Fall 2013

Full Newton Step Interior Point Method for Linear Complementarity Problem Over Symmetric Cones

Andrii Berdnikov
Georgia Southern University

FULL NEWTON STEP INTERIOR POINT METHOD FOR LINEAR COMPLEMENTARITY PROBLEM OVER SYMMETRIC CONES

by

ANDRII BERDNIKOV

(Under the Direction of Goran Lesaja)

ABSTRACT

In this thesis, we present a new Feasible Interior-Point Method (IPM) for Linear Complementarity Problem (LPC) over Symmetric Cones. The advantage of this method lies in that it uses full Newton-steps, thus, avoiding the calculation of the step size at each iteration. By suitable choice of parameters we prove the global convergence of iterates which always stay in the the central path neighborhood. A global convergence of the method is proved and an upper bound for the number of iterations necessary to find ε-approximate solution of the problem is presented.

INDEX WORDS: Linear complementarity Problem, Interior-Point Method, Jordan Euclidean Algebra, Symmetric Cones

FULL NEWTON STEP INTERIOR POINT METHOD FOR LINEAR COMPLEMENTARITY PROBLEM OVER SYMMETRIC CONES

by

ANDRII BERDNIKOV

B.S., Odessa I.I. Mechnikov National University, Ukraine, 2005

A Thesis Submitted to the Graduate Faculty of Georgia Southern University in

Partial Fulfillment of the Requirement for the Degree

MASTER OF SCIENCE

STATESBORO, GEORGIA

2013

FULL NEWTON STEP INTERIOR POINT METHOD FOR LINEAR COMPLEMENTARITY PROBLEM OVER SYMMETRIC CONES

by

ANDRII BERDNIKOV

Major Professor: Goran Lesaja

Committee: Alina Iacob

Alex Stokolos

Electronic Version Approved:

May, 2013

DEDICATION

To Ralph.

ACKNOWLEDGMENTS

The author owes more than gratitude to Dr. Alex Stokolos and Dr. Goran Lesaja, without whom neither this work nor my study at Georgia Southern University couldn't be feasible.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

LINEAR COMPLEMENTARITY PROBLEM (LCP)

In this chapter we will give the basic definitions, state important results and discuss the motivation of the work.

1.1 Overview

Problems of linear optimization (i.e. when the target function and constraints are linear) appear in all aspects of life, engineering, science and business. By 1950s Linear Optimization has developed into a thriving field of mathematics.

Linear Complementarity problem (LCP) was proposed in 1968 by Cottle and Dantzig (see [3]) and is very closely connected to Linear Optimization problem (LP), in a way it's a natural extension of LP. The Karush-Kuhn-Tucker conditions for many linear and non-linear optimization problem can be reduced to LCP. Some examples will be shown in section 1.3.

The Euclidean Jordan algebras (**EJA**) were introduced by P. Jordan in 1933 as a way of describing some aspects of Quantum Mechanics. Their algebraic structure was studied quite extensively, however, only in recent decades it has been established the connection between EJAs and the most important cases of cones used in optimization, such as non-negative orthant, second order cone, and cone of positive semidefinite symmetric matrices. Symmetric cones are cones of squares of EJAs and serve as a unifying framework for problems of conic optimization and, in a way, classify them. Faraut and Korny in their classical monograph on Jordan algebras, symmetric cones and related topics [4] provide a very sophisticated theoretical apparatus. The first to study conic formulation of convex optimization problems were Nesterov and

Nemirovski(in [14]).

Interior-point methods (IPM) for convex optimization problems in the framework of self-scaled cones, without the discovery of connection to EJA being yet made, were studied by Nesterov and Todd in [16] and [17]. The first connection between the self-scaled cones and symmetric cones, which can be completely classified in the framework of EJA, was made in [7] by Güler. The first analysis of a short-step path-following IPM for SCLO and SCLCP was made in [5, 6] by Faybusovich and, of the first infeasible IPM for SCLCP by Rangarajan in [19], and of primal-dual IPMs for SCLO by Alizadeh and Schmieta in [1], [2]. Analysis of kernel-based IPMs for SCLO was made by Vieira in [24], and of full-Newton step IPMs for SCLO by Gu in [8], in their respective Ph.D. theses. The literature on IPMs for LCPs over the non-negative orthant is quite extensive, however there are very few results on LCPs over the semidefinite cone, and even less so dealing with LCP over general symmetric cones. As a pioneer in analysis of IPMs for Nonlinear Complementarity Problems (NCP) over symmetric cones can be regarded Yoshise (see [25] and [25]).

Since late 1940s, a standard and most popular method of solving linear programming problems was **Simplex Method (SM)**, proposed by George Dantzig. Because of linearity of constraints, the surface(boundary) of the feasible region looks like a connected graph(the whole surface is a polyhidral). The method starts at a vertex of the feasible region and goes from vertex to vertex along the edges decreasing the objective function at every step. The method is finite, and it will either find the solution, or show it's non existence. Theoretically, for a LP problem in n-dimensional space, the algorithm may have to visit each and every vertex and therefore will have to have 2^n iterations(see [11]). In practice, the algorithm is remarkably efficient and usually takes $O(n)$ iterations. Exponential number of iterations are never observed in practice, they appear only in artificially designed examples.

In 1976 Nemirovsky and Yudin [15] (in 1977 they were followed by Shor [23]) introduced another great method for solving of convex optimization problems - the *ellipsoid method*. It works by encapsulating the minimizer of a convex function in a sequence of ellipsoids whose volume decreases at each iteration. In 1984 Khachiyan [10] showed that the ellipsoid method allows to solve the LP in polynomial time. This was the first polynomial time algorithm for the LP. In practice, the method was far surpassed by the SM. Nevertheless, the theoretical importance of the ellipsoid method is hard to neglect.

Another great advancement came in 1984, when Karmarkar [9] introduced his *Interior-Point Method* (IPM) for LP. IPM combines the efficiency of SM with the theoretical advantages of the ellipsoid method and works in polynomial time. Unlike the SM, which travels from vertex to vertex along the edges of the feasible region, the IPM follows approximately a central path in the interior of the feasible region and approaches optimal solution only asymptotically. Therefore the analysis of the IPMs is substantially more complex than that of the SM. There are many efficient IPMs for solving LP based on primal, dual, or primal-dual formulations of the LP. It is established that the primal-dual formulation surpasses both the primal and the dual formulation of the algorithm in efficiency. In this work we shall focus on the *primal-dual* IPM, based on using full Newton step method in a carefully controlled manner. Full Newton step IPMs for LP were first discussed by by Roos [19]. Calculation of step size takes away lot of time in Newton based methods, full Newton step algorithms avoid this by taking a full step at each iteration, which turns it into their advantage.

Theoretical foundations of IPMs for convex optimization problems were laid by Nesterov and Nemirovski [14] in 1994. They and others generalized the IPMs to solve many important optimization problems, such as semidefinite optimization, second order cone optimization, and general convex optimization problems.

1.2 Linear Complementarity Problem

The standard **linear complementarity problem(LCP)** is formulated in a following way. Find $x, s \in \mathbb{R}^n$ satisfying

$$\begin{cases} s = Mx + q, \\ x^T s = 0, \\ x, s \geq 0, \end{cases} \tag{1.1}$$

As we see (1.1) is not an optimization problem per se. However, some important optimization problems such as linear and quadratic optimization problems can be reduced to LCP. The equation $x^T s = 0$ is called a *complementarity condition*, and because both s and x belong to the first(non-negative) orthrant, it means that all $s_i x_i = 0, \;\; i = 1, \ldots, n$. Therefore it is more convenient to rewrite the complementarity condition as follows

$$xs = 0, \tag{1.2}$$

where xs denotes H'Adamard product - a vector such that it's i-th coordinate is $x_i s_i$. This way the problem is a bit more revealing, we know that for $s_i x_i$ to be zero, either x_i or s_i should be equal to zero. This way vectors x and s **complement** each other, where one has a non zero coordinate the other has to have zero.

$$\begin{cases} s = Mx + q, \\ \quad xs = 0, \\ \quad x, s \geq 0 \end{cases} \tag{1.3}$$

The solution of (1.3) is unique, if the matrix M is positive semi-definite, $(\forall x)(x^T Mx \geq 0)$(more formal definition will be given in later chapters). The feasible region of (1.3) is defined as the following set

$$F = \left\{(x, s) \in \mathbb{R}^{2n} : s = Mx + q, x \geq 0, s \geq 0\right\}, \tag{1.4}$$

and the strictly feasible region of the (1.3) is

$$F_0 = \{(x, s) \in F : x > 0, s > 0\}.$$

The solution set of the (1.3) is given by

$$F^* = \left\{(x^*, s^*) \in F : x^{*T} s^* = 0\right\}. \tag{1.5}$$

and, as a subset of the above solution set, a set of strictly complementary solutions, is given by

$$F_s^* = \{(x^*, s^*) \in F^* : x^* + s^* > 0\}. \tag{1.6}$$

We can now say that the main idea of the LCP is to find vectors x, s (a solution of the LCP) that are both feasible and complementary(i.e. $(x, s) \in F^*$). If $q \succ 0$, the LCP is always solvable with the zero vector being a trivial solution.(For both a and b vectors over $\mathbb{R}$, the notation $a \succ b$ simply means that the vector $v = a - b$ belongs to $\mathbb{R}_+$).

1.3 Examples

Importance of LCP in part comes from the fact that Karush Kuhn Tucker optimality conditions of many problems can be reduced to LCP. In this section we will present some examples of it. There are also direct applications of LCP in business, transportation and engineering.

Example 1.3.1. *Linear and Quadratic optimization problems*

Quadratic optimization problem is formulated in following way. Given a symmetric $n \times n$ matrix Q, matrix $A \in \mathbb{R}^{n\times m}$, vectors $b \in \mathbb{R}^m$ and $c \in \mathbb{R}^n$ find a vector $x \in \mathbb{R}^n_+$ satisfying

$$\begin{cases} min\{c^T x + \frac{1}{2}x^T Qx\} \\ \qquad Ax \geq b \\ \qquad x \geq 0 \end{cases} \tag{1.7}$$

A **linear optimization** problem in a standard form is usually formulated in the following way. Given a matrix $A \in \mathbb{R}^{n\times m}$, vectors $b \in \mathbb{R}^m$ and $c \in \mathbb{R}^n$ find a vector $x \in \mathbb{R}^n_+$ satisfying

$$\begin{cases} min\{c^T x\} \\ \quad Ax \geq b \end{cases} \tag{1.8}$$

Note, that when $Q = 0$ (1.7) is the same as (1.8). The Karush-Kuhn-Tucker(**KKT**) conditions for (1.7) are

$$\begin{aligned} u = c + Qx - A^T y = 0 \\ v = -b + Ax \geq 0. \end{aligned} \tag{1.9}$$

If we know denote

$$s = \begin{pmatrix} u \\ v \end{pmatrix}, \; z = \begin{pmatrix} x \\ y \end{pmatrix}, \; q = \begin{pmatrix} c \\ -b \end{pmatrix}, \; M = \begin{pmatrix} Q & -A^T \\ A & 0 \end{pmatrix}. \tag{1.10}$$

we can rewrite (1.9) in a more compact way

$$s = Mz + q, \quad s^T z = 0, \quad s, z \in \mathbb{R}_+^{2n}. \tag{1.11}$$

Example 1.3.2. *Bimatrix games.*

Consider a game with two players, with n and m pure strategical choices respectively.Their losses a reflected in matrix A for the first player and matrix B for the second one, thus $A, B \in R^{n\times m}$. In Game Theory a game is called zero sum game, if $A + B = 0$ and a bimatrix game if $A + B \neq 0$. Each player is assigned a probability vector, x for the first player and y for the second one. Player I chooses to play strategy i with probability x_i and Player II chooses to play strategy j with probability y_j. Therefore in order to comply with requirements of Probability Theory we have to have $\sum x_i = 1$ and $\sum y_j = 1$. Defined this way, the losses of Player I and II, are random variables with expectations $x^T Ay$ for player I and $x^T By$ for player II.

A player is changing his own strategy while the other player holds his strategy fixed to minimize loss. i.e,

$$\begin{aligned} \forall x \geq 0 \; \bar{x}^T A\bar{y} &\leq x^T A\bar{y} \; e_m^T x = 1, \\ \forall y \geq 0 \; \bar{x}^T B\bar{y} &\leq \bar{x}^T By \; e_n^T y = 1, \end{aligned} \tag{1.12}$$

where the e denotes the identity vector. The objective is to find $(\bar{x}, \bar{y})$ that is called Nash equilibrium pair. The problem of finding Nash equilibrium pair, with the help of following lemma, is reduced to a LCP.

Lemma 1.3.3. *Suppose $A, B \in R^{m\times n}$ are positive loss matrices representing a game $\Gamma(A,B)$ and suppose that $(s,t) \in R^{m\times n}$ solves LCP(M,q), where*

$$M = \begin{pmatrix} 0 & A \\ B^T & 0 \end{pmatrix}, q = -e_{m+n} \in R^{m+n}.$$

Then $(\bar{x}, \bar{y})$ such that,

$$\bar{x} = \tfrac{s}{e_m^T s} \text{ and } \bar{y} = \tfrac{t}{e_m^T t},$$

is the Nash equilibrium pair of $\Gamma(A,B)$.

Example 1.3.4. *The Market Equilibrium Problem*

The state of an economy where the supplies of producers and the demands of consumers are balanced at the resulting price level is called **market equilibrium**. The conditions below represent the supply side in mathematical form

Supply conditions:

$$\begin{cases} min\{c^T x\} \\ Ax \geq b \\ Bx \geq r^* \\ x \geq 0, \end{cases} \tag{1.13}$$

where c is the vector of variable cost, x is the vector of production amount. As we see (1.13) is a linear programming model. Econometric models with commodity prices as the primary independent variables generates the market demand function. Basically, we need to find a vector x^* and subsequent vectors p^* and r^* such that technological constraints on production are represented by the first condition in (1.13) and the demand requirement constraints are represented by the second condition in (1.13). Denote the dual vector of market supply prices corresponding to the second constraint in (1.13) by π^*.

Demand conditions:

$$r^* = Q(p^*) = Dp^* + d, \tag{1.14}$$

where $Q(\cdot)$ is the market demand function with p^* and r^* representing the vectors of demand prices and quantities, respectively.

Equilibrium condition:

$$p^* = \pi^* \tag{1.15}$$

The Karush-Kuhn-Tucker conditions for LP problem (1.13) lead us to the following system

$$\begin{cases} y & = c - A^T v - B^T \pi \geq 0, \quad x \geq 0, \quad y^T x = 0, \\ u & = -b + Ax \geq 0, \qquad\qquad v \geq 0, \quad u^T v = 0, \\ \delta & = -r + Bx \geq 0, \qquad\qquad \pi \geq 0, \quad \delta^T \pi = 0. \end{cases} \tag{1.16}$$

The problem (1.13) has an optimal solution vector x^* if and only if there exist vectors v^* and π^* satisfying the conditions of the system (1.16).

If for r^*, we substitute the demand function (1.14) and we use condition (1.15), then we can see that the conditions in (1.16) gives us the linear complementarity problem where

$$q = \begin{pmatrix} c \\ -b \\ -d \end{pmatrix}, \; M = \begin{pmatrix} 0 & -A^T & -B^T \\ A & 0 & 0 \\ B & 0 & -D \end{pmatrix}. \tag{1.17}$$

Observe that the matrix M in (1.17) is bisymmetric if and only if the matrix D is symmetric. The above steps give us a fairly simple method how any linear problem in general, can be expressed in the LCP framework. This can also be extended to quadratic programming problems as stated below.

$$\begin{array}{ll} \text{maximize} & d^T x + \frac{1}{2} x^T D x + b^T y \\ \text{subject to} & A^T y + B^T x \leq c \\ & x \geq 0, \quad y \geq 0. \end{array} \tag{1.18}$$

On the other hand, if D is not symmetric, then M is not bisymmetric and the connection between the market equilibrium model and the quadratic program above fails to exist.

1.4 Generalizations

So far we considered only the case of LCP where $x, s \geq 0$, i.e. when x, s belong to the non-negative orthrant. In this section we introduce generalizations we shall be working with in later chapters. First we observe, that non-negative orthrant is an example of a cone.

A **cone** $\mathscr{K}$ is a set with a following property: $(\forall t \in \mathscr{K})(\forall \alpha > 0)(\alpha x \in \mathscr{K})$.

We already observed that the non-negative orthrant $\mathbb{R}^n_+$ is an example of a cone. Another example of a cone is so called Lorenz cone L^{n+1}_+, which will be described in Example 2.4.1.

Let L^{n+1} be the $(n+1)$-dimensional real vector space whose elements are indexed from zero. Denote $(x_0, x_1, ..., x_n) \in \mathbb{R}^{n+1}$ as $(x_0; \bar{x})$, with $\bar{x} = (x_1, ..., x_n) \in \mathbb{R}^n$. The cone of squares of L^{n+1} is

$$L^{n+1}_+ = \{x \in \mathbb{R}^{n+1} : x_0 \geq \|\bar{x}\|\} \tag{1.19}$$

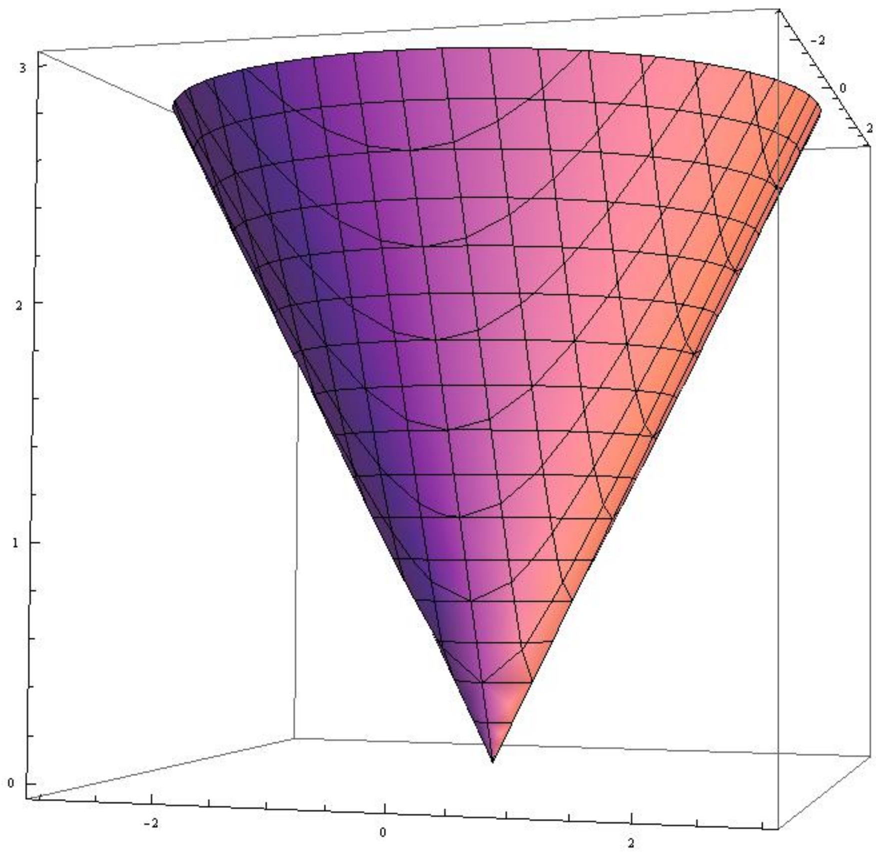

Figure 1.1: Graphical representation of ice-cream cone in $\mathbb{R}^3$

It is also known as ice-cream cone because of it's shape in $\mathbb{R}^2$ and $\mathbb{R}^3$. We include figure 1.1 and figure 1.4 with ice-cream cone in 3-dimensional and 3 dimensional spaces respectively (made in Wolfram Mathematica).

Non-negative orthrant $\mathbb{R}^n_+$ and Lorenz cone are both examples of symmetric cones. We will deal with symmetric cones at length in section (2.3).

Returning back to (1.3) it is a natural way to alter the third condition $x, s \in \mathbb{R}^n_+$ into $x, s \in L^{n+1}_+$, example (2.4.1)

We can substitute non-negative orthrant $\mathbb{R}^n_+$ with a any symmetric cone $\mathscr{K}$, given

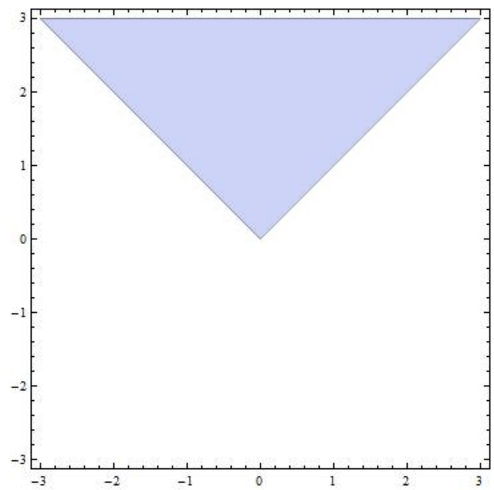

Figure 1.2: Graphical representation of ice-cream cone in $\mathbb{R}^2$

that the operations are also appropriately updated.

$$\begin{cases} s = Mx + q, \\ xs = 0, \\ x, s \in \mathscr{K} \end{cases} \tag{1.20}$$

An example of (1.20) and one way to generalize (1.3) is to change the variables from vectors to matrices. We denote $\mathscr{L}_+^n$ as a cone of positive semi-definite $n \times n$ matrices(see example 2.4.2 for details). Let now $X, S \in \mathscr{L}_+^n$. Dimension of $\mathscr{S}_+^n$ as a vector space is $\bar{n} = \frac{n(n+1)}{2}$(cause the matrices are symmetrical), so it is natural to take $q \in \mathbb{R}^{\bar{n}}$. Inner product is just the trace of a matrix product $\langle X, Y \rangle = tr\,(XY)$. Let's introduce linear operators $P, Q : \mathscr{S}_+^n \longmapsto \mathbb{R}^{\bar{n}}$, and a related to them LCP, which we call semi-definite linear complimentarity problem **(SDLCP)**

$$\begin{cases} P(X) + Q(S) = q, \\ XS = 0, \\ X, S \in \mathscr{L}_+^n \end{cases} \tag{1.21}$$

Therefore, $P(X) = (\langle P_1, X\rangle, \ldots, \langle P_{\bar{n}}, X\rangle)^T$ and $Q(S) = (\langle Q_1, S\rangle, \ldots, \langle Q_{\bar{n}}, S\rangle)^T$, where $P_i, Q_i \in \mathscr{S}_+^n$. We now can rewrite (1.21) by components

$$\begin{cases} tr\,(P_i X) + tr\,(Q_i S) = q_i, i = 1, \ldots, \bar{n} \\ XS = 0, \\ X, S \in \mathscr{S}_+^n \end{cases} \tag{1.22}$$

Because the product of two symmetrical matrices is not necessary symmetrical as a matrix and therefore is not even a binary algebraic operation (BAO) on $\mathscr{L}_+^n$, and in order to be able to use the Newton's method, we have to substitute second equation $XS = 0$ with a symmetrized one. It will now become $X \circ S = 0$, where $X \circ S = \frac{XS+SX}{2}$. This product is a BAO on $\mathscr{L}_+^n$, and a commutative one.

Cones are natural generalization of non-negative orthrant $\mathbb{R}_+^n$. This way we can work with spaces far more advanced than the standard euclidean $\mathbb{R}^n$. The most potent theoretical apparatus to work with symmetric cones lies in Euclidean Jordan algebras. Every symmetric cone is a cone of squares of some(i.e. all his elements are squared elements of) Euclidean Jordan algebra. The formal definitions and most important results shall be presented in the next chapter.

CHAPTER 2

EUCLIDEAN JORDAN ALGEBRAS AND SYMMETRIC CONES

In this chapter we introduce the concepts of Euclidean Jordan Algebras and Symmetric Cones, very briefly go over their properties and state the most important results on their characterization.

2.1 Introducing Jordan Algebras

Let us assume that $\mathscr{J}$ is finite dimensional inner product space over $\mathbb{R}$ - the field of real numbers.

Definition 2.1.1. *Bilinear mapping. A binary algebraic operation* $\circ : \mathscr{J}^2 \longrightarrow \mathscr{J}$ *is called bilinear if the following two properties(which one can call axioms of biliniarity) hold.* $\forall x, y, z \in \mathscr{J}\, and \forall \alpha, \beta \in \mathbb{R}$

- $(\alpha x + \beta y) \circ z = \alpha(x \circ z) + \beta(y \circ z)$;
- $z \circ (\alpha x + \beta y) = \alpha(z \circ x) + \beta(z \circ y)$;

Definition 2.1.2. *A finite-dimensional inner product space* $\mathscr{J}$ *over field* $\mathbb{R}$ *is called a Jordan algebra if*

1. *A bilinear maping* $\circ : \mathscr{J}^2 \longrightarrow \mathscr{J}$ *is defined. (Definition of* $\mathbb{R}$*-algebra).*
2. $x \circ y = y \circ x$. *(Commutativity).*
3. $x \circ (x^2 \circ y) = x^2 \circ (x \circ y)$, *where* $x^2 = x \circ x$. *(Jordan's axiom).*

Definition 2.1.3. *For an element* x *we define a linear map* $L(x) : \mathscr{J} \longrightarrow \mathscr{J}$ *is defined as follows*

$$\forall y \in \mathscr{J}\, (L(x)y = x \circ y)$$

We will call $L(x)$ a linear operator of x(or associated with x). Part 3 in definition 2.1.2 means that the operators $L(x)$ and $L(x^2)$ commute. We define $x^n = x \circ x^{n-1}$ recursively, and call a property $x^n \circ x^m = x^{m+n}$ power associativity. Jordan algebras are not necessary associative, but they are power associative. The proof is a double induction, by n and by m.

Definition 2.1.4. *An element e is said to be identity if*

$$\forall x \in \mathscr{J} (x \circ e = e \circ x = x)$$

Identity is unique and the proof of this fact is trivial(e_1, e_2 are both identities in $\mathscr{J}$, then, by definition, $e_1 = e_1 \circ e_2 = e_2 \circ e_1 = e_2$). From now on we will assume that identity exists in Jordan algebra $\mathscr{J}$ under our consideration. As in a basic Linear Algebra course we will denote $\mathbb{R}[X]$ a ring of polynomials with one variable over the field . For an element x from $\mathscr{J}$ we denote another ring of polynomials

$$\mathbb{R}[x] = \{p(x) | p \in \mathbb{R}[X]\} \tag{2.1}$$

Because $\mathscr{J}$ is a finite dimensional inner product space, each element x has an integer k $(\leq Dim \mathscr{J})$ such that a set of vectors $\{e, x, x^2, x^3, ..., x^k\}$ is linearly dependent. Which means that exist real numbers $a_0, a_1, a_2, ..., a_k$ such that

$$a_k x^k + ... + a_1 x + a_0 e = 0$$

We will call $p(x)$ a minimal polynomial of x, if $p(x) = 0$; plus we will demand for it to be monic and of minimal degree in order to have uniqueness. A minimal positive integer k such that there exist a monic polynomial in $\mathbb{R}[X]$ that annihilates x we will call a degree of element x. As we mentioned above $degx \leq Dim \mathscr{J}$. The minimal polynomial of x is unique, with a trivial proof.

Definition 2.1.5. *Rank of inner product space $\mathscr{J}$. We will call*

$$r = max\left\{degx|x \in \mathscr{J}\right\}$$

the rank of $\mathscr{J}$

An element x is said to be regular if $deg\left\{x\right\} = r$.

Lemma 2.1.6. *The set of all regular elements is open and dense in $\mathscr{J}$. There exist polynomials $a_1, a_2, ..., a_r$ over $\mathscr{J}$ such that the minimal polynomial of every regular element x in $\mathscr{J}$ is given by*

$$f(\lambda; x) = \lambda^r - a_1(x)\lambda^{r-1} + a_2(x)\lambda^{r-2} - ... + (-1)^r a_r(x)$$

The polynomials $a_1, a_2, ..., a_r$ are unique and a_i is homogeneous of degree i.

The polynomial $f(\lambda; x)$ is called **the characteristic polynomial** of the regular element x. Because the regular elements are dense in $\mathscr{J}$, by continuity we can extend the polynomials $a_i(x)$ and, therefore, the characteristic polynomial to all elements of $\mathscr{J}$. Note that the characteristic polynomial is a polynomial of degree r in λ, where r is the rank of $\mathscr{J}$. Moreover, the minimal polynomial coincides with the characteristic polynomial for regular elements, but it divides the characteristic polynomial of non-regular elements. The coefficient $a_1(x)$ is called the trace of x, denoted as $tr(x)$, and the coefficient $a_r(x)$ is called the determinant of x, denoted as $det(x)$. The following proposition gives us an important property of the trace.

Lemma 2.1.7. *Trace is associative as a symmetric bilinear form, i.e.*

$$tr((x \circ y) \circ z) = tr(x \circ (y \circ z)) \tag{2.2}$$

Definition 2.1.8. *An element $x \in \mathscr{J}$ is said to be invertible if*

$$\exists y \in \mathbb{R}[X](x \circ y = e)$$

We will call such y an inverse of x and denote it as x^{-1}. Because $\mathscr{J}$ is power associative, $\mathbb{R}[X]$ is associative and it is a ring, which means x^{-1} is unique. Now let us turn to the linear operator $L(x)$.

Lemma 2.1.9. *If $L(x)$ is invertible, then x is invertible and*

$$x^{-1} = L(x)^{-1}e$$

In the sequel we shall assume $\mathscr{J}$ to be a Jordan algebra over $\mathbb{R}$ with identity element e and of rank r.

Definition 2.1.10. *For $x \in \mathscr{J}$ we define*

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

where $L(x)^2 = L(x)L(x)$. P is called a quadratic operator of x(or quadratic representation of x).

Lemma 2.1.11. *An element $x \in \mathscr{J}$ is invertible if and only if $P(x)$ is invertible, moreover*

$$P(x)x^{-1} = x$$

and

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

Lemma 2.1.12. $\forall x, y \in \mathscr{J}$ *the following holds*

1. *The differential of the map $x \longrightarrow x^{-1}$ is $P(x)^{-1}$.*

2. *If x and y are invertible, then $P(x)y$ is invertible and*

$$(P(x)y)^{-1} = P(x^{-1})y^{-1} \tag{2.3}$$

3. $P(P(y)x) = P(y)P(x)P(y)$.

In particular, the equation 3 of Lemma 2.1.12 is usually called **the fundamental formula**.

2.2 Euclidean Jordan algebras

Let's again denote $\mathscr{J}$ a Jordan algebra with identity over $\mathbb{R}$.

Definition 2.2.1. *Jordan algebra is said to be Euclidean if there exists a positive definite symmetric bilinear form on $\mathscr{J}$ which is (quasi)associative; in other words, there exists an inner product denoted by $\langle \cdot, \cdot \rangle$, such that*

$$\forall x, y, z \in \mathscr{J} \left(\langle x \circ y, z \rangle = \langle x, y \circ z \rangle \right)$$

From now on we will assume $\mathscr{J}$ to be Euclidean Jordan algebra unless otherwise explicitly mentioned. As in a standard Linear Algebra course an element c for which $c^2 = c$ is called idempotent. Elements c_1 and c_2 are orthogonal if $c_1 \circ c_2 = 0$.
Because

$$\langle c_1, c_2 \rangle = \langle c_1^2, c_2 \rangle = \langle c_1, c_1 \circ c_2 \rangle = \langle c_1, 0 \rangle = 0 \tag{2.4}$$

orthogonal(with respect to $\circ$) elements are orthogonal with respect to Euclidean inner product too. An idempotent is called primitive if it cannot be represented as a sum of two orthogonal non-zero idempotents.

Definition 2.2.2. *(Jordan Frame) A set* $\{c_1, \ldots, c_r\}$ *is called a Jordan Frame(or alternatively a complete system of orthogonal primitive idempotents) if each of* c_i *is a primitive idempotent,* $\forall i \neq j (c_i \circ c_j = 0)$ *and*

$$\sum_{i=1}^{r} c_i = e$$

The notion of Jordan frame plays an extremely significant role in building the apparatus of analysis per se on Euclidean Jordan algebras. So far they seem "orphaned" - there are only the most basic algebraic operations(addition, inner product, bilinear form) defined here. Using Jordan frames we will be able to extend the definition of any continuous real valued function of one variable into Euclidean Jordan algebras. We will start with two crucial theorems.

Theorem 2.2.3. *(First spectral theorem). For every element* $x \in \mathcal{J}$ *there exist unique real numbers* $\lambda_1, \lambda_2 \ldots \lambda_k$*, all distinct, and a unique Jordan frame* $\{c_1, c_2, \ldots c_k\}$ *such that*

$$x = \sum_{i=1}^{k} \lambda_i c_i. \tag{2.5}$$

All $c_i \in \mathbb{R}[x]$*.*

The numbers λ_i are called the eigenvalues of x and the equation (2.5) its spectral decomposition.

As promised, given a real valued function $f(\cdot)$, we define

$$f(x) = f(\lambda_1)c_1 + f(\lambda_2)c_2 + \ldots + f(\lambda_k)c_k. \tag{2.6}$$

In light of the goal of this work, we would particularly interested in the inverse(provided that neither of the eigenvalues is zero), square root(only if all the eigenvalues are non negative) and square functions.

$$x^{-1} = \lambda_1^{-1}c_1 + \ldots + \lambda_k^{-1}c_k. \tag{2.7}$$

$$x^{\frac{1}{2}} = \lambda_1^{\frac{1}{2}} c_1 + \ldots + \lambda_k^{\frac{1}{2}} c_k. \tag{2.8}$$

$$x^2 = \lambda_1^2 c_1 + \ldots + \lambda_k^2 c_k. \tag{2.9}$$

What is wonderful about the above definitions, is that they perfectly fit into the "natural" algebraic way of defining inverse, square root and square. I.e. it is easy to verify that

$$x^{-1} \circ x = x \circ x^{-1} = e \tag{2.10}$$

as well as

$$x^2 = x \circ x \tag{2.11}$$

and

$$x^{\frac{1}{2}} \circ x^{\frac{1}{2}} = x \tag{2.12}$$

Theorem 2.2.4. *(Second spectral theorem) Let* r *be the rank of* $\mathscr{J}$*. Given* $x \in \mathscr{J}$ *there exist a Jordan frame* $\{c_1, c_2, \ldots c_r\}$ *and real numbers* $\lambda_1, \lambda_2 \ldots \lambda_r$*, such that*

$$x = \sum_{i=1}^{r} \lambda_i c_i. \tag{2.13}$$

The numbers $\lambda_1, \lambda_2 \ldots \lambda_r$ *as well as their multiplicities are uniquely determined by* x*. Moreover,*

$$a_k(x) = \sum_{1 \le i_1 < \ldots < i_k \le r} \lambda_{i_1}, \lambda_{i_2} \ldots \lambda_{i_k}, \tag{2.14}$$

where a_k *is the polynomial defined in Lemma* 2.1.6.

In particular

$$tr(x) = \sum_{i=1}^{r} \lambda_i, det(x) = \prod_{i=1}^{r} \lambda_i. \tag{2.15}$$

To emphasize their dependence on x, we will write the eigenvalues as vector function $\lambda(x) \in \mathbb{R}^r$ whose i-th component is $\lambda_i(x)$. We will denote $\lambda_{min}(x)$ and $\lambda_{max}(x)$ as the smallest and the largest eigenvalues of x respectively.

Definition 2.2.5. *Two elements x and y are called similar if they share the same set of eigenvalues. We denote the relation of similarity with $x \sim y$.*

Another important theorem follows. It will help us to close the circle around the Euclidean Jordan algebra concept connecting the inner product to the trace.

Theorem 2.2.6. *Let $\mathscr{J}$ be Jordan algebra over $\mathbb{R}$ with identity element e. The following statements are equivalent.*

i $\mathscr{J}$ is an Euclidean Jordan algebra.

ii The symmetric bilinear form $tr(x \circ y)$ is positive definite, i.e. $\forall x \in \mathscr{J}$

$$tr(x \circ x) > 0 \tag{2.16}$$

From now on we will denote $\langle x, y\rangle = tr(x \circ y)$ and call it the trace inner product. We will also define the Frobenius norm associated with this inner product

$$\|x\|_F = \sqrt{\langle x, x\rangle} = \sqrt{tr(x^2)} = \sqrt{\sum_{i=1}^{r} \lambda_i^2} = \|\lambda(x)\|. \tag{2.17}$$

The sequence of equations (2.17) is significant because it shows that the Frobenius norm of element x equals standard Euclidean norm of its spectral vector $\lambda(x)$.

2.3 Symmetric Cones

In this section we will introduce the notion of Symmetric cones, state the most important results concerning them and discuss some of their properties, all in the light of the goal of this work. But first we need some definitions.

A set $\mathscr{K}$ is called **convex** if $\forall x, y \in \mathscr{K}$ and for any $\alpha \in [0, 1]$, we have $\alpha x + (1-\alpha)y \in$

$\mathscr{K}$. A set $\mathscr{K}$ is called a **cone** if $\forall x \in \mathscr{K} \forall \alpha \geq 0 (\alpha x \in \mathscr{K})$. Naturally, we will call a set $\mathscr{K}$ a **convex cone** if it is both convex and a cone, which in practice means that for all $x, y \in \mathscr{K}$ and $\alpha, \beta \geq 0$ we have $\alpha x + \beta y \in \mathscr{K}$. An interior set of the cone $\mathscr{K}$ is defined containing all the points that belong to $\mathscr{K}$ with some neighborhood
A set $K^* = \{y \in \mathscr{J} | < x, y > \geq 0, forall x \in \mathscr{K}\}$ is called a **dual cone** of $\mathscr{K}$(provided that $\mathscr{K}$ itself is a cone). $\mathscr{K}^*$ is always a convex cone, even when the original one $\mathscr{K}$ is not. If a cone coincides with its dual, we call it **self-dual**. Self-dual cones have non empty interior. Convex cone $\mathscr{K}$ is called **homogeneous** if for any $x, y \in int\mathscr{K}$ there exist an invertible linear operator such that $g\mathscr{K} = \mathscr{K}$ and $gx = y$.

Definition 2.3.1. *(Symmetric cone). A cone $\mathscr{K}$ is called symmetric if it's self-dual and homogeneous.*

Let's denote the set of all invertible linear maps from $\mathscr{J}$ into itself by $GL(\mathscr{J})$. A linear map g is orthogonal if $g^* = g^{-1}$, where g^* denotes a conjugate map. We shall cal a map $g \in GL(\mathscr{J})$ an **automorphism** of $\mathscr{J}$ if for every x and y in $\mathscr{J}$, we have $g(x \circ y) = g(x) \circ g(y)$, which is equivalent to $gL(x)g^{-1} = L(gx)$. The set of all automorphisms of $\mathscr{J}$ is denoted as $Aut(\mathscr{J})$. A map $g \in GL(\mathscr{J})$ shall be called an **automorphism** of $\mathscr{K}$ if $g\mathscr{K} = \mathscr{K}$. The set of all automorphisms of $\mathscr{K}$ is denoted as $Aut(\mathscr{K})$. We will denote a set of all orthogonal isomorphisms on the cone as

$$OAut(\mathscr{K}) = \{g \in Aut(\mathscr{K}) : g^* = g^{-1}\}. \tag{2.18}$$

The following facst are of importance to us.

$$Aut(\mathscr{J}) = OAut(\mathscr{K}). \tag{2.19}$$

Lemma 2.3.2. *The trace and the determinant are invariant under $Aut(\mathscr{J})$.*

We now introduce a relation of partial ordering between the elements of symmetric cones.

Definition 2.3.3. *If $x - y \in \mathscr{K}$ then we define $x \succeq_{\mathscr{K}} y$. If $x - y \in int\mathscr{K}$ then we define $x \succ_{\mathscr{K}} y$.*

Definition 2.3.4. *Let $\mathscr{J}$ be Euclidean Jordan algebra, it's* ***cone of squares*** *is defined as a set*

$$\mathscr{K}(\mathscr{J}) = \{x^2 | x \in \mathscr{J}\} \tag{2.20}$$

Cone of squares is the first link between cones and the realm of Euclidean Jordan algebras. The following theorem specifies this connection.

Theorem 2.3.5. *Let $\mathscr{J}$ be Euclidean Jordan algebra, then $\mathscr{K}(\mathscr{J})$ is a symmetric cone, and is the set of elements $x \in \mathscr{J}$ for which $L(x)$ is positive semi definite. Moreover, if x is invertible, then*

$$P(x) int\mathscr{K}(\mathscr{J}) = int\mathscr{K}(\mathscr{J}) \tag{2.21}$$

Theorem 2.3.6. *A cone is symmetric if and only if it is a cone of squares of some Euclidean Jordan algebra.*

The above crucial and beautiful results show one to one correspondence between the classes of Euclidean Jordan algebras and Symmetric cones. And now we can use the apparatus of the former to study the latter. From now on, unless otherwise specified, we shall have $\mathscr{J}$ as Euclidean Jordan algebra and assume $\mathscr{K}$ to be its cone of squares $\mathscr{K}(\mathscr{J})$.

Another important results follows.

Theorem 2.3.7. *If $x, s \in \mathscr{K}$, then $tr(x \circ s) \geq 0$. Moreover,*

$$(tr(x \circ s) = 0) \Longleftrightarrow (x \circ s = 0) \tag{2.22}$$

We will call an $\mathbb{R}$-algebra $\mathscr{J}$ simple if it contains only trivial ideals- 0 and itself.

Theorem 2.3.8. *If $\mathscr{J}$ is an Euclidean Jordan algebra, then it is in a unique way a direct sum of simple Euclidean Jordan algebras.*

Definition 2.3.9. *A symmetric cone $\mathscr{K}$ in a Euclidean space $\mathscr{J}$ is called irreducible(simple), if there doesn't exist non trivial subspaces $\mathscr{J}_1$, $mathbbmathscr J_2 \subset \mathscr{J}$ and symmetric cones $\mathscr{K}_1 \subset \mathscr{J}_1$,$\mathscr{K}_2 \subset \mathscr{J}_2$, such that*

$$\begin{aligned} \mathscr{J} &= \mathscr{J}_1 \bigoplus \mathscr{J}_2 \\ \mathscr{K} &= \mathscr{K}_1 \bigoplus \mathscr{K}_2, \end{aligned} \tag{2.23}$$

where $\bigoplus$ denotes direct sum.

Theorem 2.3.10. *If $\mathscr{K}$ is a symmetric cone, then it is in a unique way a direct sum of irreducible symmetric cones.*

Theorem 2.3.11. *If $\mathscr{J}$ is a simple Jordan algebra, then it is isomorphic to one of the following algebras.*

1. *The algebra in space $\mathbb{R}^{n+1}$ with Jordan multiplication defined as*

$$x \circ y = (x^T y; x_0 \bar{y} + y_0 \bar{x}), \tag{2.24}$$

   where $x = (x_0; \bar{x})$ and $y = (y_0; \bar{y})$ with $x_0, y_0 \in \mathbb{R}$ and $\bar{x}, \bar{y} \in \mathbb{R}^n$.

2. *The algebra of real symmetric matrices with Jordan multiplication defined as*

$$X \circ Y = \frac{XY + YX}{2}. \tag{2.25}$$

3. *The algebra of complex Hermitian matrices with Jordan multiplication defined as in (2.25).*

4. *The algebra of quaternion Hermitian matrices, with Jordan multiplication defined as in (2.25).*

5. *The algebra of 33 octonion Hermitian matrices with Jordan multiplication defined as in (2.25).*

Above theorem holds a tremendous significance. Any Jordan algebra can be reduced to(or decomposed into) the combination of the cases enumerated in theorem (2.25). We will continue with another way of decomposing Euclidean Jordan algebras.

Theorem 2.3.12. *(First Peirce decomposition theorem). If $\mathscr{J}$ is a Euclidean Jordan algebra and c an idempotent, then $\mathscr{J}$ as a vector space can be decomposed in a direct sum of*

$$\mathscr{J} = \mathscr{J}_0 \bigoplus \mathscr{J}_{\frac{1}{2}} \bigoplus \mathscr{J}_1, \tag{2.26}$$

where

$$\mathscr{J}_i(c) = \{x : c \circ x = ix\}. \tag{2.27}$$

The subspaces $\mathscr{J}_0$ and $\mathscr{J}_1$ are subalgebras of $\mathscr{J}$, which are orthogonal in terms

$$\mathscr{J}_0 \circ \mathscr{J}_1 = \{0\}. \tag{2.28}$$

Moreover,

$$(\mathscr{J}_0(c) \bigoplus \mathscr{J}_1(c)) \circ \mathscr{J}_{\frac{1}{2}}(c) \subseteq \mathscr{J}_{\frac{1}{2}}(c), \tag{2.29}$$

$$\mathscr{J}_{\frac{1}{2}}(c) \circ \mathscr{J}_{\frac{1}{2}}(c) \subseteq \mathscr{J}_0(c) \bigoplus \mathscr{J}_1(c). \tag{2.30}$$

We shall have even more sophisticated decomposition.

Theorem 2.3.13. *(Second Peirce decomposition theorem). Let $\mathscr{J}$ be a Euclidean Jordan algebra and $\{c_1, \ldots, c_r\}$ a Jordan frame. Then*

1. $\mathscr{J}$ *as a vector space can be decomposed in a direct sum of*

$$\mathscr{J} = \bigoplus_{i \leq j} \mathscr{J}_{ij}, \tag{2.31}$$

where

$$\mathscr{J}_{ii} = \mathscr{J}_i(c_i) = \mathbb{R}c_i, \tag{2.32}$$

$$\mathscr{J}_{ij} = \mathscr{J}_{\frac{1}{2}}(c_i) \cap \mathscr{J}_{\frac{1}{2}}(c_j), i \neq j. \tag{2.33}$$

2. *If we denote by* $\mathscr{P}_{ij}$ *the orthogonal projection on* $\mathscr{J}_{ij}$*, then*

$$\mathscr{P}_{ii} = \mathscr{P}(c_i), \tag{2.34}$$

$$\mathscr{P}_{ij} = 4L(c_i)L(c_j), i \neq j. \tag{2.35}$$

3.

$$\mathscr{J}_{ij} \circ \mathscr{J}_{ij} \subseteq \mathscr{J}_{ii} + \mathscr{J}_{jj}, \tag{2.36}$$

$$\mathscr{J}_{ij} \circ \mathscr{J}_{jk} \subseteq \mathscr{J}_{ik}, i \neq k, \tag{2.37}$$

$$\mathscr{J}_{ij} \circ \mathscr{J}_{kl} \subseteq \{0\}, if \{i,j\} \cap \{k,l\} = \emptyset. \tag{2.38}$$

Corollary 2.3.14. *Let* $x \in \mathscr{J}$ *and* $x = \sum_{i=1}^r \lambda_i c_i$ *be its spectral decomposition. Then the following statements hold*

1. *Matrices* $L(x)$ *and* $P(x)$ *commute and thus share a common system of eigenvectors; moreover,* c_i *are among their common eigenvectors.*

2. *The eigenvalues of* $L(x)$ *are of the form*

$$\frac{\lambda_i + \lambda_j}{2}, 1 \leq i \leq j \leq r. \tag{2.39}$$

3. *The eigenvalues of* $P(x)$ *are of the form*

$$\lambda_i \lambda_j, 1 \leq i \leq j \leq r. \tag{2.40}$$

We end this section with a very important lemma. We will use it in the next chapter in order to have unique solution of the Newton method system of equation.

Lemma 2.3.15. *Given $x, s \in int\mathscr{K}$, there exists a unique $w \in int\mathscr{K}$ such that*

$$x = P(w)s.$$

Moreover,

$$w = P(x)^{\frac{1}{2}}\left(P(x)^{\frac{1}{2}}s\right)^{-\frac{1}{2}} = \left[P(s)^{-\frac{1}{2}}\left(P(s)^{\frac{1}{2}}x\right)^{\frac{1}{2}}\right]. \tag{2.41}$$

We will call w defined as in (2.41) the scaling point of x and s, in this order.

2.4 Examples

In this section we will present some examples of Euclidean Jordan algebras and symmetric cones.

Example 2.4.1. *The quadratic terms algebra L^{n+1}.*

This algebra has several names(quadratic terms algebra, Jordan spin algebra among them) and is used in relativistic mechanic, where the Jordan algebras were introduced in 1930-s. In optimization it is valued because of a second order cone associated with it. Let L^{n+1} be the $(n+1)$-dimensional real vector space whose elements are indexed from zero. Denote $(x_0, x_1, ..., x_n) \in \mathbb{R}^{n+1}$ as $(x_0; \bar{x})$, with $\bar{x} = (x_1, ..., x_n) \in \mathbb{R}^n$. Define the vector product in this space the following way

$$x \circ y = (x^T y; x_0\bar{y} + y_0\bar{x}). \tag{2.42}$$

The pair $(L^{n+1}, \circ)$ is a Jordan algebra - this fact has a trivial proof - with an identity element $e = (1, 0, ..., 0)$. Every vector x in this algebra satisfies the equation

$$x^2 - 2x_0x + (x_0^2 - \|\bar{x}\|^2)e = 0 \tag{2.43}$$

Therefore, the rank of Jordan spin algebra L^{n+1} is 2, regardless of the dimension of the real vector space it is based on. Each element has two eigenvalues $x_0 \pm \|\bar{x}\|$, and therefore, $tr(x) = 2x_0$ and $det(x) = x_0^2 - \|\bar{x}\|^2$. Linear and Quadratic operators have the following matrices

$$L(x) = \begin{pmatrix} x_0 & \bar{x}^T \\ \bar{x} & x_0 I \end{pmatrix}, \tag{2.44}$$

$$Q(x) = 2L(x)^2 - L(x^2) = \begin{pmatrix} x^T x & 2x_0\bar{x}^T \\ 2x_0\bar{x} & det(x)I + 2\bar{x}\bar{x}^T \end{pmatrix}. \tag{2.45}$$

The spectral decomposition of it's elements is

$$x = \lambda_1 c_1 + \lambda_2 c_2, \tag{2.46}$$

where, $\lambda_1 = x_0 - \|\bar{x}\|$, $\lambda_2 = x_0 + \|\bar{x}\|$ and the idempotents are

$$c_1 = \frac{1}{2}\begin{bmatrix} 1 \\ -\frac{\bar{x}}{\|x\|} \end{bmatrix}, c_2 = \frac{1}{2}\begin{bmatrix} 1 \\ \frac{\bar{x}}{\|x\|} \end{bmatrix}. \tag{2.47}$$

The associative trace inner product, is

$$\langle x, y \rangle = tr(x \circ y) = 2x^T y. \tag{2.48}$$

The cone of squares of L^{n+1} is

$$L_+^{n+1} = \{x \in \mathbb{R}^{n+1} : x_0 \geq \|\bar{x}\|\}. \tag{2.49}$$

It is known as Lorenz cone, the ice-cream(because of it's shape in $\mathbb{R}^2 and \mathbb{R}^3$) cone and by other names.

Example 2.4.2. *Jordan algebra S^n of symmetric matrices and positive semi-definite cone L_+^n.*

Denote S^n the matrix space of n-dimensional symmetric real valued matrices. We will define a symmetric multiplication by

$$X \circ Y = \frac{XY + YX}{2}. \tag{2.50}$$

It is rather easy to show that both commutativity and Jordan's axiom are satisfied and, therefore, $(S^n, \circ)$ is a Jordan algebra. More important, $X \circ X = XX$, therefore characteristic and minimal polynomials, eigenvalues will not change due to symmetrization of the matrix product. As we are dealing with symmetric(therefore diagonalizible with real eigenvalues) matrices, the $degX$ is the number of distinct eigenvalues of X and so it is at most n. Thus, $rankS^n = n$.

As we know, symmetric matrices are diagonalizible, and have only real eigenvalues. Therefore, its cone of squares L^n_+ combines those of symmetric matrices that have non negative eigenvalues, i.e. positive semidefinite. Let $A = Q\Lambda Q^T$, where Λ is a diagonal matrix with eigenvalues, while Q is unitary(orthogonal). The column vectors q_i of Q are mutually orthogonal. We can write

$$A = \sum_{i=1}^{n} \lambda_i q_i q_i^T, \tag{2.51}$$

where λ_i is i-th eigenvalue of A. It is easy to verify that $q_i q_i^T$ form a Jordan frame, using the properties of orthogonal(unitary) matrices.

CHAPTER 3

FEASIBLE INTERIOR POINT METHOD

3.1 System

Let $\mathscr{J}$ be a Euclidean Jordan Algebra with rank r, and $\mathscr{K}$ be its respective cone of squares. As we already know, given a square matrix M and a column vector q a Linear Complementary Problem is formulated as a system

$$\begin{cases} s = Mx + q \\ x \circ s = 0 \\ x, s \in \mathscr{K}. \end{cases} \tag{3.1}$$

As was discussed in the section 1.4, the change from (1.3) is the third line, in addition, the appropriate operations have to be taken into account. Instead of positive orthrant we require for x and s to belong to a symmetric cone.

Before setting a Newton system, we need to parametrize (3.1). Newton method is known to have stability problems when (at least)one of the coordinates is close (or equal) to zero. Therefore we change the second line in the system by parametrizing it, in stead of zeros we will have $x \circ s = \mu e$, where μ is positive scalar, and e is a unit vector. Now we also require strict feasibility $x, s \in int\mathscr{K}$. The system now looks the following way:

$$\begin{cases} s = Mx + q \\ x \circ s = \mu e \\ x, s \in int\mathscr{K}. \end{cases} \tag{3.2}$$

For each $\mu > 0$ the system (3.2) has a unique solution $(x(\mu), s(\mu))$, when M is positive semi-definite, and we call the pair $(x(\mu), s(\mu))$ the μ-center of the problem (3.1). The set of μ-centers (with μ running through all positive real numbers) shall be known as the **Central path** of (3.1). If $\mu \longrightarrow 0$, then the limit of the central path exists and since the limit points satisfy the complementarity condition, the limit yields optimal solution for the problem (3.1) [6]. The main idea of the method is to approximately trace a central path while reducing μ to 0.

Now everything is ready to set up the Newton system. First of all, let's rewrite the above system as follows

$$\begin{cases} s - Mx + q = 0 \\ x \circ s - \mu e = 0. \end{cases} \tag{3.3}$$

Denote

$$z = \begin{pmatrix} x \\ s \end{pmatrix}$$

and

$$F(z) = \begin{pmatrix} s - Mx - q \\ x \circ s - \mu e \end{pmatrix}. \tag{3.4}$$

The gradient of F is, thus,

$$\nabla F(z) = \begin{pmatrix} -M & I \\ S & X \end{pmatrix}$$

.

The Newtonian system, therefore, is

$$\begin{pmatrix} -M & I \\ S & X \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta s \end{pmatrix} = -F(z) \tag{3.5}$$

or equivalently

$$\begin{cases} -M\Delta x + \Delta s & = s - Mx - q = 0 \\ s\Delta x + x\Delta s & = \mu e - x \circ s. \end{cases} \tag{3.6}$$

Because $L(x)$ and $L(s)$ do not commute in general, - this in it's turn is an implication of the absence of associativity in Jordan algebras - the above system doesn't necessarily poses a unique solution. We have to scale it in the same way as in [8]. As $x, s \in int\mathscr{K}$, from Lemma 2.1.12 we have

$$x \circ s = \mu e$$

if and only if

$$x = \mu s^{-1}.$$

Let's act with $P(u)$, where $u \in int\mathscr{K}$, on the last identity, we obtain

$$P(u)x = \mu P(u)s^{-1} = \mu(P(u^{-1})s)^{-1},$$

which is equivalent (just multiply by $P(u^{-1})s$ from the right) to

$$P(u)x \circ P(u^{-1})s = \mu e.$$

We will replace second equation in parametrized system (3.4) with the one above. This scaling scheme depends on the choice of u. Our new Newtonian system looks as follows

$$\begin{cases} s = Mx + q \\ P(u)x \circ P(u^{-1})s = \mu e \end{cases}. \tag{3.7}$$

Again denote

$$z = \begin{pmatrix} x \\ s \end{pmatrix}$$

and

$$F(z) = \begin{pmatrix} s - Mx - q \\ P(u)x \circ P(u^{-1})s - \mu e \end{pmatrix}.$$

The Newtonian system is

$$\begin{pmatrix} -M & I \\ P(u^{-1})s \circ P(u) & P(u)X \circ P(u^{-1}) \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta s \end{pmatrix} = -F(z) \tag{3.8}$$

or equivalently

$$\begin{cases} -M\Delta x + \Delta s & = 0 \\ P(u^{-1})s \circ P(u)\Delta x + P(u)x \circ P(u^{-1})\Delta s & = \mu e - P(u)x \circ P(u^{-1})s. \end{cases} \tag{3.9}$$

3.2 NT-scaling

There are different choices of u. We will choose Nesterov-Todd scaling scheme(NT-scaling scheme) and the resulting directions are called Nesterov-Todd directions(NT-directions). This scaling scheme was first proposed by Nesterov and Todd for self-scaling cones in [16, 17] and then applied by Faybusovich in [5, 6] for symmetric cones. According to NT-scaling scheme $u = w^{-\frac{1}{2}}$, where w is a scaling point of x and s. From Lemma 2.3.15 we know that

$$w = P(x)^{\frac{1}{2}} \left(P(x)^{\frac{1}{2}} s\right)^{-\frac{1}{2}} = \left[P(s)^{-\frac{1}{2}} \left(P(s)^{\frac{1}{2}} x\right)^{\frac{1}{2}}\right]. \tag{3.10}$$

Define a variation vector v as follows

$$v = \frac{P(w)^{-\frac{1}{2}} x}{\sqrt{\mu}} = \left[\frac{P(w)^{\frac{1}{2}} s}{\sqrt{\mu}}\right] \tag{3.11}$$

We also introduce slightly altered direction vectors, which would be used purely for analysis of the algorithm.

$$d_x = \frac{P(w)^{-\frac{1}{2}}\Delta x}{\sqrt{\mu}}\ ,\ d_s = \frac{P(w)^{\frac{1}{2}}\Delta s}{\sqrt{\mu}} \tag{3.12}$$

Thus

$$\Delta x = \sqrt{\mu}P(w)^{\frac{1}{2}}d_x\ ,\ \Delta s = \sqrt{\mu}P(w)^{-\frac{1}{2}}d_s \tag{3.13}$$

And the system (3.9) transforms into

$$\begin{cases} \sqrt{\mu}P(w)^{-\frac{1}{2}}d_s - M\sqrt{\mu}P(w)^{\frac{1}{2}}d_x & = 0 \\ \sqrt{\mu}P(w)^{\frac{1}{2}}s \circ P(w)^{-\frac{1}{2}}P(w)^{\frac{1}{2}}d_x & + \\ +\sqrt{\mu}P(w)^{-\frac{1}{2}}x \circ P(w)^{\frac{1}{2}}P(w)^{-\frac{1}{2}}d_s & = \mu e - P(w)^{-\frac{1}{2}}x \circ P(w)^{\frac{1}{2}}s. \end{cases} \tag{3.14}$$

And after we simplify it

$$\begin{cases} \sqrt{\mu}P(w)^{-\frac{1}{2}}d_s - M\sqrt{\mu}P(w)^{\frac{1}{2}}d_x & = 0 \\ \sqrt{\mu}P(w)^{\frac{1}{2}}S \circ d_x + \sqrt{\mu}P(w)^{-\frac{1}{2}}x \circ d_s & = \mu e - P(w)^{-\frac{1}{2}}x \circ P(w)^{\frac{1}{2}}s. \end{cases} \tag{3.15}$$

Multiply first equation by $\frac{P(w)^{\frac{1}{2}}}{\sqrt{\mu}}$ and divide the second by by μ to obtain

$$\begin{cases} d_s - P(w)^{\frac{1}{2}}MP(w)^{\frac{1}{2}}d_x & = 0 \\ \frac{P(w)^{\frac{1}{2}}s\circ d_x}{\sqrt{\mu}} + \frac{P(w)^{-\frac{1}{2}}x\circ d_s}{\sqrt{\mu}} & = e - v^2. \end{cases} \tag{3.16}$$

If we now denote $\bar{M} = P(w)^{\frac{1}{2}} M P(w)^{\frac{1}{2}}$ and recall the definition of the variation vector, we have

$$\begin{cases} d_s - \bar{M} d_x & = 0 \\ v \circ d_x + v \circ d_s & = e - v^2. \end{cases} \tag{3.17}$$

Dividing the second equation by v we'll come to the system

$$\begin{cases} d_s - \bar{M} d_x & = 0 \\ d_x + d_s & = v^{-1} - v. \end{cases} \tag{3.18}$$

3.3 The Algorithm

Previous discussion can be summarized in the following Algorithm.

Algorithm: Feasible Full Newton-step Interior Point Algorithm for LCP over Symmetric Cones

Start

Input:

An accuracy parameter $\varepsilon > 0$;

set the threshold parameter $\tau = \frac{1}{\sqrt{2}}$;

set barrier update parameter $\theta = \frac{1}{r\sqrt{2}}$;

a starting point $(x_0, s_0) \in \mathscr{K}$, $\mu_0 > 0$

such that $s_0 = \bar{M}x_0 + q_0$ and $\delta(x_0, s_0, \mu_0) \leq \tau$.

begin

$x := x_0$; $s := s_0$; $\mu := \mu_0$;

while $(\mu \geq \varepsilon)$ or $(\|x \circ s\|_F \geq \varepsilon)$ **do**.

begin

Calculate w using formula (3.10);

update the the barrier parameter $\mu := (1-\theta)\mu$;

calculate search directions Δx,Δs by solving

the system (3.9) for $u = w^{-\frac{1}{2}}$;

update $x := x + \Delta x$ and $s := s + \Delta s$.

end

end

Output: ε-approximate solution (x, s)

Finish

Figure 3.1: Feasible Full Newton-step Interior Point Algorithm for LCP over Symmetric Cones

CHAPTER 4

ANALYSIS OF THE ALGORITHM

4.1 Feasibility of iterates

In this and following section we will present and discuss the Feasible IPM for LCP over symmetric cones. First we state two lemmas which will be of utmost importance in the analysis of the algorithm. We will start with a lemma estimating the range of eigenvalues of a $\circ$ product.

Lemma 4.1.1. ***(Analog of Lemma 4.49 in*** [**8**]***)*** *Let J be Euclidean-Jordan algebra, $x, s \in J$, and $\langle x, s\rangle \geq 0$, then*

$$-\frac{1}{4}\|x+s\|_F^2 e \preceq_K x \circ s \preceq_K \frac{1}{4}\|x+s\|_F^2 e \tag{4.1}$$

Proof. We write

$$x \circ s = \frac{1}{4}((x+s)^2 - (x-s)^2).$$

Since $(x+s)^2 \in K$, then

$$x \circ s + \frac{1}{4}(x-s)^2 \in K.$$

We know that

$$(x-s)^2 \preceq_K \lambda_{max}\{(x-s)^2\}e \preceq_K \|x-s\|_F^2 e. \tag{4.2}$$

Hence,

$$x \circ s + \frac{1}{4}\|x-s\|_F^2 e \in \mathscr{K} \tag{4.3}$$

and

$$-\frac{1}{4}\|x-s\|_F^2 e \preceq_K x \circ s. \tag{4.4}$$

Using the same argument we have

$$x \circ s = \frac{1}{4}((x+s)^2 - (x-s)^2) \preceq_K \frac{1}{4}(x+s)^2, \tag{4.5}$$

and

$$(x+s)^2 \preceq_K \lambda_{max}\{(x+s)^2\}e \preceq_K \|x+s\|_F^2 e. \tag{4.6}$$

Hence,

$$x \circ s \preceq_K \frac{1}{4}\|x+s\|_F^2 e. \tag{4.7}$$

We, therefore, have

$$-\frac{1}{4}\|x-s\|_F^2 e \preceq_K x \circ s \preceq_K \frac{1}{4}\|x+s\|_F^2 e. \tag{4.8}$$

Notice that since $\langle x, s\rangle \geq 0$ we have

$$\begin{aligned} \|x+s\|_F^2 = \langle x+s, x+s\rangle = \\ = \langle x,x\rangle + \langle s,s\rangle + 2\langle x,s\rangle \geq \\ \geq \langle x,x\rangle + \langle s,s\rangle \geq \\ \geq \langle x,x\rangle + \langle s,s\rangle - 2\langle x,s\rangle = \|x-s\|_F^2, \end{aligned} \tag{4.9}$$

which implies that

$$\|x+s\|_F^2 e \succeq_K \|x-s\|_F^2 e \tag{4.10}$$

and

$$-\frac{1}{4}\|x+s\|_F^2 e \preceq_K x \circ s \preceq_K \frac{1}{4}\|x+s\|_F^2 e \tag{4.11}$$

Q.E.D. □

The following lemma connects the norm of a product with norm of a sum.

Lemma 4.1.2. *Let J be Euclidean-Jordan algebra, $x, s \in J$, and $\langle x, s\rangle \geq 0$. Then*

$$\|x \circ s\|_F \leq \frac{1}{2\sqrt{2}}\|x+s\|_F^2. \tag{4.12}$$

Proof. As in the previous lemma

$$\|x \circ s\|_F^2 = \|\frac{1}{4}((x+s)^2 - (x-s)^2)\|_F^2 = \frac{1}{16} tr[((x+s)^2 - (x-s)^2)^2] =$$

$$= \frac{1}{16}[tr((x+s)^4) + tr((x-s)^4) - 2tr((x+s)^4)] \leq$$

$$\leq \frac{1}{16}[tr((x+s)^4) + tr((x-s)^4)] =$$

$$= \frac{1}{16}\|(x+s)^2\|_F^2 + \|(x-s)^2\|_F^2 \leq$$

$$= \frac{1}{8}\|(x+s)^2\|_F^2.$$

Thus,

$$\|x \circ s\|_F \leq \frac{1}{2\sqrt{2}}\|x+s\|_F^2 \tag{4.13}$$

Q.E.D. □

We shall call x, s such that $x \circ s = \mu e$, the $\mu-$centers and denote them as $x(\mu), s(\mu)$. Looking at the second equation of the system (3.18) it is quite natural to take

$$\delta(x, s, \mu) = \delta(v) = \frac{1}{2}\|v - v^{-1}\|_F \tag{4.14}$$

as a measure of closeness of our solution x, s to the $\mu-$centers $x(\mu), s(\mu)$. One can show that $v = e$ is equivalent to x, s being the $\mu-$centers.

We will use Frobenius norm as defined in Chapter 2.

$$\begin{aligned} 4\delta^2(v) &= \|v - v^{-1}\|_F^2 = \langle v - v^{-1}; v - v^{-1}\rangle = \\ &\langle v; v\rangle + \langle v^{-1}; v^{-1}\rangle - 2\langle v; v^{-1}\rangle = tr(v^2) + tr(v^{-2}) - 2tr(e). \end{aligned} \tag{4.15}$$

We can rewrite x^+ and s^+ as follows

$$\begin{aligned} x^+ &= x + \Delta x = x + \sqrt{\mu}P(w)^{\frac{1}{2}}d_x = \sqrt{\mu}P(w)^{\frac{1}{2}}(v + d_x) \\ s^+ &= s + \Delta s = s + \sqrt{\mu}P(w)^{-\frac{1}{2}}d_s = \sqrt{\mu}P(w)^{-\frac{1}{2}}(v + d_s) \end{aligned} \tag{4.16}$$

As we know from [8] both $P(w)^{-\frac{1}{2}}$ and $P(w)^{\frac{1}{2}}$ are automorphisms on $\mathscr{K}$, and therefore x^+,s^+ are feasible (strictly feasible) if and only if $v+d_x, v+d_s \in \mathscr{K}$ $(v+d_x, v+d_s \in int\mathscr{K})$.

The following result will be used in proving the feasibility of iterates.

Lemma 4.1.3. *If $\delta(v) \leq 1$, then $e+d_x \circ d_s \in K$, if $\delta(v) < 1$, then $e+d_x \circ d_s \in intK$.*

Proof. Since $\langle d_x \circ d_s \rangle \geq 0$ according to Lemma 4.1.1,

$$-\frac{1}{4}\|d_x + d_s\|_F^2 e \preceq_K d_x \circ d_s \preceq_K \frac{1}{4}\|d_x + d_s\|_F^2 e \tag{4.17}$$

implying

$$|\lambda|\{d_x \circ d_s\}| \leq \frac{1}{4}\|d_x + d_s\|_F^2 = \delta^2(v) \tag{4.18}$$

and

$$-\delta^2(v) \leq \lambda\{d_x \circ d_s\}) \leq \delta^2(v) \tag{4.19}$$

$$1 - \delta^2(v) \leq \lambda\{e + d_x \circ d_s\} \leq 1 + \delta^2(v) \tag{4.20}$$

Therefore, if $\delta(v) \leq 1$

$$\lambda\{e + d_x \circ d_s\} \geq 0 \Longleftrightarrow e + d_x \circ d_s \in \mathscr{K} \tag{4.21}$$

and if $\delta(v) < 1$

$$\lambda\{e + d_x \circ d_s\} > 0 \Longleftrightarrow e + d_x \circ d_s \in int\mathscr{K} \tag{4.22}$$

Q.E.D. □

The following theorem will be our main result.

Theorem 4.1.4. *The iterates with full NT-step are feasible if $\delta(v) \leq 1$ and strictly feasible if $\delta(v) < 1$.*

Proof. For $0 \leq \alpha \leq 1$ denote

$$v_x^\alpha = v + \alpha d_x || v_s^\alpha = v + \alpha d_s \tag{4.23}$$

and consider

$$v_x^\alpha \circ v_s^\alpha = (v + \alpha d_x) \circ (v + \alpha d_s) = \tag{4.24}$$

$$\begin{aligned} &= v^2 + \alpha v \circ (d_x + d_s) + \alpha^2 d_x \circ d_s = v^2 + \alpha v \circ (v^{-1} - v) + \alpha^2 d_x \circ d_s = \\ &\qquad = (1-\alpha)v^2 + \alpha e + \alpha^2 d_x \circ d_s \end{aligned} \tag{4.25}$$

Case $\delta(v) \leq 1$.

From Lemma 4.2.1 we know that $e + d_x \circ d_s \in K$ thus $d_x \circ d_s \succeq_K -e$, hence

$$v_x^\alpha \circ v_s^\alpha \succeq_K (1-\alpha)v^2 + \alpha e - \alpha^2 e =$$

$$= (1-\alpha)(v^2 + \alpha e).$$

For $\alpha \in [0; 1)$, $(1-\alpha)(v^2 + \alpha e) \in intK$ and

$$v_x^\alpha \circ v_s^\alpha \succ_K 0.$$

If we now use Lemma 4.51[8] we shall have for $\alpha \in [0; 1)$

$$det(v_x^\alpha) \neq 0,$$

and

$$det(v_s^\alpha) \neq 0$$

Since

$$det(v_x^0) = det(v_s^0) = det(v) > 0,$$

because of continuity of both vectors for $\alpha \in [0;1)$

$$det(v_x^\alpha), det(v_s^\alpha) > 0.$$

Thus, for any eigenvalues of respective vectors we have (assuming again that $\alpha \in [0;1)$) $\lambda(v_x^\alpha) > 0$ and $\lambda(v_s^\alpha) > 0$. Using the continuity with respect to α once more we have

$$\lambda(v_x^1) \geq 0 \ , \ \lambda(v_s^1) \geq 0, \tag{4.26}$$

and we are done taking into consideration our remark on feasibility.

Case $\delta(v) < 1$.

Using Lemma 4.2.1 as above for $\alpha \in [0;1]$

$$v_x^\alpha \circ v_s^\alpha \succ_K (1-\alpha)v^2 + \alpha e - \alpha^2 e =$$

$$= (1-\alpha)(v^2 + \alpha e) \geq 0,$$

which means that $v_x^\alpha \circ v_s^\alpha \in intK$. Again by lemma 4.51[8] we have

$$det(v_x^\alpha) \neq 0,$$

$$det(v_s^\alpha) \neq 0.$$

and, because

$$det(v_x^0) = det(v_s^0) = det(v) > 0, rt$$

by continuity for $\alpha \in [0;1]$

$$det(v_x^\alpha) > 0, \tag{4.27}$$

$$det(v_s^\alpha) > 0,$$

which leads us to

$$v + d_x, v + d_s \in int\mathscr{K}. \tag{4.28}$$

Q.E.D. $\square$

4.2 Quadratic convergence of iterates

In this section we shall discuss the question of convergence of the Algorithm stated in Figure 3.1. To start we state some auxiliary lemmas which then will be used in the proof of the convergence of the Algorithm.

Lemma 4.2.1. *Let $x, s \in intK$, $\mu > 0$. Then*

$$\langle x^+; s^+\rangle = \mu(tr(e) + \langle d_x; d_s\rangle). \tag{4.29}$$

Proof.

$$\langle x^+; s^+\rangle = \langle \sqrt{\mu}P(w)^{\frac{1}{2}}(v + d_x); \sqrt{\mu}P(w)^{-\frac{1}{2}}(v + d_s)\rangle =$$

$$= \mu(\langle v; v\rangle + \langle v; d_x + d_s\rangle + \langle d_x; d_s\rangle) =$$

$$= \mu(\langle v; v\rangle + \langle v; v^{-1} - v\rangle + \langle d_x; d_s\rangle) =$$

$$= \mu(\langle v; v\rangle + \langle v; v^{-1}\rangle - \langle v; v\rangle + \langle d_x; d_s\rangle) =$$

$$= \mu(tr(e) + \langle d_x; d_s\rangle)$$

Q.E.D. □

Lemma 4.2.2. ***(Analog of Lemma 5.5 in** [**8**]**)*** *If we denote*

$$v^+ = \frac{P(w^+)^{-\frac{1}{2}}x^+}{\sqrt{\mu}} = \frac{P(w+)^{\frac{1}{2}}s^+}{\sqrt{\mu}}, \tag{4.30}$$

then

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

Proof. By the scaling point lemma(Lemma 4.46 in [8]) we have

$$\sqrt{\mu}v^+ = P(w^+)^{\frac{1}{2}}s^+ \sim (P(x^+)^{\frac{1}{2}}s^+)^{\frac{1}{2}}. \tag{4.31}$$

By Lemma 4.45 in [8] we have

$$P(x^+)^{\frac{1}{2}}s^+ = \mu P(P(w)^{\frac{1}{2}}(v+d_x))^{\frac{1}{2}} \cdot P(w)^{\frac{-1}{2}}(v+d_s) \sim \mu P(w)^{\frac{1}{2}}(v+d_x)(v+d_s) \tag{4.32}$$

Thus,

$$v^+ \sim (P(v+d_x)^{\frac{1}{2}}(v+d_s))^{\frac{1}{2}},$$

which concludes the proof. □

Lemma 4.2.3. ***(Lemma 4.56 in*** [***8***]***)*** *Let* $x, s \in int\mathscr{K}$, *then*

$$\|P(x)^{\frac{1}{2}}s - e\|_F \leq \|x \circ s - e\|_F \tag{4.33}$$

Lemma 4.2.4. ***(Lemma 4.58 in*** [***8***]***)*** *Let* $x, s \in int\mathscr{K}$, *then*

$$\lambda_{min}\left(P(x)^{\frac{1}{2}}s\right) \geq \lambda_{min}\left(x \circ s\right). \tag{4.34}$$

What follows shall be named Quadratic Convergence Theorem.

Theorem 4.2.5. *If* $\delta = \delta(v) < 1$, *then the full NT-step, defined in algorithm in figure* 3.1, *is strictly feasible and*

$$\delta(x^+, s^+, \mu) = \delta(v^+) \leq \frac{\delta^2}{\sqrt{2(1-\delta^2)}}. \tag{4.35}$$

Proof. From Theorem 4.1.4 we - because $\delta < 1$ - have $v+d_x, v+d_s, (v+d_x)\circ(v+d_s) \in intK$ Let us denote(as in [8])

$$u = P(v+d_x)^{\frac{1}{2}}(v+d_s) \in int\mathscr{K}, \tag{4.36}$$

$$\bar{u} = (v+d_x) \circ (v+d_s) \in int\mathscr{K}. \tag{4.37}$$

From Lemma 4.2.2 it follows that $v^+ \sim u^{\frac{1}{2}}$, thus,

$$v^+ - (v^+)^{-1} \sim u^{\frac{1}{2}} - u^{-\frac{1}{2}}. \tag{4.38}$$

Now

$$2\delta(v^+) = \|v^+ - (v^+)^{-1}\|_F = \|u^{\frac{1}{2}} - u^{-\frac{1}{2}}\|_F = \|u^{-\frac{1}{2}}(u-e)\|_F \leq$$
$$\leq \frac{\|u-e\|_F}{\lambda_{min}(u^{\frac{1}{2}})} = \frac{\|u-e\|_F}{\lambda_{min}(u)^{\frac{1}{2}}}.$$

At the same time, using Lemma 4.2.3, we get

$$\|P(v+d_x)^{\frac{1}{2}}(v+d_s)\|_F \leq \|(v+d_x)\circ(v+d_s)\|_F \tag{4.39}$$

in our terms it is

$$\|u-e\|_F \leq \|\bar{u}-e\|_F \tag{4.40}$$

and with the help of Lemma 4.2.4 we arrive at

$$\lambda_{min}(P(v+d_x)^{\frac{1}{2}}(v+d_s)) \geq \lambda_{min}((v+d_x)\circ(v+d_s)) \tag{4.41}$$

which is again simplified with our notation to

$$\lambda_{min}(u) \geq \lambda_{min}(\bar{u}). \tag{4.42}$$

Hence,

$$2\delta(v^+) \leq \frac{\|u-e\|_F}{\lambda_{min}(u)^{\frac{1}{2}}} \leq \frac{\|\bar{u}-e\|_F}{\lambda_{min}(\bar{u})^{\frac{1}{2}}}. \tag{4.43}$$

Recall that

$$\begin{aligned}\bar{u} &= (v+d_x)\circ(v+d_s) = v^2 + v\circ(d_x+d_s) + d_x\circ d_s = \\ &= v^2 + v\circ(v^{-1}-v) + dx\circ d_s = v^2 + e - v^2 + d_x\circ d_s = e + d_x\circ d_s\end{aligned} \tag{4.44}$$

If we now combine two last formulas, by substituting (4.44) into (4.43) we will get

$$2\delta(v^+) \leq \frac{\|e+d_x\circ d_s - e\|_F}{\lambda_{min}(e+d_x\circ d_s)^{\frac{1}{2}}} \leq \frac{\|d_x\circ d_s\|_F}{1+\lambda_{min}(d_x\circ d_s)^{\frac{1}{2}}}. \tag{4.45}$$

Now, we can use Lemma 4.1.1 in order to get

$$\lambda_{min}(d_x\circ d_s) \geq -\frac{1}{4}\|d_x+d_s\|_F^2 = -\delta^2(v) = -\delta^2, \tag{4.46}$$

$$1 + \lambda_{min}(d_x \circ d_s) \geq 1 - \delta^2. \tag{4.47}$$

From Lemma 4.1.2 we have

$$\|d_x \circ d_s\|_F \leq \frac{1}{2\sqrt{2}} \|d_x + d_s\|_F^2 = \sqrt{2}\delta^2 \tag{4.48}$$

Bottom line, we have

$$2\delta(v^+) \leq \frac{\sqrt{2}\delta^2}{\sqrt{1-\delta^2}} \tag{4.49}$$

and the inequality

$$\delta(v^+) \leq \frac{\delta^2}{\sqrt{2 \cdot (1-\delta^2)}} \tag{4.50}$$

concludes the proof.

Q.E.D. □

4.3 Updating the barrier parameter

Lemma 4.3.1. *Let $x, s \in intK, \delta = \delta(x, s; \mu)$, and $tr(x \circ s) = \mu(tr(e) + tr(d_x \circ d_s))$. If $\mu^+ = (1-\theta)\mu$ for some $\mu \in (0;1)$, then*

$$\delta(x, s; \mu^+)^2 \leq \frac{2\theta - \theta^2}{4(1-\theta)} \cdot \delta^2 tr(e) + (1-\theta)\delta^2 + \frac{\theta^2}{4(1-\theta)} tr(e). \tag{4.51}$$

Proof. Denote $\mu^+ = (1-\theta)\mu$, then $v(x, s, \mu^+) = \frac{v}{\sqrt{1-\theta}}$. Recall that because of commutativity of the inner product(trace in our case) we can write

$$\mu tr(v^2) = \mu tr\left(\frac{P(w)^{-\frac{1}{2}}x}{\sqrt{\mu}} \circ \frac{P(w)^{\frac{1}{2}}s}{\sqrt{\mu}}\right) =$$

$$= tr(x \circ s) = \mu(tr(e) + tr(d_x \circ d_s)). \tag{4.52}$$

Thus,

$$\|v\|_F^2 = tr(v^2) = tr(e) + tr(d_x \circ d_s). \tag{4.53}$$

Now, consider

$$4\delta^2(x,s;\mu^+) = \|\frac{v}{\sqrt{1-\theta}} - \sqrt{1-\theta}v^{-1}\|_F^2 = \|\frac{v}{\sqrt{1-\theta}} - \sqrt{1-\theta}v + \sqrt{1-\theta}v - \sqrt{1-\theta}v^{-1}\|_F^2 =$$

$$= \|\frac{\theta v}{\sqrt{1-\theta}} + \sqrt{1-\theta}(v - v^{-1})\|_F^2 = \frac{\theta^2\|v\|_F^2}{1-\theta} + (1-\theta)\|v - v^{-1}\|_F^2 + 2\theta\langle v; v - v^{-1}\rangle =$$

$$= (\frac{\theta^2}{1-\theta} + 2\theta)\|v\|_F^2 + 4(1-\theta)\delta^2 - 2\theta tr(e) =$$

$$= (\frac{\theta^2}{1-\theta} + 2\theta)(tr(e) + tr(d_x \circ d_s)) + 4(1-\theta)\delta^2 - 2\theta tr(e) =$$

$$= (\frac{\theta^2}{1-\theta} + 2\theta)tr(d_x \circ d_s) + 4(1-\theta)\delta^2 + \frac{\theta^2}{1-\theta}tr(e) \tag{4.54}$$

Now, we recall that due to the Lemma 4.1.1 we have

$$-\frac{1}{4}\|d_x + d_s\|_F^2 e \preceq_K d_x \circ d_s \preceq_K \frac{1}{4}\|d_x + d_s\|_F^2 e, \tag{4.55}$$

which means, that we can apply the trace-operator to the second inequality and obtain

$$tr(d_x \circ d_s) \leq \frac{1}{4}\|d_x + d_s\|_F^2 tr(e) = \delta^2 tr(e) \tag{4.56}$$

Continuing with the formula for $\delta^2(x,s;\mu^+)$

$$4\delta^2(x,s;\mu^+) = (\frac{\theta^2}{1-\theta} + 2\theta)tr(d_x \circ d_s) + 4(1-\theta)\delta^2 + \frac{\theta^2}{1-\theta}tr(e) \leq$$

$$\leq (\frac{\theta^2}{1-\theta} + 2\theta) \cdot \frac{1}{4}\|d_x + d_s\|_F^2 tr(e) + 4(1-\theta)\delta^2 + \frac{\theta^2}{1-\theta}tr(e) =$$

$$= \frac{2\theta - \theta^2}{1-\theta} \cdot \delta^2 tr(e) + 4(1-\theta)\delta^2 + \frac{\theta^2}{1-\theta}tr(e)$$

□

4.4 Iteration bound

Let again $\mathscr{J}$ be a Euclidean Jordan Algebra with rank r, and $\mathscr{K}$ be its respective cone of squares. In the Algorithm in the Figure 3.1 set $\tau = \frac{1}{\sqrt{2}}, \theta = \frac{1}{\sqrt{2r}}$. Therefore, $\delta(x_0, s_0, \mu_0) \leq \tau = \frac{1}{\sqrt{2}}$. First, we update the current point. Using Theorem 4.2.5 we have

$$\delta(x^+, s^+; \mu) \leq \frac{\delta^2}{\sqrt{2(1-\delta^2)}} \leq \frac{\frac{1}{2}}{\sqrt{2(1-\frac{1}{2})}} = \frac{1}{2} < \frac{1}{\sqrt{2}}. \tag{4.57}$$

It is crucial that $\delta(x^+, s^+; \mu) \leq \frac{1}{2}$. We will rely on this estimate heavily in order to prove that the Algorithm in the Figure 3.1 is well defined.

Next, we update the barrier parameter $\mu^+ = (1-\theta)\mu$, and using Lemma 4.3.1 we obtain, using abbreviation $\delta^+ = \delta(x^+, s^+; \mu)$,

$$\begin{aligned}\delta(x^+, s^+; \mu^+)^2 &\leq \frac{2\theta - \theta^2}{4(1-\theta)}(\delta^+)^2 tr(e) + (1-\theta)(\delta^+)^2 + \frac{\theta^2}{4(1-\theta)} tr(e) = \\ &= \frac{2\theta - \theta^2}{4(1-\theta)}(\delta^+)^2 r + (1-\theta)(\delta^+)^2 + \frac{\theta^2}{4(1-\theta)} r\end{aligned} \tag{4.58}$$

Now, using the estimate (4.57), we check whether the inequality $\delta(x^+, s^+; \mu^+) \leq \frac{1}{\sqrt{2}}$ is satisfied.

$$\begin{aligned}\delta(x^+, s^+; \mu^+)^2 &\leq \frac{2\theta - \theta^2}{16(1-\theta)} \cdot r + \frac{(1-\theta)}{4} + \frac{\theta^2}{4(1-\theta)} r = \\ &= \frac{2\theta}{16(1-\theta)} \cdot r - \frac{\theta^2}{16(1-\theta)} \cdot r + \frac{(1-\theta)}{4} + \frac{\theta^2}{4(1-\theta)} r = \\ &= \frac{1}{8\sqrt{2}(1-\theta)} - \frac{1}{32r(1-\theta)} + \frac{(1-\theta)}{4} + \frac{1}{8r(1-\theta)} = \\ &= \frac{1}{8\sqrt{2}(1-\theta)} + \frac{(1-\theta)}{4} + \frac{3}{32r(1-\theta)} \leq \frac{15}{32} < \frac{1}{2}\end{aligned} \tag{4.59}$$

The last non trivial inequality in (4.59) is due to the fact that the function $f(\theta) = \frac{1}{8\sqrt{2}(1-\theta)} + \frac{(1-\theta)}{4} + \frac{3}{32r(1-\theta)}$ is convex on $[0; \frac{1}{2}]$ and $f(0), f(\frac{1}{2}) \leq \frac{15}{32}$ Thus, the algorithm is well defined, we start with the property $\delta \leq \tau = \frac{1}{\sqrt{2}} = \tau$ and maintain it through every iteration.

Now we are ready to calculate the upper bound on the number of iterations the Algorithm in figure 3.1 needs to obtain an ε-approximate solution (i.e. a solution pair x, s satisfying $\langle x; s\rangle \leq \varepsilon$). At each iteration the duality gap, due to Lemma 4.2.1, is

$$\langle x^+; s^+\rangle = \mu(tr(e) + \langle d_x; d_s\rangle) \leq \mu r(1+\delta^2) \leq \mu\left(\frac{3r}{2}\right). \tag{4.60}$$

After each iteration μ is reduced by the factor $(1-\theta)$. Therefore after n iterations the duality gap would satisfy

$$\langle x^+; s^+\rangle \leq (1-\theta)^n \mu_0 \left(\frac{3r}{2}\right). \tag{4.61}$$

As we need $\langle x^+; s^+\rangle \leq \varepsilon$, it suffices to require $(1-\theta)^n\mu_0\left(\frac{3r}{2}\right) \leq \varepsilon$. Taking logarithm of this inequality we obtain

$$n \log 1 - \theta + \log \mu_0 \frac{3r}{2} \leq \log \varepsilon \tag{4.62}$$

or equivalently

$$\log \mu_0 \frac{3r}{2} - \log \varepsilon \leq -n \log 1 - \theta \tag{4.63}$$

Because $(\forall \theta \in [0,1))(-\log(1-\theta) \geq \theta)$ we have

$$\log \mu_0 \frac{3r}{2} - \log \varepsilon \leq n\theta \tag{4.64}$$

or in other words

$$n = \left[\frac{1}{\theta} \log \mu_0 \frac{3r}{2\varepsilon}\right] + 1, \tag{4.65}$$

where $[z]$ means the integral part of z, would be more than enough.

Thus we just proved

Theorem 4.4.1. *If $\tau = \frac{1}{\sqrt{2}}$, $\theta = \frac{1}{r\sqrt{2}}$, then for arbitrary $\varepsilon > 0$ the maximum number of iterations the Algorithm in Figure 3.1 needs to obtain ε-approximate solution is*

$$n = \left[\frac{1}{\theta} \log \mu_0 \frac{3r}{2\varepsilon}\right] + 1, \tag{4.66}$$

which can be written in asymptotic Landau symbolic notation as

$$O\left(r \log\left(\frac{r}{\varepsilon}\right)\right). \tag{4.67}$$

CHAPTER 5
CONCLUSION

Linear Complementarity Problem is very important in Optimization theory. It is not an optimization problem per se, however it is closely connected to optimization problems because optimality conditions of several classes of important optimization problems can be formulated as LCP.

This makes finding efficient methods for solving LCP an important issue. The most popular among these methods used to belong to finite simplex type pivoting based algorithms, such as Lemke's method.

Recently a new group of powerful methods,Interior Point Methods (IPM), have been developed. These methods are based on Newton method; and are, therefore, iterative methods.

In this thesis we have considered a generalization of LCP from it's traditional formulation (over non-negative orthrant in $\mathbb{R}^n$) to LCP over symmetric cones. We developed a feasible IPM for this type of problem. This algorithm is full Newton step method, when a Newton method is being used on each iteration in order to find the search directions. The advantage of full step approach lies in the fact that we avoid calculating the step size which consumes processing time and resources. In order to develop and analyze the algorithm we used sophisticated theoretical apparatus of Euclidean Jordan algebras.

We proved that with appropriate choice of threshold parameter τ and barrier decrease parameter θ - namely $\tau = \frac{1}{\sqrt{2}}$ and $\theta = \frac{1}{r\sqrt{2}}$ - the algorithm converges globally. It arrives at ε-approximate solution in at most $O(r \log\left(\frac{r}{\varepsilon}\right))$ iterations.

BIBLIOGRAPHY

[1] F. Alizadeh and S. H. Schmieta. Associative and Jordan algebras, and polynomial time interior-point algorithms for symmetric cones. *Mathematics of Operations Research,* 26(3): 543-564, 2001.

[2] F. Alizadeh and S. H. Schmieta. Extension of primal-dual interior-point algorithms to symmetric cones. *Mathematical Programming,* 96(3): 409-438, 2003.

[3] R. W. Cottle and G. B. Dantzig. Complementary pivot theory of mathematical programming. *Linear Algebra and its Applications,* 1:103-125, 1968.

[4] J. Faraut and A. Koràny. *Analysis on Symmetric Cones.* Oxford Mathematical Monographs. The Clarendon Press Oxford University Press, New York, 1994.

[5] L. Faybusovich, Euclidean-Jordan algebras and interior-point algorithms, *Positivity*, 1: 331-335, 1997.

[6] L. Faybusovich, Linear systems in Jordan algebras and primal-dual interior-point algorithms, *J. Comput. Appl. Math.*, 86(1): 149-175, 1997.

[7] O. Güler. Barrier functions in interior-point methods. *Mathematics of Operations Research,* 21(4): 860-885, 1996.

[8] G. Gu, Full-Step Interior-Point Methods for Symmetric Optimization, *TU Delft Ph.D. thesis*, 2009.

[9] Karmarkar N., *A New Polynomial Time Algorithm for Linear Programming*, Combinatorica, Vol 4, No. 4, (1984).

[10] Khachiyan L.G., *A Polynomial Algorithm in Linear Programming*, Soviet Mathematics Doklady, 20, pp. 373-395 (1984).

[11] Klee V. and Minty G.J., *How Good is the Simplex Algorithm?* Inequalities, III, pp. 159-175, Academic Press, New York, NY, (1972).

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

[13] Mansouri H., *Full-Newton-Step Interior-point Methods for conic Optimization*, Ph.D. Thesis, TU Delft, Netherland, (2008).

[14] Nesterov Y., Nemirovski A., Interior-Point Polynomial Algorithms in Convex Programming. *SIAM Studies in Applied Mathematics*, Philadelphia, PA, (1994).

[15] Nemirovski A.S., Yudin D.B. Informational complexity and effective methods for the solution of convex extremal problems. *Ekon. Mat. Metody* 12: 357-369, 1976.

[16] Y. Nesterov and M. J. Todd. Self-scaled barriers and interior-points methods for convex programming. *Mathematics of Operations Research*, 22(1): 1-42, 1997.

[17] Y. Nesterov and M. J. Todd. Primal-dual interior-point methods for self-scaled cones. *SIAM Journal on Optimization,* 8(2): 324-364, 1998.

[18] Peng J.S., Roos C., Terlak T., *Self-Regularity: A New Paradigm for Primal-Dual Interior-Point Algorithms*, Princeton University Press, (2002)

[19] B.K. Rangarajan. Polynomial convergence of infeasible-interior-point methods over symmetric cones. *SIAM Journal on Optimization,* 16(4): 1211-1229, 2006.

[20] Roos C., Terlaky T., Vial J.P., *Theory and Algorithms for Linear Optimization*, John Wiley and Sons, Chichester, UK (1997).

[21] J. Peng, C. Roos, and T. Terlaky. *Self-Regularity: A New Paradigm for Primal-Dual Interior-Point Algorithms.* Princeton University Press, 2002.

[22] C. Roos, T. Terlaky, and J. Ph. Vial. *Theory and Algorithms for Linear Optimization. An Interior-Point Approach.* Springer Science, 2005.

[23] Shor N.Z., *Cut-off Method with Space Extension in Convex Programming Problems*, Cybernetics 13, pp. 94-96, (1977).

[24] M. Vieira, Jordan Algebraic Approach to Symmetric Optimization. *Ph.D. Thesis*, TU Delft, 2007.

[25] A. Yoshise. Interior point trajectories and homogeneous model for nonlinear complementarity problems over symmetric cones. *SIAM J. Optim.*, 17(4): 1129-1153, 2006.

[26] A. Yoshise. Complementarity problems over symmetric cones: a survey of recent developments in several aspects. in Anjos M.F., Lasserre J.B. (eds.): *Handbook of Semidefinite, Conic and Polynomial Optimization Theory, Algorithms, Software and Applications.* International series in Operations Research and Management Science. Volume 166: 339-376, Springer, 2012.