, Q) \mid \lambda \in \mathbb{R}^r, \omega \in \mathbb{R}^r, \mathbf{y} \in \mathbb{R}^m, Q \in L_{\mathcal{K}},\ \lambda > 0,\ \omega > 0; \\ &(\lambda_i)_j(\omega_i)_j \ge \gamma_c\frac{\lambda^T\omega}{r}\ (j = 1, \ldots, r_i;\ i = 1, \ldots, n); \\ &\lambda^T\omega \ge \gamma_p\|A\mathbf{x} - \mathbf{b}\| \text{ and } \|A\mathbf{x} - \mathbf{b}\| \le \zeta_p, \text{ or } \|A\mathbf{x} - \mathbf{b}\| \le \epsilon_p; \\ &\lambda^T\omega \ge \gamma_d\left\|A^\top\mathbf{y} + \mathbf{z} - \mathbf{c}\right\| \text{ and } \left\|A^\top\mathbf{y} + \mathbf{z} - \mathbf{c}\right\| \le \zeta_d, \text{ or } \left\|A^\top\mathbf{y} + \mathbf{z} - \mathbf{c}\right\| \le \epsilon_d. \Big\} \end{aligned}$$

Other parts of the algorithm are the same as in the previous section.

Bounds on step sizes

Next we give a lower bound on the step size α to ensure that each iterate is in $\tilde{N}$.

Assume $\|\mathbf{r}_p^k\| \le \zeta_p$. Then

$$\begin{aligned}
&\left\| AQ^k \exp(\alpha S) \sum_{i=1}^{n} \sum_{j=1}^{r_i} \left[(\lambda_i^k)_j + \alpha \Delta(\lambda_i)_j \right] (\mathbf{d}_i)_j - \mathbf{b} \right\| \\
&\qquad \le (1-\alpha) \left\| AQ^k \sum_{i=1}^{n} \sum_{j=1}^{r_i} (\lambda_i^k)_j (\mathbf{d}_i)_j - \mathbf{b} \right\| + \alpha^2 \left\| AQ^k S \sum_{i=1}^{n} \sum_{j=1}^{r_i} \Delta(\lambda_i)_j (\mathbf{d}_i)_j \right\| \\
&\qquad + \alpha^2 \|AQ^k\| \exp(\alpha \|S\|) \|S\|^2 \left\| \sum_{i=1}^{n} \sum_{j=1}^{r_i} \left[(\lambda_i^k)_j + \alpha \Delta(\lambda_i)_j \right] (\mathbf{d}_i)_j \right\| \le (1-\alpha)\zeta_p + 2\alpha^2 \eta.
\end{aligned}$$

The first inequality is from (24a) and (39), and the last one form the definition of η and $\|\mathbf{r}_p^k\| \le \zeta_p$. Similarly, assume $\|\mathbf{r}_p^k\| \le \zeta_p$:

$$\begin{aligned}
&\left\| A^\top (\mathbf{y}^k + \alpha \Delta \mathbf{y}) + Q^k \exp(\alpha S) \sum_{i=1}^{n} \sum_{j=1}^{r_i} \left[(\omega_i^k)_j + \alpha \Delta(\omega_i)_j \right] (\mathbf{d}_i)_j - \mathbf{c} \right\| \\
&\qquad \le (1-\alpha) \left\| A^\top \mathbf{y}^k + Q^k \sum_{i=1}^{n} \sum_{j=1}^{r_i} (\omega_i^k)_j (\mathbf{d}_i)_j - \mathbf{c} \right\| + \alpha^2 \left\| Q^k S \sum_{i=1}^{n} \sum_{j=1}^{r_i} \Delta(\omega_i)_j (\mathbf{d}_i)_j \right\| \\
&\qquad + \alpha^2 \exp(\alpha \|S\|) \|S\|^2 \left\| \sum_{i=1}^{n} \sum_{j=1}^{r_i} \left[(\omega_i^k)_j + \alpha \Delta(\omega_i)_j \right] (\mathbf{d}_i)_j \right\| \le (1-\alpha)\zeta_d + 2\alpha^2 \eta.
\end{aligned}$$

The first inequality is from (24b) and (39), and the last one is from the definition of η and $\|\mathbf{r}_d^k\| \le \zeta_d$.

Let α^* be the lower bound of the step size of the extended KMM algorithm in the previous section. Then

$$\alpha^{**} \stackrel{\text{def}}{=} \min\left(\frac{\zeta_p}{2\eta}, \frac{\zeta_d}{2\eta}, \alpha^* \right)$$

is a lower bound for the modified algorithm of this section. Therefore, the analysis of the last section carries over to the algorithm presented in this section.

boundedness of iterates

Consider the perturbed system:

$$\begin{aligned}
&\mathbf{z} + A^T \mathbf{y} = \mathbf{c} + \tilde{\mathbf{c}} \\
&A\mathbf{x} = \mathbf{b} + \tilde{\mathbf{b}} \\
&\mathbf{x} \succcurlyeq \mathbf{0} \\
&\mathbf{z} \succcurlyeq \mathbf{0}.
\end{aligned} \tag{41}$$

Assume $\|\tilde{\mathbf{b}}\| \le \zeta_p$, $\|\tilde{\mathbf{c}}\| \le \zeta_d$. Then (40) has a solution $(\tilde{\mathbf{x}}, \tilde{\mathbf{z}}, \tilde{\mathbf{y}})$ with its primal and dual eigenvalues satisfying $\left\|\tilde{\lambda}\right\|_\infty \le \frac{1}{2}\delta_p$ and $\|\tilde{\omega}\|_\infty \le \frac{1}{2}\delta_d$. Let

$$\breve{\mathbf{x}} \stackrel{\text{def}}{=} \hat{\mathbf{x}} + \tilde{\mathbf{x}}, \quad \breve{\mathbf{y}} \stackrel{\text{def}}{=} \hat{\mathbf{y}}, \quad \breve{\mathbf{z}} \stackrel{\text{def}}{=} \hat{\mathbf{z}} + \tilde{\mathbf{z}}.$$

Since $\delta_p \mathbf{1} \le \hat{\lambda} \le \chi_p \mathbf{1}$, we have

$$\hat{\mathbf{x}} - \delta_p \mathbf{e} \succcurlyeq \mathbf{0}, \qquad \chi_p \mathbf{e} - \hat{\mathbf{x}} \succcurlyeq \mathbf{0}.$$

Similarly, because $|\tilde{\lambda}| \leq \frac{1}{2}\delta_p\mathbf{1}$, we get

$$\frac{1}{2}\delta_p\mathbf{e} + \tilde{\mathbf{x}} \succcurlyeq \mathbf{0}, \qquad \frac{1}{2}\delta_p\mathbf{e} - \tilde{\mathbf{x}} \succcurlyeq \mathbf{0}.$$

By convexity of $\mathcal{K}$,

$$(\hat{\mathbf{x}} - \delta_p\mathbf{e}) + \left(\frac{1}{2}\delta_p\mathbf{e} + \tilde{\mathbf{x}}\right) \succcurlyeq \mathbf{0}, \qquad (\chi_p\mathbf{e} - \hat{\mathbf{x}}) + \left(\frac{1}{2}\delta_p\mathbf{e} - \tilde{\mathbf{x}}\right) \succcurlyeq \mathbf{0}.$$

The above inequalities imply:

$$\frac{1}{2}\delta_p\mathbf{1} \leq \breve{\lambda} \leq \chi_p\mathbf{1} + \frac{1}{2}\delta_p\mathbf{1}.$$
$$\frac{1}{2}\delta_d\mathbf{1} \leq \breve{\omega} \leq \chi_d\mathbf{1} + \frac{1}{2}\delta_d\mathbf{1}.$$

And we get the following

Lemma 4.1 *Each iterate in $\tilde{\mathcal{N}}$ is bounded.*

Proof: The k^{th} iterate is a solution of

$$A\mathbf{x} = \mathbf{b} + \mathbf{r}_b^k$$
$$A^\top\mathbf{y} + \mathbf{z} = \mathbf{c} + \mathbf{r}_c^k.$$

Since $\|\mathbf{r}_b^k\| \leq \delta_p$, $\|\mathbf{r}_c^k\| \leq \delta_d$, by the analysis above, there exists $(\breve{\mathbf{x}}, \breve{\mathbf{y}}, \breve{\mathbf{z}})$ satisfying the above perturbed constraints with

$$\frac{1}{2}\delta_p\mathbf{1} \leq \breve{\lambda} \leq \chi_p\mathbf{1} + \frac{1}{2}\delta_p\mathbf{1}, \qquad \frac{1}{2}\delta_d\mathbf{1} \leq \breve{\omega} \leq \chi_d\mathbf{1} + \frac{1}{2}\delta_d\mathbf{1}.$$

To avoid clutter, we omit the superscript k in the remainder of the proof. Let $(\mathbf{x}, \mathbf{y}, \mathbf{z})$ denote the k^{th} iterate. Then

$$A(\mathbf{x} - \breve{\mathbf{x}}) = \mathbf{0}, \quad A^\top(\mathbf{y} - \breve{\mathbf{y}}) + \mathbf{z} - \breve{\mathbf{z}} = \mathbf{0}.$$

Hence

$$\langle \mathbf{x} - \breve{\mathbf{x}}, \mathbf{z} - \breve{\mathbf{z}} \rangle = -\langle \mathbf{x} - \breve{\mathbf{x}}, A^\top(\mathbf{y} - \breve{\mathbf{y}}) \rangle = -\langle A(\mathbf{x} - \breve{\mathbf{x}}), \mathbf{y} - \breve{\mathbf{y}} \rangle = 0.$$

Therefore,

$$\begin{aligned}
\langle \mathbf{x}, \mathbf{z} \rangle + \langle \breve{\mathbf{x}}, \breve{\mathbf{z}} \rangle &= \langle \mathbf{x}, \breve{\mathbf{z}} \rangle + \langle \breve{\mathbf{x}}, \mathbf{z} \rangle \\
&= \langle \mathbf{x}, \frac{1}{2}\delta_d\mathbf{e} \rangle + \langle \mathbf{x}, \breve{\mathbf{z}} - \frac{1}{2}\delta_d\mathbf{e} \rangle + \langle \frac{1}{2}\delta_p\mathbf{e}, \mathbf{z} \rangle + \langle \breve{\mathbf{x}} - \frac{1}{2}\delta_p\mathbf{e}, \mathbf{z} \rangle \\
&\geq \frac{1}{2}\delta_d\langle \mathbf{x}, \mathbf{e} \rangle + \frac{1}{2}\delta_p\langle \mathbf{e}, \mathbf{z} \rangle \\
&= \frac{1}{2}\delta_d \left\|\lambda\right\|_1 + \frac{1}{2}\delta_p \left\|\omega\right\|_1.
\end{aligned}$$

The inequality follows because of self-duality of $\mathcal{K}$ and $\breve{\mathbf{z}} - \frac{1}{2}\delta_d\mathbf{e} \succcurlyeq \mathbf{0}$, $\breve{\mathbf{x}} - \frac{1}{2}\delta_p\mathbf{e} \succcurlyeq \mathbf{0}$. On the other hand,

$$\begin{aligned}
\langle \mathbf{x}, \mathbf{z} \rangle + \langle \breve{\mathbf{x}}, \breve{\mathbf{z}} \rangle &\leq \lambda^T\omega + \|\breve{\mathbf{x}}\| \cdot \|\breve{\mathbf{z}}\| \\
&\leq {\lambda^0}^T\omega^0 + n\left(\chi_p + \frac{1}{2}\delta_p\right)\left(\chi_d + \frac{1}{2}\delta_d\right).
\end{aligned}$$

The last inequality is due to the fact that the duality gap is reduced at each iteration and the bounds on the eigenvalues of $\breve{\mathbf{x}}$, $\breve{\mathbf{z}}$. Combine the two inequalities:

$$\frac{1}{2}\delta_d \|\lambda\|_1 + \frac{1}{2}\delta_p \|\omega\|_1 \leq {\lambda^0}^T \omega^0 + n\left(\chi_p + \frac{1}{2}\delta_p\right)\left(\chi_d + \frac{1}{2}\delta_d\right).$$

∎

Hence, under the assumptions of this section, the algorithm converges to an optimum in finite number of steps.

5 Numerical Examples

In this section we report on some computational performance of our algorithms. We have implemented the the algorithms in MATLAB for the second order cone programming problem. Below are the results of our test on 1,000 randomly generated problems with known solutions. For the step sizes, we simply choose $\alpha = \min(1, \tau\alpha')$, $\beta = \min(1, \tau\beta')$, $\gamma = \sqrt{\alpha\beta}$, where α' and β' are the maximum step sizes to the boundary of the second-order cone.

We used $\mathbf{x}_i = (2; 1; \mathbf{0})$, $\mathbf{z}_i = (2; -1; \mathbf{0})$, $\mathbf{y} = \mathbf{0}$ as starting point. We picked $\sigma = 0.25$, $\tau = 0.99$, which may not be the best choice of parameters. Our code reduced the l_2 norm of primal infeasibility, l_2 norm of dual infeasibility, and l_1 norm of duality gap to less than $5.0e-12$ for all the problems. The range of every element in our randomly generated problem is $(-0.5, 0.5)$; therefore, we didn't use relative measurement for accuracy. Note that our accuracy requirement is much more stringent than most other algorithms. Below we summarize the results.

bk	dimension of each block	type of each block	m	r_p0	r_d0	it
10	[2,2,2,2,2,2,2,2,2,2]	[b,i,o,b,i,b,o,i,i,b]	12	342.20	45.59	27.07
10	[10,10,10,10,10,10,10,10, 10,10]	[b,o,i,b,b,i,o,b,b, o]	30	299.69	142.30	34.16
10	[3,10,8,9,12,4,6,3,14,8]	[b,i,o,b,i,o,i,i,b,o]	45	539.07	146.97	31.46
10	[20,10,8,9,12,15,6,3,14,8]	[b,i,b,i,i,o,b,i,b,o]	55	861.28	190.32	33.31
10	[20,15,15,15,15,15,15,15, 15,15]	[b,i,b,i,i,o,b,i,b,o]	75	1331.71	269.07	32.16
12	[10,10,10,10,10,10,10,10, 10,10,10,10]	[b,o,i,b,b,i,o,b,b,o, b,i]	50	420.43	197.80	31.96
15	[10,10,10,10,10,10,10,10 10, 10,10,10,10,10,10]	[b,o,i,b,b,i,o,b,b,o, b,o,i,i,o]	70	558.19	262.10	32.46
15	[15,15,15,15,15,15,15,15, 15,15,15,15,15,15,15,]	[i,o,b,i,i,b,o,i,b,b, i,o,b,b,o]	100	1748.47	375.81	33.46
20	[10,20,13,20,24,20,3,8,26, 30,9,12,21,3,11,23,5,2,20, 18]	[b,o,i,b,b,i,o,b,b,o, b,b,i,o,i,b,b,b,i,b]	130	1478.57	496.35	31.97
20	[20,20,20,20,20,20,20,20 20,20,20,20,20,20,20,20, 20,20,20,20]	[b,o,i,b,b,i,o,b,b,o, b,b,i,o,i,b,b,b]	130	1348.60	572.31	33.94

Figure 1: Numerical results for the Q method applied to SOCP problems

- Each row is the summary of 100 instances of a problem with the same number of blocks, dimension of each block, optimum variable type, and number of equality constraints.
- "bk" represents the number of blocks;
- "type of each block" shows at optimum, whether each block is in the boundary(b), zero(o), or in the interior(i);

- "m" is the number of constraints;
- "r_p0" is the average l_2 norm of initial primal infeasibility for the 100 instances;
- "r_d0" is the average l_2 norm of initial dual infeasibility for the 100 instances;
- "it" is the average number of iterations for the 100 instances.

All the instances were terminated at ϵ accuracy within 50 iterations. This shows that our algorithm is indeed stable and can get high accuracy. Notice that the problem type and size have little effect on the total number of iterations, which is expected in well-behaved interior point methods.

References

[1] F. Alizadeh and D. Goldfarb. Second-order cone programming. *Math. Programming Series B*, 2002. To appear.

[2] F. Alizadeh and D. Goldfarb. Second-order cone programming. *Math. Program.*, 95(1, Ser. B):3–51, 2003.

[3] F. Alizadeh, J.P. Haeberly, and M.L. Overton. A new primal-dual interior point method for semidefinite programming. In *Proc. 5th SIAM Conference on Appl. Linear Algebra*, Snowbird, Utah, 1994.

[4] F. Alizadeh and S.H. Schmieta. Symmetric Cones, Potential Reduction Methods and Word-By-Word Extensions. In R. Saigal, L. Vandenberghe, and H. Wolkowicz, editors, *Handbook of Semidefinite Programming, Theory, Algorithms and Applications*, pages 195–233. Kluwer Academic Publishers, 2000.

[5] Farid Alizadeh, Jean-Pierre A. Haeberly, and Michael L. Overton. Primal-dual interior-point methods for semidefinite programming: convergence rates, stability and numerical results. *SIAM J. Optim.*, 8(3):746–768 (electronic), 1998.

[6] Jacques Faraut and Adam Korányi. *Analysis on symmetric cones*. The Clarendon Press Oxford University Press, New York, 1994. Oxford Science Publications.

[7] L. Faybusovich. Euclidean Jordan algebras and interior-point algorithms. *Positivity*, 1(4):331–357, 1997.

[8] Leonid Faybusovich. Linear systems in Jordan algebras and primal-dual interior-point algorithms. *J. Comput. Appl. Math.*, 86(1):149–175, 1997. Special issue dedicated to William B. Gragg (Monterey, CA, 1996).

[9] Roland W. Freund, Florian Jarre, and Shinji Mizuno. Convergence of a class of inexact interior-point algorithms for linear programs. *Math. Oper. Res.*, 24(1):50–71, 1999.

[10] O. Güler. Barrier Functions in Interior Point Methods. *Math. of Oper. Res.*, 21:860–885, 1996.

[11] R. Horn and C. Johnson. *Topics in Matrix Analysis*. Cambridge University Press, Cambridge, 1990.

[12] N. Jacobson. *Structure and Representation of Jordan Algebras*, volume XXXIX of *Colloqium Publications*. American Mathematical Society, Providence, Rhode Island, 1968.

[13] Koecher, Max. *The Minnesota Notes on Jordan Algebras and Their Applications*. Springer-Verlag, 1999. Edited by Kreig, Aloys and Walcher, Sebastian based on Lectures given in The University of Minnesota, 1960.

[14] M. Kojima, S. Mizuno, and A. Yoshise. A primal-dual interior-point algorithm for linear programming. In N. Megiddo, editor, *Progress in Mathematical Programming*, Berlin, 1989. Springer Verlag.

[15] Masakazu Kojima, Nimrod Megiddo, and Shinji Mizuno. A primal-dual infeasible-interior-point algorithm for linear programming. *Math. Programming*, 61(3, Ser. A):263–280, 1993.

[16] N. Megiddo. Pathways to the optimal set in linear programming. In N. Megiddo, editor, *Progress in Mathematical Programming*. Springer Verlag, Berlin, New York, 1989.

[17] R.D.C. Monteiro. Primal-dual path-following algorithms for semidefinite programming. *SIAM J. Optim.*, 7:663–678, 1997.

[18] R.D.C. Monteiro and I. Adler. Interior path following primal-dual algorithms. Part I: Linear programming. *Math. Programming*, 44:27–41, 1989.

[19] S.H. Schmieta. *Application of Jordan algebras to the design and analysis of interior-point algorithms for linear, quadratically constrained quadratic, and semi-definite programming.* PhD thesis, Rutgers Center for Operations Research, Rutgers University, 1999.

[20] S.H. Schmieta and F. Alizadeh. Associative and Jordan Algebras, and Polynomial Time Interior-Point Algorithms for Symmetric Cones. *Math. of Oper. Res.*, 26(3):543–564, 2001.

[21] S.H. Schmieta and F. Alizadeh. Extension of Commutative Class of Primal-Dual Interior Point Algorithms to Symmetric Cones. *Math. Programming*, 96:409–438, 2003.

[22] T. Tsuchiya. A convergence analysis of the scaling-invariant primal-dual path-following algorithms for second-order cone programming. *Optim. Methods Softw.*, 11:141–182, 1999.

[23] Y. Zhang. On extending primal-dual interior-point algorithms from linear programming to semidefinite programming. *SIAM J. Optim.*, 8:356–386, 1998.