Inner approximation of convex cones via primal-dual ellipsoidal norms

by

Miaolan Xie

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Mathematics
in
Combinatorics and Optimization

Waterloo, Ontario, Canada, 2016

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

We study ellipsoids from the point of view of approximating convex sets. Our focus is on finding largest volume ellipsoids with specified centers which are contained in certain convex cones. After reviewing the related literature and establishing some fundamental mathematical techniques that will be useful, we derive such maximum volume ellipsoids for second order cones and the cones of symmetric positive semidefinite matrices. Then we move to the more challenging problem of finding a largest pair (in the sense of geometric mean of their radii) of primal-dual ellipsoids (in the sense of dual norms) with specified centers that are contained in their respective primal-dual pair of convex cones.

Acknowledgments

Table of Contents

List of Figures

Chapter 1

Introduction

In this introduction chapter, we will introduce some fundamental concepts, and describe the problems we studied with their motivation. After that, we will present the overall structure for the rest of the thesis.

Some subsets of finite dimensional Euclidean spaces have the property that for every pair of points in the set, the line segment that joins the pair of points is entirely contained in the set. Such sets are called *convex*.

A compact convex set with non-empty interior is called a *convex body*, and a *convex hull* of a set C is the smallest (with respect to set inclusion) convex set that contains C, denoted by $\text{conv}(C)$.

Let $\mathbb{R}^n$ denote the n-dimensional Euclidean space with the standard inner-product, and $\mathbb{E}$ denote a general Euclidean space of dimension n without a specific inner product. Let $\mathbb{R}^n_+$ denote the non-negative orthant of $\mathbb{R}^n$, which is the set of vectors whose entries are all nonnegative. The Euclidean ball centered at $\bar{x} \in \mathbb{R}^n$ with radius r is denoted as:

$$B(\bar{x}, r) := \{x \in \mathbb{R}^n : \langle (x - \bar{x}), I(x - \bar{x}) \rangle \leq r^2\}.$$

By the above definition on convexity, we have the empty set, $\mathbb{R}^n$, the non-negative orthant, and any Euclidean ball are all convex sets.

A *closed half space* is the set
$$\{x \in \mathbb{R}^n : \langle a, x \rangle \leq b\},$$
for some $a \in \mathbb{R}^n$ and $b \in \mathbb{R}$. Notice that when a is the zero vector, $\langle a, x \rangle = 0$, if $b \geq 0$, the set is the whole space $\mathbb{R}^n$, and if $b < 0$, the set is the empty set. Every closed half space is a closed convex set.
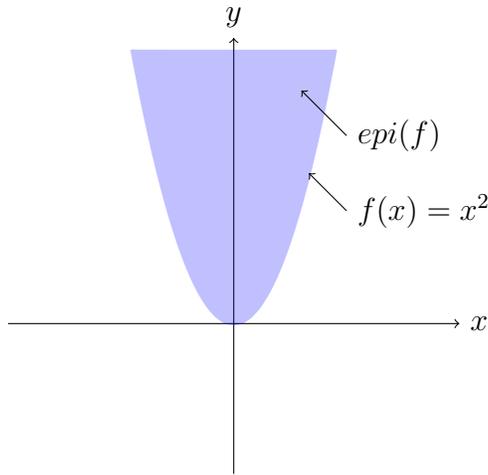
Figure 1.1: $epi(f) = \{(x, \mu) : x, \mu \in \mathbb{R}, \mu \geq x^2\}$.

The intersection of any collection (finite or infinite) of convex sets is always convex. Hence, the intersection of any collection of closed half spaces is again closed and convex. A *polyhedron* is the intersection of a finite set of closed half spaces, and a *polytope* is a polyhedron that is bounded. An equivalent definition for polytope is that it is the convex hull of finitely many points. This equivalence has connections to Minkowski's work from more than 100 years ago. A *face* of a polyhedron P is the intersection of P and some of its defining closed half spaces.

We can represent a function using a geometric set, by defining the notion of *epigraph*. The *epigraph* of a function $f : \mathbb{E} \rightarrow [-\infty, +\infty]$ is the set of points lying on or above its graph, i.e.,

$$epi(f) := \{(x, \mu) : x \in \mathbb{E}, \mu \in \mathbb{R}, \mu \geq f(x)\}.$$

Consider the function $f : \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = x^2$ as illustrated in Figure 1.1. The shaded area is the epigraph of f.

A *convex function* is a function whose epigraph is a convex set. For instance, the epigraph of the function $g : \mathbb{R} \rightarrow \mathbb{R}$, $g(x) := x^3, x \in \mathbb{R}$ is not convex, so g is not a convex function.

Convex optimization deals with the problem of minimizing convex functions over convex sets. Any convex optimization problem can be written as:

$$\min : f(x)$$
$$x \in \mathcal{F},$$

where $f : \mathbb{E} \to \mathbb{R}$ is a convex function, and $\mathcal{F} \subseteq \mathbb{E}$ is a convex set. This optimization problem is equivalent to the problem of the following form, with the introduction of a new variable t:

$$\min : t$$
$$f(x) \leq t$$
$$x \in \mathcal{F}.$$

The set of feasible solutions for the constraint $f(x) \leq t$ is precisely the set:

$$epi(f) = \left\{ \begin{pmatrix} x \\ t \end{pmatrix} \in \mathbb{E} \oplus \mathbb{R} : f(x) \leq t \right\}.$$

Since f is a convex function, by definition, its epigraph is a convex set. $\mathcal{F}$ is convex to begin with, so the intersection of $\mathcal{F}$ and $epi(f)$ must again be a convex set. Hence, the feasible region of the new formulation is convex. With this new equivalent formulation, we see that any convex optimization problem can be reduced to a problem of minimizing a linear function over a convex feasible region. Thus, redefining $\mathbb{E}$, $\mathcal{F}$ and x, we have the following general form for convex optimization problems:

$$\min : \langle c, x \rangle$$
$$x \in \mathcal{F},$$

where $c \in \mathbb{E}^*$ defines the linear objective function, and $\mathcal{F} \subseteq \mathbb{E}$ is a convex set.

A *local minimum solution* of a convex optimization problem is a solution that has the minimum objective value within a neighboring set of possible solutions, and a *global minimum solution* is a solution that has the minimum objective value among all possible solutions. Convex optimization problems are in general "easier" to deal with than general optimization problems, since any local minimum solution is automatically a global minimum solution. Moreover, it is actually useful enough a tool even for solving general optimization problems. The reasons are as follows:

A non-convex optimization problem may be transformed to a convex optimization problem by taking the "convex hull" of the feasible region. Recall that the *convex hull* of a set C is the smallest (with respect to set inclusion) convex set that contains C, denoted by $conv(C)$. Without loss of generality, we may assume the objective function of the non-convex optimization problem is linear, since we can use the same technique as before to push any non-linear objective function into the constraints by adding a new variable and

constraint. Now, consider the following general non-convex optimization problem:

$$\min : \langle c, x \rangle$$
$$x \in \mathcal{U},$$

where $\mathcal{U}$ is a non-convex set. The convex optimization problem which corresponds to the above non-convex problem is:

$$\min : \langle c, x \rangle$$
$$x \in \operatorname{conv}(\mathcal{U})$$

In principle, we can solve the non-convex problem by solving its convex counterpart after this transformation. It turns out they have the same optimal objective value. Furthermore, for "most" problems, the optimal solution for the convex problem is unique and is also an optimal solution for the original non-convex problem (see Theorem 9.1 in [48]). Of course, this may not always be the most efficient approach computationally.

Since we can push any objective function and all the complexities into the constraints, the convex set constraints might be hard to deal with in general. We may try to first reduce the convex set constraint to some compact convex set constraint, and then use a much simpler convex set, say an ellipsoid, to approximate the resulting compact convex set and obtain an approximation for the original problem.

Ellipsoids are generalizations of Euclidean balls, which are very simple objects to work with. Formally, an ellipsoid centered at $\bar{x}$ can be defined as the set of x that satisfies the inequality:
$$(x - \bar{x})^T H (x - \bar{x}) \leq 1.$$

The self-adjoint positive definite matrix H in the above equation defines the shape of the ellipsoid. Specifically, the eigenvectors and the eigenvalues of H determine the axes and their lengths of the ellipsoid.

Ellipsoids are simple enough objects that can be used to "approximate" more complex feasible regions. The minimum volume ellipsoid containing a set and the maximum volume ellipsoid contained in a set provide global approximations to a (bounded) convex set. These extremal ellipsoids are called *Löwner-John ellipsoids*. For a detailed history and the development of the study on Löwner-John ellipsoids please refer to Henk [23] and the references therein. Just as we can use ellipsoids to facilitate solving convex optimization problems, the same idea can be applied to non-convex problems via taking the convex hull of the feasible regions. In interior-point methods, ellipsoidal approximations to convex sets

are also used locally to define local variable metrics. Ellipsoids are indeed very useful for a wide class of optimization problems.

Moreover, ellipsoids have many desirable properties, for instance, they are intrinsically associated with convex quadratic functions, and any affine transformation of an ellipsoid is still an ellipsoid. In general, ellipsoids are used extensively in the fields of science, engineering and mathematics. We list some of their uses here:

1. The ellipsoid method for solving convex optimization problems:

   The *ellipsoid method* was first proposed by Yudin and Nemirovski [26], as well as by Shor [44] for solving convex minimization problems. The method makes extensive use of the idea of using an ellipsoid to approximate the underlying feasible region. This is achieved by finding an ellipsoid that contains the feasible region (or a subset of the feasible region which contains an optimal solution) at each iteration. In particular, the method generates a sequence of ellipsoids whose volume decreases proportionally at every step. Each time the method generates a new ellipsoid, it does so by computing a minimum volume ellipsoid which contains the desired subset of the current ellipsoid.

   In 1979, Khachiyan proved Linear Programing with rational data can be solved exactly, in polynomial-time, using the ellipsoid method. This result is of great theoretical importance. Khachiyan showed that within a polynomial number of iterations of the ellipsoid method, the objective value of the current iterate will be "close enough" to the optimal objective value so that every extreme point solution with at least as good an objective value has to be an optimal solution. Then, the last step of finding an exact optimal solution can be done in strongly polynomial time. Please refer to Khachiyan [21] [22] for more details.

2. Trust-region methods in non-linear Programming (NLP):

   The trust-region method is a common iterative method for solving non-linear minimization problems (typically unconstrained). In each iteration, it uses a local quadratic function to approximate the non-linear objective function over a small "trusted region". If the function is twice continuously differentiable, the Taylor approximation of degree two can be used. In each iteration of the algorithm, a minimizer of the quadratic approximation is found over an ellipsoid, which is the current trust region, and then another local quadratic approximation of the function is constructed with respect to this new point, with a new ellipsoid "trusted region" centered at that point. Radii of the ellipsoids defining the trust-region are adjusted throughout the

execution of the algorithm based on the amount of progress made with respect to minimization. Ellipsoids again play an important role in this algorithm. Please see Conn et al. [10], Pong and Wolkowicz [39] and the references therein for a treatment of this subject.

3. Determining uncertainty regions in robust optimization:

   In robust optimization, the parameters are usually uncertain and stretch over some ranges. One way to represent the uncertainty set over the data space is to use an ellipsoid or an intersection of ellipsoids, which is called the *ellipsoidal uncertainty set*. Ellipsoid plays an important role in this subject as well. A good comprehensive reference for robust optimization as of 2002 is Ben-Tal and Nemirovski [4], and another main reference is the book robust optimization by Ben-Tal et al. [2].

4. Upper and lower bounding for extremal geometric problems:

   In extremal geometry, the problem of finding good bounds for a functional over convex bodies is studied. Good approximations of the convex bodies themselves will give rise to desired upper and lower bounds of the functional. As ellipsoids are geometrically simple and easy to evaluate over functionals and quadratic functions, they are used often in practice to obtain reliable bounds. In particular, the maximum volume ellipsoid contained in the target geometric objects or the minimum volume ellipsoid containing the target geometric objects is a reasonable first attempt for the approximation. Please refer to Henk [23] for more discussion.

5. Describing regions of stability or controllability in system and control theory:

   Ellipsoids are also used in the field of engineering. In system and control theory, ellipsoids are used to describe regions of stability or controllability. Specifically, invariant ellipsoids are used to interpret quadratic stability. The problem of finding the smallest invariant ellipsoid containing a polytope and the problem of finding the largest invariant ellipsoid contained in a polytope are both important, as the polytope represents the allowable or safe operating region for the system. The largest invariant ellipsoid contained in a polytope can be interpreted as the region of initial conditions for which the state always remains in the safe operating region. Boyd et al. [7] and and the convex optimization book by Boyd and Vandenberghe [8] have good treatment on this subject.

6. Clustering and classification in data mining and estimation in statistics:

   In data mining and statistics, we can use ellipsoids to solve problems in clustering and classification. In order to find k clusters for a given set of data, we may find k

minimum volume covering ellipsoids, one for each cluster, which cover all the data points and minimize the geometric mean of the volumes of each cluster's covering ellipsoids. Please refer to Shioda and Tuncel [43] and the references therein for details. Finding the minimum volume covering ellipsoids (please see Sun and Freund [45]) is useful in identifying the outliers of the set, as the outliers will correspond to the data points lying on the boundary of the ellipsoid. Another related problem in statistics is finding the minimum volume ellipsoid that covers k of the n data points. The center of the ellipsoid is called the minimum volume ellipsoid location estimator, and the positive definite form of the ellipsoid is called the minimum scatter estimator (please see Croux et al. [11]).

7. Optimal experimental design in statistics:

   In statistics, the problem of finding the optimal experimental design is related to the problem of finding the minimum volume circumscribing ellipsoid. The design problem tries to identify the points at which the experiments will be carried out so that the estimation of the desired parameter will be "optimal". The problem is called the $D - optimal$ design problem, when the criterion is about the determinant of a matrix related to the inverse of the covariance matrix. In fact the $D - optimal$ design problem is exactly dual to the problem of finding the minimum volume circumscribing ellipsoid. Please see Gürtuna [20] for details and derivation.

8. Other applications and usage of the ellipsoids can be found in computational geometry, computer graphics and pattern recognition.

There is also publicly available software for various optimization problems involving extremal ellipsoids. For example, MVE package of Zhang and Gao [50] is a MATLAB based package, that given a full dimension polytope $P := \{x : Ax \leq b\}$, computes the maximum volume ellipsoid contained in P.

Convex cones are of main interests in modern optimization for their nice intrinsic properties. A *cone* is a set such that if x is in the set, λx is also in the set, for any $\lambda \geq 0$. A *convex cone* is a cone that is convex. In this thesis, we studied the problem of finding the largest volume ellipsoids with specified centers which are contained in certain convex cones.

The *second order cones* SOC^n are defined as

$$SOC^n := \{(x, y) \in \mathbb{R}^{n+1} : \langle x, x \rangle^{\frac{1}{2}} = ||x||_2 \leq y\}.$$

By definition, the intersection of any second order cone and a hyperplane orthogonal to the last axis of $\mathbb{R}^{n+1}$ is a n dimension ball, which shows the second order cones are highly symmetric.

A matrix M is *positive semidefinite*, if $x^T M x$ is non-negative for any vector $x \in \mathbb{R}^n$, and a matrix M is *positive definite*, if $x^T M x$ is positive for any vector $x \in \mathbb{R}^n \setminus \{0\}$. We denote the space of n-by-n symmetric matrices as $\mathbb{S}^n$. The *positive semidefinite cones* are defined as:
$$\mathbb{S}^n_+ := \{X \in \mathbb{S}^n : X \text{ is positive semidefinite}\}.$$

They are of central importance in semidefinite optimization. We use the notation $X \succeq 0$ to indicate X is a symmetric positive semidefinite matrix, and use $X \succ 0$ to denote X is a symmetric positive definite matrix. For $U, V \in \mathbb{S}^n$, we say $U \succeq V$, if and only if $U - V \succeq 0$, and $U \succ V$, if and only if $U - V \succ 0$. The order defined by $\succeq$ and $\succ$ as described is called the *Löwner order*. In the thesis, when we talk about positive definite or positive semidefinite matrix, we always assume that the matrix is symmetric or Hermitian.

It is straightforward to verify that the non-negative orthant, second order cones and positive semidefinite cones are also convex cones.

For any given cone $\mathcal{K}$, we can define the notion of a *dual cone* associated to it as follows:

Definition 1.0.1. *Suppose $\mathcal{K}$ is a convex cone in the vector space $\mathbb{E}$. The* dual cone *$\mathcal{K}^*$ of this convex cone is:*
$$\{s \in \mathbb{E}^* : \langle x, s \rangle \geq 0, \forall x \in \mathcal{K}\}.$$

Now, let us define the notion of *dual norm* which will be associated with a dual cone: Given a dual pairing $\langle \cdot, \cdot \rangle : \mathbb{E} \oplus \mathbb{E}^* \to \mathbb{R}$ and a norm $||\cdot||$ on $\mathbb{E}$, the *dual norm* $||\cdot||^*$ is defined by:
$$||s||^* := \sup\{\langle s, x \rangle : ||x|| \leq 1\}, \forall s \in \mathbb{E}^*.$$

Let H be an arbitrary self-adjoint positive definite operator. For every $x \in \mathbb{E}$, the *norm defined by H* is:
$$\|x\|_H := \langle x, Hx \rangle^{1/2}.$$

By the definition of dual norm, the dual norm for $||\cdot||_H$ is:
$$\|s\|_H^* := \sup_{\|x\|_H \leq 1} \langle s, x \rangle,$$

for every $s \in \mathbb{E}^*$.

By the above definitions, we see that the primal-dual norms defined by H have direct linkage with ellipsoids having shapes defined by H: an ellipsoid centered at $\bar{x}$ with radius

one, defined by the self-adjoint positive semidefinite form H corresponds to the set of points x such that $||x - \bar{x}||_H \leq 1$.

The dual norm of a norm defined by the self-adjoint positive definite operator H is actually a corresponding local norm that is defined by the operator H^{-1}. Before we state and prove this proposition on dual norms, let us first introduce some definitions and a result in linear algebra.

Let $X \in \mathbb{R}^{n \times n}$ be given. The roots $\lambda_1, \ldots, \lambda_n$ of the polynomial equation $\det(X - \lambda I) = 0$ are the *eigenvalues* of X. We can write them as a vector $\lambda(X) \in \mathbb{C}^n$. If X is a symmetric matrix over the reals, i.e., $X \in \mathbb{S}^n$, then all the eigenvalues of X are real. Let $\mathrm{Diag} : \mathbb{R}^n \to \mathbb{S}^n$ denote the linear map that maps a real vector $(x_1, \ldots, x_n)^T$ to the diagonal matrix:

$$\begin{pmatrix} x_1 & 0 & \ldots & 0 \\ 0 & x_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & x_2 \end{pmatrix}.$$

Theorem 1.0.2 (Spectral Decomposition). *For every $X \in \mathbb{S}^n$, there exists $Q \in \mathbb{R}^{n \times n}$, orthogonal $(Q^T Q = I)$ such that*

$$X = Q \, \mathrm{Diag}(\lambda(X)) Q^T.$$

With the spectral decomposition theorem, for $X \in \mathbb{S}^n_{++}$, we define its unique *symmetric positive definite square root* as:

$$X^{\frac{1}{2}} := Q[\mathrm{Diag}(\lambda(X))]^{\frac{1}{2}} Q^T.$$

This definition can be extended to $\mathbb{S}^n_+$, the set of symmetric positive semidefinite matrices.

Now, we can state and prove our proposition on dual norms (please see for instance Hiriart-Urruty and Lemaréchal [24]):

Proposition 1.0.3. *The dual norm (over $\mathbb{E}^*$) of the norm defined by a self-adjoint positive definite operator H over $\mathbb{E}$ is again defined by its inverse H^{-1}, i.e., $||s||_H^* = \langle s, H^{-1} s \rangle^{1/2}$.*

Proof. By definition, the dual norm for $s \in \mathbb{E}^*$ is

$$||s||_H^* = \sup_{||x||_H \leq 1} \langle s, x \rangle,$$
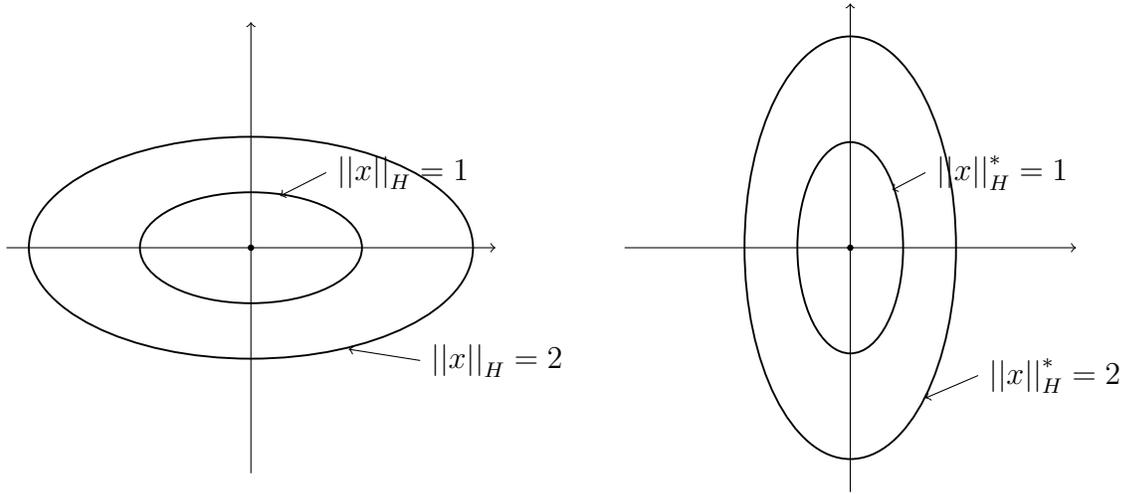
Figure 1.2: Illustration for dual norms in $\mathbb{R}^2$.

and it is clearly equal to $\sup_{\langle x, Hx \rangle \leq 1} \langle s, x \rangle$. Let us do a change of variable by taking

$$u := H^{\frac{1}{2}} x.$$

We obtain

$$\sup_{\langle x, Hx \rangle \leq 1} \langle s, x \rangle = \sup_{||u||^2 \leq 1} \left\langle u, H^{-\frac{1}{2}} s \right\rangle.$$

Here $||\cdot||$ denotes the standard Euclidean norm. For $s \neq 0$, we know the unique solution u that achieves the supremum is $u = \frac{H^{-\frac{1}{2}} s}{\left\| H^{-\frac{1}{2}} s \right\|}$, by Cauchy-Schwarz (see Corollary 2.3.4 in Chapter 2), and the supremum is $\langle s, H^{-1} s \rangle^{\frac{1}{2}}$. Hence, the result follows. $\square$

Please see a pictorial illustration of pairs of primal-dual norms on $\mathbb{R}^2$, in Figure 1.2.

Finding the extremal ellipsoids for a given convex set is clearly a fundamental problem both in optimization and other fields of science, and in modern convex optimization, cones are used extensively in the optimization formulation. In this thesis, we studied the problem of finding the largest volume ellipsoids with specified centers which are contained in certain convex cones. Specifically, we derive such maximum volume ellipsoids for non-negative orthant, second order cones and the cones of symmetric positive semidefinite matrices. The solution of this problem for non-negative orthant is relatively well-known. The solutions for the second order cones has been solved by Güler and Gürtuna [19], and we will use

different techniques to derive them in the thesis. We will also derive the solutions for the positive semidefinite cones in different ways, which are original results of this thesis.

We will then move to the more challenging problem of finding the largest pair (in the sense of geometric mean of their radii) of primal-dual ellipsoids (in the sense of dual norms) with specified centers that are contained in their respective primal-dual pair of convex cones. This problem has been proposed and solved by Todd [46] for all symmetric cones K.

Finding the largest dual ellipsoids inscribed in dual cones has connections to interior point methods (IPM). In IPM primal-dual symmetric methods, ellipsoidal approximations to convex sets are used locally to define local variable metrics. Specifically, in the iterations of many IPM, we want to find the "best" search directions and step sizes in both the primal and dual spaces, while remaining in the cone. This is guaranteed by the methods as they restrict the "search area" within a pair of primal-dual ellipsoids centered at the current iterates and contained in the primal-dual cones respectively, at each iteration. We will study the primal-dual version of the ellipsoidal approximation problem as described above in Chapter 5. The main reference is Todd [46]. There are also algorithms by Chua [9], Litvinchev [34], Renegar and Sondjaja [40] which utilize conic representations of ellipsoids to approximate the feasible region.

The overall structure of the remainder of this paper is as follows:

Chapter 2 introduces some more concepts related to convex sets, convex functions and convex cones. An introduction to a special class of convex functions called self-concordant barriers is also presented in Chapter 2. After that, some convex optimization formulations and some useful optimality conditions for the formulations are discussed. We will also see why optimizing a linear function over an ellipsoid is an easy problem.

Chapter 3 introduces the concept of Löwner-John ellipsoid and Löwner's theorem, which are of central interest of this thesis. We will derive various optimality conditions for finding the minimum volume ellipsoid containing and the maximum volume ellipsoid contained in a convex set. We will also discuss some general observations on maximum volume ellipsoids in convex cones.

Chapter 4 starts with a discussion on symmetries of ellipsoids and various convex cones. We derive the maximum volume ellipsoids with specified centers contained in various convex cones. We will start with the cones of nonnegative orthant, second order cones and then move to positive semidefinite and homogeneous cones.

Chapter 5 focuses on the problem of finding the largest primal-dual pairs of ellipsoids with specified centers in a pair of primal-dual convex cones. We start with the work done by Todd [46] for symmetric cones, then move on to its generalization by Lim [32], and end with some discussion on generalizing the results to homogeneous and hyperbolic cone settings.

11

Chapter 6 summarizes the thesis and concludes with some future research directions.

Chapter 2

Convex sets, convex cones, self-concordant barrier functions and some optimality conditions

In this chapter, we will begin with an introduction to a few more relevant concepts to convex sets, convex functions and convex cones. We will also introduce a special class of convex functions called self-concordant barriers. A self-concordant barrier function intrinsically encodes a convex cone (or in general a convex set) by having real values in the interior of the cone and when a sequence converging to a point on the boundary, the function value goes to infinity. We will then introduce some convex optimization formulations and some useful optimality conditions for them. In the last section, we will optimize a linear function over an ellipsoid, and see that it is indeed an easy problem.

2.1   Fundamental concepts

In this section, we will continue to introduce fundamental concepts and properties related to convex sets and cones which will be useful in the rest of the thesis. Let us first fix some notations. A convention we use in this thesis is that all the vectors are column vectors.

We will refer to $\{e_1, ..., e_n\}$ as the standard basis for $\mathbb{R}^n$, where:
$e_1 := (1, 0, ..., 0)^T, ..., e_n := (0, ..., 0, 1)^T$. We will denote the vector of all ones by e. The default norm used unless otherwise noted is the 2-norm induced by the standard Euclidean inner product.

We consider the matrix representations of symmetric bilinear positive definite forms as elements of $\mathbb{S}^n_{++}$. The default inner product we use in the symmetric matrix space is the trace inner product:

$$\langle X, Y \rangle := \mathrm{Tr}(X^T Y) = \sum_{i=1}^n \sum_{j=1}^n X_{ij} Y_{ij},$$

and the default norm of the symmetric matrix space is the norm associated to the trace inner product if not otherwise noted. This norm is called *the Frobenius norm* or *the operator 2-norm*. *The nuclear norm* on $\mathbb{S}^n$ or in general, the space of Hermitian matrices is defined as:

$$||X||_* := \mathrm{Tr}(\sqrt{X^* X}),$$

where X^* is the adjoint of X.

Let s2vec $: \mathbb{S}^n \to \mathbb{R}^{\frac{n(n+1)}{2}}$ be the isometry between $\mathbb{S}^n$ and $\mathbb{R}^{\frac{n(n+1)}{2}}$ that takes the lower triangular part of a symmetric matrix column-wise and multiplies the strict lower triangular part by $\sqrt{2}$. Explicitly,

$$\mathrm{s2vec}(A) := (A_{11}, \sqrt{2}A_{21}, \ldots, \sqrt{2}A_{n1}, A_{22}, \sqrt{2}A_{22}, \ldots, \sqrt{2}A_{n2}, \ldots, A_{nn})^T.$$

The inner product for A, B in $\mathbb{S}^n$ exactly corresponds to the standard vector inner product between s2vec(A) and s2vec(B). We denote sMat $: \mathbb{R}^{\frac{n(n+1)}{2}} \to \mathbb{S}^n$ as the inverse of s2vec.

We can also express the matrix representation of certain linear transformations on $\mathbb{S}^n$ (identified with $\mathbb{R}^{\frac{n(n+1)}{2}}$) using the *symmetric Kronecker product*. For every $X \in \mathbb{S}^n$, we require:

$$(A \overset{s}{\otimes} B)\,\mathrm{s2vec}(X) := \mathrm{s2vec}\left(\frac{1}{2}(BXA^T + AXB^T)\right).$$

Notice that $A \overset{s}{\otimes} B = B \overset{s}{\otimes} A$.

We use $\nabla f(x)$ or $f'(x)$ to denote the derivative of a function for simplicity. $f''(x)$ is the Hessian of f. For higher order derivatives, $D^n f(x)$ is used to denote the n-th derivative of a function $f(x)$, and $D^2 f(x)[h^{(1)}, h^{(2)}]$ denotes the second derivative of f evaluated along the directions $h^{(1)}, h^{(2)} \in \mathbb{E}$. We also use the notation $\langle f''(x)h^{(1)}, h^{(2)} \rangle$ to denote the same quantity. When we need to put emphasis on the set of variables we are differentiating with respect to, we will use the $\frac{\partial f}{\partial x}$ notation.

The interior of a cone K is denoted as $\mathrm{int}(K)$, and the boundary of the cone K is denoted as $\mathrm{bd}(K)$.

Now, let us begin by defining an important notion in convex analysis: the *convex conjugate function*.

Definition 2.1.1. *For a function $f : \mathbb{E} \to \mathbb{R} \cup \{+\infty\}$, the* convex conjugate *(also called Fenchel conjugate or Legendre-Fenchel conjugate)* $f_* : \mathbb{E}^* \to \mathbb{R} \cup \{+\infty\}$ *is defined as:*

$$f_*(y) := \sup\{-\langle y, x \rangle - f(x) : x \in \mathbb{E}\}.$$

Some of the literature use instead:

$$f^*(y) := \sup\{\langle y, x \rangle - f(x) : x \in \mathbb{E}\} \tag{2.1}$$

to define the convex conjugate of f. As we will see later, the notion of convex conjugacy is a notion of duality for functions. Convex conjugacy has a natural correspondence with the notions of duality for cones and polarity for sets as defined in Definition 2.1.5. Duals of convex cones as defined in Definition 1.0.1 fit nicely with our Definition 2.1.1. Please see Theorem 2.1.9. If we were to use the definition given by (2.1) instead, we would be forced to use many awkwardly placed minus signs or polars of cones.

Let us derive the convex conjugates for some important convex functions. We will explicitly work out the supremum for each given y by using the first order optimality condition of the derivative being zero.

1. We first consider the function $f : \mathbb{R} \to \mathbb{R}_+$, where $f(x) := \frac{1}{2}x^2$.

   For any given real value y, the derivative of the expression $-\langle y, x \rangle - f(x)$ in terms of x is $-y - \partial f(x)$. To find the x such that the expression achieves the maximum, we use the first order optimality condition that $-y - \nabla f(x) = 0$, which is $-y - x = 0$. Solving this equality, we have $x = -y$. We plug this x value back into $-\langle y, x \rangle - f(x)$ and obtain
   $$\sup\{-\langle y, x \rangle - f(x) : x \in \mathbb{R}\} = \frac{1}{2}y^2.$$

   Hence, the convex conjugate function $f_*(y) = \frac{1}{2}y^2$. We have $f = f_*$, and such functions are called *self-dual*.

2. The second function we consider is $f : \mathbb{R}^n_{++} \to \mathbb{R}$, where $f(x) := -\sum_{j=1}^n \ln(x_j)$.

   This function is defined on the interior of the nonnegative orthant, and it has the property that as a sequence converging to a point on the boundary of the nonnegative orthant, the function value goes to infinity. We will later see that this is a "self-concordant barrier function" for the nonnegative orthant which encodes relevant information about the cone.

For any given point y, the gradient of the expression $-\langle y, x \rangle - f(x)$ in terms of x is $-y - \nabla f(x)$. To find the x such that the expression achieves the maximum, we still use the first order optimality condition that

$$-y - \nabla f(x) = 0,$$

which is $-y + (\frac{1}{x_1}, ..., \frac{1}{x_n})^T = 0$. Solving this equality, we have $x_j = \frac{1}{y_j}$ for $j = 1, ..., n$. We plug this x value back into $-\langle y, x \rangle - f(x)$ and obtain

$$\sup\{-\langle y, x \rangle - f(x) : x \in \mathbb{R}^n_{++}\} = -n - \sum_{j=1}^{n} \ln(y_j).$$

Hence, the convex conjugate function $f_*(y) = -\sum_{j=1}^{n} \ln(y_j) - n$.

3. The third function we consider is $f : \mathbb{R}^n_{++} \to \mathbb{R}$, where $f(x) := \sum_{j=1}^{n} x_j \ln(x_j)$. This function is also defined on the interior of the nonnegative orthant. It is called the entropy function, which is widely used in mathematical sciences and engineering.

For any given point y, the gradient of the expression $-\langle y, x \rangle - f(x)$ in terms of x is $-y - \nabla f(x)$. To find the x such that the expression achieves the maximum, we still use the first order optimality condition that

$$-y - \nabla f(x) = 0,$$

which is: $-y - (\ln(x_1) + 1, ..., \ln(x_n) + 1)^T = 0$. Solving this equality, we have: $x_j = e^{-y_j - 1}$ for $j = 1, ..., n$. We plug this x value back into $-\langle y, x \rangle - f(x)$ and obtain

$$\sup\{-\langle y, x \rangle - f(x) : x \in \mathbb{R}^n_{++}\} = \sum_{j=1}^{n} e^{-y_j - 1}.$$

Hence, the convex conjugate function $f_*(y) = \sum_{j=1}^{n} e^{-y_j - 1}$.

4. The last function we consider is $f : \mathbb{S}^n_{++} \to \mathbb{R}$, where $f(X) := -\ln \det(X)$. This function is defined on the interior of the positive semidefinite cone, and it has the property that as a sequence converging to a point on the boundary of the positive semidefinite cone, the function value goes to infinity. We will later see that this is a "self-concordant barrier function" for the positive semidefinite cone.

For any given point Y, the gradient of the expression $-\langle Y, X \rangle - f(X)$ in terms of X is $-Y - \nabla f(X)$. To find the X such that the expression achieves the maximum, we still use the first order optimality condition:

$$-Y - \nabla f(X) = 0,$$

16

which is $-Y + X^{-1} = 0$. Solving this equality, we have $X = Y^{-1}$. We plug this X value back into $-\langle Y, X \rangle - f(X)$ and obtain

$$\sup\{-\langle Y, X \rangle - f(X) : X \in \mathbb{S}^n_{++}\} = -\ln \det(Y) - n.$$

Hence, the convex conjugate function $f_*(Y) = -\ln \det(Y) - n$.

Note that the functions $f(x) = \frac{1}{2}x^2$, $f(x) = -\sum_{j=1}^n \ln(x_j)$ and $f(X) = -\ln \det(X)$ are all "self-dual" in the sense that the functions and their corresponding conjugates differ only by a constant.

Let us continue to define some notions and state some results related to convex sets and convex cones specifically. First, recall Definition 1.0.1 of the dual cone in Chapter 1, and we have the following proposition:

Proposition 2.1.2. *For every $\mathcal{K} \subseteq \mathbb{E}$, the dual cone $\mathcal{K}^*$ is a closed convex cone.*

Proof. By definition, the dual cone is the intersection of possibly infinitely many closed half spaces. Since the intersection of any collection of closed sets is still closed, and the intersection of any collection of convex sets is still convex, the dual cone must be closed and convex. $\qquad\square$

If upon identifying $\mathbb{E}$ with $\mathbb{E}^*$, there exists an inner product on $\mathbb{E}$ under which $\mathcal{K} = \mathcal{K}^*$, then $\mathcal{K}$ is called *self-dual*. $\mathbb{R}^n_+$, $\mathbb{S}^n_+$ and SOC^n are all self-dual cones under the usual Euclidean inner product, which in case of $\mathbb{S}^n_+$ coincides with the $\mathrm{Tr}\,(\cdot)$ inner product.

A cone is *homogeneous* if for every pair of points x and u in $\mathrm{int}(K)$, there is a linear bijection $f : \mathbb{E} \to \mathbb{E}$ with

$$f(K) = K \text{ and } f(x) = u.$$

Cones that are self-dual and homogeneous are called *symmetric*. $\mathbb{R}^n_+$, $\mathbb{S}^n_+$ and SOC^n are all homogeneous, hence are all symmetric cones.

A non-empty cone is *pointed* if the only linear subspace contained in it is the zero vector. The nonnegative orthant, second order cones and positive semidefinite cones are all pointed cones. The following propositions describe the correspondence between pointedness and the property of having non-empty interior for primal and dual cones (see for instance Boyd et al. [8]).

Proposition 2.1.3. *If the cone $\mathcal{K}$ has non-empty interior, then $\mathcal{K}^*$ is pointed.*

Proposition 2.1.4. *If the closure of the cone $\mathcal{K}$ is pointed, then $\mathcal{K}^*$ has non-empty interior.*

The notion of duality of the cones can be extended to general convex sets by "polarity", and we will see that the problems of finding the minimum volume ellipsoid containing a certain geometric object and the maximum volume ellipsoid contained in a certain object are related to each other via polarity.

Definition 2.1.5. *Given a set $G \subseteq \mathbb{E}$, the* polar set *of G is defined as:*

$$G^\circ := \{x \in \mathbb{E}^* : \langle x, u \rangle \le 1, \forall u \in G\}.$$

Another related notion for polarity is *gauge function*. In fact, the convex conjugate $\gamma^*(G, \cdot)$ of the gauge function $\gamma(G, \cdot)$ is the indicator function for G°.

Definition 2.1.6. *Given a set $G \subseteq \mathbb{E}$ (typically assume $0 \in \text{int}(G)$) and a point $u \in \mathbb{E}$, the* gauge function *is defined as:* $\gamma(G, u) = \inf\{\lambda \ge 0 : u \in \lambda G\}$.

By the definition of polar set, we can obtain the following proposition:

Proposition 2.1.7. *For every pair of closed convex sets $G_1, G_2 \subseteq \mathbb{E}$, both containing the origin, we have $G_1 \subseteq G_2 \iff G_2^\circ \subseteq G_1^\circ$.*

By this proposition, we see that the problem of finding the maximum volume ellipsoid contained in a convex body can be turned into a problem of finding the minimum volume ellipsoid containing a convex body and vice versa:

Let convex body G and the minimum volume ellipsoid $B_H(0, 1)$ containing G be given. Without loss of generality we may assume the ellipsoid is centered at the origin, as we can always translate the sets so that this happens. By the above proposition, the polar set G° must contain $B_H(0, 1)^\circ = B_{H^{-1}}(0, 1)$, which is also an ellipsoid. By equation (2.4), we know this ellipsoid must be the maximum volume ellipsoid contained in G°, otherwise it will contradict with the fact that $B_H(0, 1)$ is the minimum volume ellipsoid containing G. Also notice that

$$\text{vol}(B_H(0, 1)) \cdot \text{vol}([B_H(0, 1)]^\circ) = \text{vol}(B_H(0, 1)) \cdot \text{vol}(B_{H^{-1}}(0, 1))$$

is a constant for any given space $\mathbb{E}$.

There is a special class of convex functions called self-concordant barriers. A self-concordant barrier function intrinsically encodes a convex cone (or, in general, a convex set) by having real values in the interior of the cone and when a sequence converging to a point on the boundary the function value goes to infinity.

Definition 2.1.8. *(Self-concordance) Let $F : \mathrm{int}(\mathcal{K}) \to \mathbb{R}$ be a $\mathcal{C}^3$-smooth convex function such that F is a barrier for $\mathcal{K}$ (i.e. for every sequence in the interior of $\mathcal{K}$, converging to a boundary point of $\mathcal{K}$, the corresponding function values $F(x) \to +\infty$) and there exists $\vartheta \geq 1$ such that for each $t > 0$, $F(tx) = F(x) - \vartheta \ln(t)$, and*

$$|D^3 F(x)[h, h, h]| \leq 2 \left(D^2 F(x)[h, h] \right)^{3/2} \tag{2.2}$$

for all $x \in \mathrm{int}(\mathcal{K})$ and for all $h \in \mathbb{E}$. Then, F is called a ϑ-logarithmically homogeneous self-concordant barrier for $\mathcal{K}$.

As mentioned before, the notion of convex conjugacy is a notion of duality for functions. Convex conjugacy has a natural correspondence with the notions of duality for cones (see Rockafellar [41]). Below is the theorem by Nesterov and Nemirovskii [35], in the context of logarithmically homogeneous self-concordant barriers.

Theorem 2.1.9. *Let $\mathcal{K}$ be a closed convex pointed cone with a nonempty interior in $\mathbb{E}$ and let F be a ϑ-logarithmically homogeneous self-concordant barrier for $\mathcal{K}$. Then its Legendre-Fenchel conjugate*

$$F_*(y) = \sup\{-\langle y, x \rangle - F(x) : x \in \mathrm{int}(\mathcal{K})\},$$

is a ϑ-logarithmically homogeneous self-concordant barrier for the dual cone $\mathcal{K}^*$.

We refer to F_* simply as the *conjugate barrier*.

Once we have a ϑ-logarithmically homogeneous self-concordant barrier F for $\mathcal{K}$, at every point $x \in \mathrm{int}(\mathcal{K})$, the Hessian of F defines a local metric. For every $h \in \mathbb{E}$ the *local norm induced by F at x* is

$$\|h\|_x := \langle F''(x)h, h \rangle^{1/2}.$$

It is not hard to see from the definition of logarithmically homogeneous self-concordant barrier that $F''(x) : \mathbb{E} \to \mathbb{E}^*$ is self adjoint and positive definite, $[F''(x)]^{-1} : \mathbb{E}^* \to \mathbb{E}$ is well-defined and is also a self-adjoint positive definite operator. By Proposition 1.0.3, it makes sense to define

$$\|u\|_x^* := \left\langle u, [F''(x)]^{-1} u \right\rangle^{1/2}$$

for every $u \in \mathbb{E}^*$.

By Cauchy–Schwarz inequality, we have:

$$|\langle u, h \rangle| \leq \|u\|_x^* \|h\|_x, \quad \forall u \in \mathbb{E}^*, h \in \mathbb{E}.$$

Logarithmically homogeneous self-concordant barrier F and its conjugate barrier F_* interact very nicely because of the elegant and powerful analytic structure imposed by the convex conjugate function and the primal-dual symmetric conic set-up.

Suppose F is a ϑ-logarithmically homogeneous self-concordant barrier for the cone $\mathcal{K}$, and it further satisfies

- $F''(w)(K) \subseteq K^*$ for all $w \in \text{int}(K)$, and

- $F_*(F''(w)x) = F(x) - 2F(w) - \theta$ for all $w, x \in \text{int}(K)$,

then we call F a *self-scaled barrier* (see Nesterov and Todd [36], [37]). A cone that admits a self-scaled barrier is called *self-scaled*. It turns out that self-scaled cones are exactly symmetric cones.

For a symmetric cone K and a self-scaled barrier F, $F''(w)$ gives a linear isomorphism between $\text{int}(K)$ and $\text{int}(K^*)$, and $F''(w)\mathcal{K} = \mathcal{K}^*$ for every choice of $w \in \text{int}(K)$. (see for instance Nesterov and Todd [36], [37]).

2.2 Convex optimization problem formulations and optimality conditions

In this section, we will introduce several formulations of convex optimization problems, and some well-known optimality conditions for them.

The classical convex optimization problem formulation (P_1) is of the following form:

$$
\begin{aligned}
(P_1) \quad \inf : \ & f(x) \\
s.t. \quad & g_i(x) \leq 0, i \in \{1, ..., m\} \\
& h_j(x) = 0, j \in \{1, ..., l\} \\
& x \in \mathcal{C}.
\end{aligned}
$$

In the formulation, $\mathcal{C}$ is a convex set in $\mathbb{E}$, $f(x)$ and $g_i(x), i \in \{1, ..., m\}$ are convex functions on $\mathcal{C}$, and $h_j(x), j \in \{1, ..., l\}$ are affine functions on $\mathbb{E}$. We assume all the functions are continuously differentiable on $\mathcal{C}$.

For this formulation, we may get rid of the affine constraints $h_j(x) = 0, j \in \{1, ..., l\}$ by reformulating the original problem in a smaller dimensional space. We can achieve this

by eliminating variables using the affine constraints. The dimension of the new space is the same as that of the null space of the affine functions. Specifically, we can parametrize the affine space defined by the affine constraints by representing any point in the original space via the parametrization of the null space together with a translation. Then, we can reformulate (P_1) using this new representation of the space. We end up with a formulation with only inequality constraints. The transformed formulation will have the form as follows:

$$
\begin{aligned}
(P_2) \quad &\inf : f(x) \\
&s.t. \quad g_i(x) \leq 0, i \in \{1, ..., m\} \\
&\qquad x \in \mathcal{C}.
\end{aligned}
$$

However, notice as we had a transformation of the space, the functions f and g_i all have different representation in the new space now. The new $\mathcal{C}$ is the intersection of the affine space and the original $\mathcal{C}$, and the space $\mathbb{E}$ has a smaller dimension now. For the sake of brevity, we keep the same notation.

To provide suffcent conditions for optimality for a convex optimization problem, we use the notion of *Slater point* (a so-called *constraint qualification*). A point $x \in \mathbb{E}$ is a *Slater point* for problem (P_2) if it is in the interior of $\mathcal{C}$, and $g_i(x) < 0, \forall i \in \{1, ..., m\}$.

We can use the following well-known optimality conditions to characterize optimality for problems in the form of (P_2).

Theorem 2.2.1 (Karush, 1939 [27], Kuhn-Tucker, 1951 [29]). *Suppose the convex optimization problem (P_2) has a Slater point, then $x^* \in \mathbb{E}$ is a global minimum if and only if there exist some $s \in \mathbb{R}^m$ such that the following constraints are satisfied:*

1. *Stationarity: $-\partial f(x) = \sum_{i=1}^{m} s_i \partial g_i(x)$,*

2. *Primal feasibility: $g_i(x) \leq 0, i \in \{1, ..., m\}, x \in \mathcal{C}$,*

3. *Dual feasibility: $s_i \geq 0, i \in \{1, ..., m\}$,*

4. *Complementary slackness: $s_i g_i(x) = 0, i \in \{1, ..., m\}$.*

We will call these conditions *Karush-Kuhn-Tucker (KKT) conditions*, and the above theorem, *KKT Theorem*.

We may extend the KKT Theorem to handle affine equality constraints directly. The notion of *relative interior* is helpful to extend the theorem. The *relative interior* of a set $\mathcal{C}$ is defined as:
$$
\text{relint}(\mathcal{C}) := \{x \in \mathcal{C} : \exists \epsilon > 0, B(x, \epsilon) \cap \text{aff}(\mathcal{C}) \subseteq \mathcal{C}\},
$$

where $B(x, \epsilon)$ is a ball centered at x with radius ϵ, and $\mathrm{aff}(\mathcal{C})$ is the affine hull of $\mathcal{C}$ in Euclidean space $\mathbb{E}$, which is the smallest affine set containing $\mathcal{C}$.

A point $x \in \mathbb{E}$ is a *Slater point* for problem (P_1) if it is in $\mathrm{relint}(\mathcal{C} \cap J)$, where

$$J := \{x : h_j(x) = 0, j \in \{1, ..., l\}\} \text{ and } g_i(x) < 0, \forall i \in \{1, ..., m\}.$$

With this extended notion of "Slater point", we obtain the following KKT optimality conditions for the classical convex formulation (P_1).

Theorem 2.2.2. *Suppose the convex optimization problem (P_1) has a Slater point, then x^* is a global minimum if and only if there exist some s such that the following constraints are satisfied:*

1. *Stationarity:* $-\nabla f(x) = \sum_{i=1}^{m} s_i \nabla g_i(x) + \sum_{j=1}^{l} s_j \nabla h_j(x)$,

2. *Primal feasibility:* $g_i(x) \leq 0, i \in \{1, ..., m\}$, $h_j(x) = 0, j \in \{1, ..., l\}$, $x \in \mathcal{C}$,

3. *Dual feasibility:* $s_i \geq 0, i \in \{1, ..., m\}$,

4. *Complementary slackness:* $s_i g_i(x) = 0, i \in \{1, ..., m\}$.

Modern convex optimization also utilizes conic formulations of the problem, which possess many nice properties. We can formulate any convex optimization problem in the conic form as (P_3) below, under some mild assumptions.

$$\begin{aligned} (P_3) \quad \inf : \ &\langle c, x \rangle \\ &\mathcal{A}(x) = b \\ &x \in \mathcal{K} \end{aligned}$$

In this form, we are optimizing a linear function over a cone subject to an affine set constraints. We can transform the optimization problem (P_1) to the form of (P_3) by simply pushing the non-linear objective function into the constraints and then push the non-affine constraints into the cone constraint.

The dual problem (D_3) for (P_3) is defined as:

$$\begin{aligned} (D_3) \quad \sup : \ &\langle b, y \rangle_D \\ &s := c - \mathcal{A}^*(y) \in \mathcal{K}^*. \end{aligned}$$

$\langle \cdot, \cdot \rangle_D$ denotes the scalar product for the dual objective function on $(\mathbb{Y}, \mathbb{Y}^*)$, and $\mathcal{A}^* : \mathbb{Y} \to \mathbb{E}^*$ denotes the adjoint of the linear operator $\mathcal{A} : \mathbb{E} \to \mathbb{Y}^*$. $\mathcal{A}^*$ is defined by the equations:

$$\langle \mathcal{A}^*(y), x \rangle = \langle \mathcal{A}(x), y \rangle_D, \quad \forall\, x \in \mathbb{E}, \ y \in \mathbb{Y}.$$

We may assume without loss of generality that $\mathcal{A}$ is surjective, i.e., $\mathcal{A}(\mathbb{E}) = \mathbb{Y}^*$. $b \in \mathbb{Y}^*$, $c \in \mathbb{E}^*$ and $\mathcal{K} \subset \mathbb{E}$ is a pointed, closed, convex cone with nonempty interior. In modern theory interior-point methods, $\mathcal{K}$ is described via $F : \text{int}(\mathcal{K}) \to \mathbb{R}$, a *logarithmically homogeneous self-concordant barrier function for $\mathcal{K}$* as defined in Definition 2.1.8. $\mathcal{K}^*$ is the dual of the primal cone $\mathcal{K}$ with respect to $\langle \cdot, \cdot \rangle$.

We have the following optimality conditions for (P_3), which are analogous to the classical KKT conditions:

Theorem 2.2.3. *Suppose the convex optimization problem (P_3) has a Slater point, then $x^* \in \mathbb{E}$ is a global minimum if and only if there exist some $s \in \mathbb{E}^*$ such that the following constraints are satisfied:*

1. *Primal feasibility: $\mathcal{A}(x) = b$, $x \in \mathcal{K}$,*

2. *Dual feasibility: $c = \mathcal{A}^*(y) + s$, $s \in \mathcal{K}^*$,*

3. *Complementary slackness: $\langle s, x \rangle = 0$.*

Fritz John theorem provides another way to look at optimality conditions:

Theorem 2.2.4 (Fritz John, 1948 [25]). *Consider the optimization problem:*

$$(P) \quad \min : f(x)$$
$$g(x, y) \leq 0, \quad \forall y \in \mathcal{Y},$$

where f(x) is a continuously differentiable function defined on an open set $\mathcal{X} \subseteq \mathbb{R}^n$, and $g(x,y)$ and $\frac{\partial g}{\partial x}$ are continuous functions defined on $\mathcal{X} \times \mathcal{Y}$, where $\mathcal{Y}$ is a compact set in some Euclidean space. If x is a local minimizer of the above problem, then there exist at most n active constraints $g(x, y_1) = 0, ..., g(x, y_k) = 0$ and a non-trivial, non-negative multiplier vector $\lambda \in \mathbb{R}_+^{k+1} \setminus \{0\}$ such that

$$\lambda_0 \frac{\partial f(x)}{\partial x} + \sum_{i=1}^{k} \lambda_i \frac{\partial g(x, y_i)}{\partial x} = 0.$$

This theorem is in some sense more general than the KKT conditions, however with a weaker statement (since λ_0 may be zero).

Remark 2.2.5. An upper bound for the number of active constraints $g(x, y_i)$ needed in the equality $\lambda_0 \frac{\partial f(x)}{\partial x} + \sum_{i=1}^{k} \lambda_i \frac{\partial g(x, y_i)}{\partial x} = 0$ is at most the dimension of the variables x in the optimization problem (P).

2.3   Optimizing a linear function over an ellipsoid

Balls and ellipsoids are very simple geometric objects, and they are easy to deal with in general. As discussed in the introduction chapter, it would be helpful if we can approximate complex geometric objects using ellipsoids well. In the later chapters, we will consider the problems of approximating cones using ellipsoids centered at given points. Please see Figure 2.1 and 2.2 for illustration.

Recall the definitions of Euclidean balls and ellipsoids as follows:

Definition 2.3.1. *An* Euclidean ball *in Euclidean space* $\mathbb{E}$ *centered at* $\bar{x}$ *with radius* r *is* $B(\bar{x}, r) := \{x \in \mathbb{E} : \langle (x - \bar{x}), I(x - \bar{x}) \rangle \leq r^2\}.$

Definition 2.3.2. *An* ellipsoid *in Euclidean space* $\mathbb{E}$*, centered at* $\bar{x}$*, with radius* r *and its shape defined by a self-adjoint positive definite operator* $H : \mathbb{E} \to \mathbb{E}^*$ *is:*

$$B_H(\bar{x}, r) := \{x \in \mathbb{E} : \langle (x - \bar{x}), H(x - \bar{x}) \rangle \leq r^2\}.$$

In other words, $B_H(\bar{x}, r) = \{x \in \mathbb{E} : ||x - \bar{x}||_H \leq r\}.$

We notice that:

$$B_H(\bar{x}, 1) = \bar{x} + H^{-\frac{1}{2}} B(0, 1). \tag{2.3}$$

That is, every ellipsoid is an image of the Euclidean unit ball centered at the origin under an affine isomorphism. It is also easy to check by equation (2.3) that the volume of the above ellipsoid is:

$$\text{vol}(B_H(\bar{x}, 1)) = \sqrt{\det(H^{-1})}\, \text{vol}(B(0, 1)). \tag{2.4}$$

If we assume that the positive definite matrix H defining the ellipsoid has determinant equal to 1, the r in Definition 2.3.2 behaves just like the radii of Euclidean balls as far as volumes are concerned.
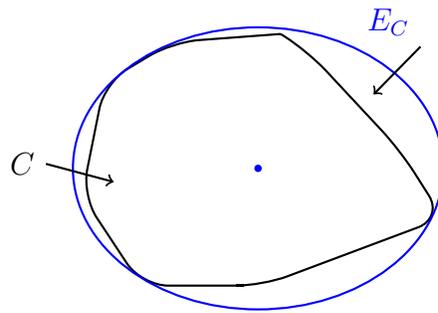
Figure 2.1: Using ellipsoid E_C as outer approximation for a convex set C.
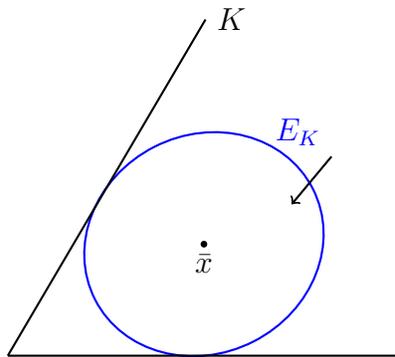
Figure 2.2: Using ellipsoid E_K centered at $\bar{x}$ as a local, inner approximation for cone K.

We will show it is very easy to optimize a linear function over an ellipsoid. Let us first show that it is easy to optimize a linear function over the unit Euclidean ball centered at the origin.

Proposition 2.3.3. *For every $u \in \mathbb{E} \setminus \{0\}$, the optimization problem:*

$$\max\{\langle u, x \rangle : x \in B(0,1)\}$$

has a unique maximizer $\frac{u}{||u||}$, and the optimal value is $||u||$.

Proof. By Cauchy-Schwarz inequality and the fact that u is a non-zero vector, we have for every $x \in B(0,1)$:

$$\langle u, x \rangle \leq ||u|| \cdot 1 = ||u|| \neq 0.$$

Moreover, equality is achieved if and only if x is some positive multiple of u. Thus, $x = \frac{u}{||u||}$ is the unique maximizer, with objective value equal to $||u||$. $\qquad\square$

We have the following equivalent result for a minimization problem as a Corollary:

Corollary 2.3.4. *For every $u \in \mathbb{E} \setminus \{0\}$, the optimization problem:*

$$\min\{\langle u, x \rangle : x \in B(0,1)\}$$

has a unique minimizer $-\frac{u}{||u||}$, and the optimal value is $-||u||$.

This result can be generalized to any arbitrary ellipsoid $B_H(\bar{x}, r)$. The following proposition provides a formula for minimizing a linear function over an ellipsoid:

Proposition 2.3.5. *For every $u \in \mathbb{E} \setminus \{0\}$, we have:*

$$\min\{\langle u, x \rangle : x \in B_H(\bar{x}, r)\} = \langle u, \bar{x} \rangle - r \cdot ||u||_H^* ,$$

and the unique minimizer is obtained at $x = \bar{x} - (r/ ||u||_H^*)H^{-1}u$.

Proof. By definition and elementary properties of inner product, we see the following three optimization problems are equivalent:

$$
\begin{aligned}
(P_1) \quad & \min\{\langle u, x \rangle : x \in B_H(\bar{x}, r)\}, \\
(P_2) \quad & \min\{\langle u, x \rangle : \langle (x - \bar{x}), H(x - \bar{x}) \rangle \leq r^2\}, \\
(P_3) \quad & \min\{\langle u, x \rangle : \left\langle \frac{H^{\frac{1}{2}}}{r}(x - \bar{x}), \frac{H^{\frac{1}{2}}}{r}(x - \bar{x}) \right\rangle \leq 1\}.
\end{aligned}
$$

Let $z := \frac{1}{r}H^{\frac{1}{2}}(x - \bar{x})$. Then, we have equivalently:

$$x = \bar{x} + rH^{-\frac{1}{2}}z. \tag{2.5}$$

Note that this is essentially observation (2.3). With this change of variable, the optimization problem becomes: $\min\{\left\langle u, \bar{x} + rH^{-\frac{1}{2}}z \right\rangle : \langle z, z \rangle \leq 1\}$, or equivalently:

$$(P_4) \quad \min\{\langle u, \bar{x} \rangle + \left\langle rH^{-\frac{1}{2}}u, z \right\rangle : z \in B(0,1)\}.$$

Notice that $\langle u, \bar{x} \rangle$ is a constant, and we may think of $rH^{-\frac{1}{2}}u$ as our new linear objective function. Thus, we can apply the previous proposition, and obtain the optimal value:

$$
\begin{aligned}
\langle u, \bar{x} \rangle - \left\| rH^{-\frac{1}{2}}u \right\| &= \langle u, \bar{x} \rangle - r \left\| H^{-\frac{1}{2}}u \right\| \\
&= \langle u, \bar{x} \rangle - r \left\langle u, H^{-1}u \right\rangle^{\frac{1}{2}} \\
&= \langle u, \bar{x} \rangle - r \left\| u \right\|_H^* .
\end{aligned}
$$

The unique optimizer of this problem is $z = -\dfrac{rH^{-\frac{1}{2}}u}{\left\| rH^{-\frac{1}{2}}u \right\|}$.

It is easy to see that z^* is optimal for (P_4) if and only if its corresponding x^*, by equality (2.5), is optimal for the original problem. Thus, by equation (2.5), we obtain the unique optimizer x of the original problem:

$$
\begin{aligned}
\bar{x} - rH^{-\frac{1}{2}}\frac{rH^{-\frac{1}{2}}u}{\left\| rH^{-\frac{1}{2}}u \right\|} &= \bar{x} - \frac{r}{\left\| H^{-\frac{1}{2}}u \right\|}H^{-1}u \\
&= \bar{x} - \frac{r}{\left\langle u, H^{-1}u \right\rangle^{\frac{1}{2}}}H^{-1}u \\
&= \bar{x} - \frac{r}{\left\| u \right\|_H^*}H^{-1}u.
\end{aligned}
$$

$\square$

We see that indeed optimizing a linear function over an ellipsoid is an easy problem. Proposition 2.3.5 provides a simple formula for its unique solution which is also easy to evaluate.

Chapter 3

Ellipsoids approximating convex sets, Löwner-John ellipsoid and Löwner's theorem

In this chapter we will introduce the concept of Löwner-John ellipsoid and Löwner's theorem, which are of central interest of this thesis. We will derive various optimality conditions for finding the minimum volume ellipsoid containing and the maximum volume ellipsoid contained in a convex set. For a detailed history and the development of the study on Löwner-John ellipsoids, please refer to Henk [23] and the references therein.

3.1 Motivation and Löwner-John theorem

As discussed in Chapter 1, it is useful to approximate complicated convex sets by simple ones such as ellipsoids. By Proposition 2.3.5 and its derivation, we see ellipsoids are indeed "simple enough" convex sets to deal with. On the other hand, by Löwner-John Theorem which will be introduced shortly afterwards, ellipsoids have some very nice properties in approximating compact convex sets. Hence, it is of general interest to be able to find a good approximation of a compact convex set using the kind of ellipsoids described in the Löwner-John Theorem.

Theorem 3.1.1 (Löwner-John Theorem). *Every compact convex set $C \in \mathbb{E}$ with $\mathrm{int}(C) \neq \emptyset$ has a unique minimum volume ellipsoid "E" (the Löwner-John ellipsoid) containing C.*

Moreover, contracting this ellipsoid by a factor of $\dim(\mathbb{E}) =: n$ *around its center yields an ellipsoid contained in* C*. That is, taking the center to be the origin,* $\frac{1}{n}E \subseteq C \subseteq E$*.*

The containment result is from John's 1948 paper [25], and the uniqueness part is largely believed to be discovered first by Löwner, but it seems that he never published this result. We note that for all compact convex C with $\text{int}(C) \neq \emptyset$, there exists also a unique maximal volume ellipsoid contained in C.

The bound in the theorem is actually tight. For instance, we can show that any unit equilateral simplex in $\mathbb{R}^n$ must have its Löwner-John ellipsoid being shrunk by a factor of $\frac{1}{n}$ to be contained in the original simplex:

Let us first embed the equilateral simplex, say $\text{conv}(\{v_0, \ldots, v_n\})$ into $\mathbb{R}^{n+1}$ by identifying each vertex of the simplex with one of the vector e_i in the standard basis of the vector space $\mathbb{R}^{n+1}$. The equilateral simplex and ellipsoids both possess lots of symmetries. Given the rich symmetry properties of the simplex around the centroid $c := \frac{1}{n+1} \sum_{i=0}^{n} v_i$ and the symmetries of the ellipsoids around the centers, it turns out that the centers of the minimum volume ellipsoid containing the simplex and the maximum volume ellipsoid contained in the simplex coincide with the centroid of the simplex. Again by the rich symmetries of the simplex, we can obtain that the minimum and maximum volume ellipsoids must take on the shape of a ball.

To contain the unit simplex, the radius of the ball must be at least the length from the origin to its farthest vertex. The distances from the origin to all vertices of the simplex are the same, and we denote it by d. By convexity, the ball containing all the vertices of the simplex must also contain the convex hull of the vertices, which is the simplex itself. Thus, the minimum volume ellipsoid containing the unit simplex is the ball centered at the centroid of the simplex with radius d. By our choice of the embedding of the simplex, we have the centroid is located at $\bar{c} := (\frac{1}{n+1}, \cdots, \frac{1}{n+1})^T$ in $\mathbb{R}^{n+1}$. Hence,

$$d = ||e_1 - \bar{c}|| = \sqrt{\frac{n}{n+1}}.$$

On the other hand, the minimum distance between the centroid and the boundary of the simplex is achieved on each facet (the maximal proper faces) of the simplex. Every facet of the simplex can be represented as the solutions of the following set of constraints for some l:

$$x_l = 0, \quad \sum_{i=1}^{n+1} x_i = 1, \quad x_i \geq 0, \forall i \in \{1, \cdots, n+1\}.$$
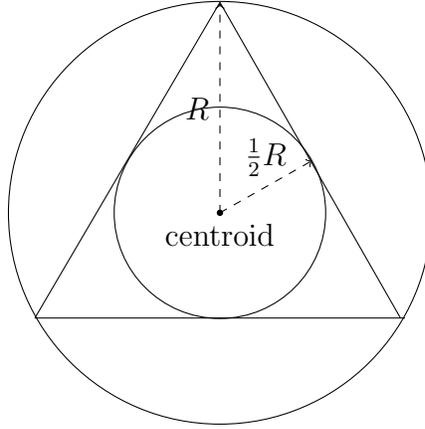
Figure 3.1: The minimum volume ellipsoid containing and the maximum volume ellipsoid contained in a equilateral simplex in $\mathbb{R}^2$.

The problem of finding the minimum distance between $\bar{c}$ and the facet where $x_{n+1} = 0$ can now be reduced to a convex optimization problem as follows:

$$\min : \sum_{i=1}^{n+1}(x_i - \frac{1}{n+1})^2$$
$$s.t. \quad x_{n+1} = 0$$
$$\sum_{i=1}^{n+1} x_i = 1$$
$$x_i \geq 0, \forall i \in \{1, \cdots, n+1\}.$$

Using the KKT optimality conditions, we obtain the optimal solution of this optimization problem is achieved at $x^* := (\frac{1}{n}, \cdots, \frac{1}{n}, 0)^T$. It is straightforward to check that the distance between $\bar{c}$ and x^* is $\sqrt{\frac{1}{n(n+1)}}$. As a result, we need to shrink the ball containing the simplex by a factor of $\frac{1}{n}$ to have it completely contained in the unit simplex.

For $n = 2$, please see Figure 3.1 for illustration. We see that the centers of both ellipsoids are located at the centroid of the simplex, and the radius of the maximum volume ellipsoid B_1 contained in the simplex is indeed $\frac{1}{2}$ of the radius of the minimum volume ellipsoid B_2 containing the simplex. We need to shrink B_2 by at least a factor of $\frac{1}{2}$ for it to be contained in the simplex.

The following propositions are straightforward to verify by the definitions, properties and the uniqueness of the Löwner-John ellipsoids of C.

Proposition 3.1.2. *If $A \subseteq B$ in $\mathbb{R}^n$, $B \subseteq B_H(\bar{x}, r)$, and the ellipsoid $B_H(\bar{x}, r)$ is the minimum volume ellipsoid containing A, then $B_H(\bar{x}, r)$ is the minimum volume ellipsoid containing B.*

Proposition 3.1.3. *If A, B are convex sets in $\mathbb{R}^n$, $A \subseteq B$, $B_H(\bar{x}, r) \subseteq A$, and the ellipsoid $B_H(\bar{x}, r)$ is the maximum volume ellipsoid contained in B, then $B_H(\bar{x}, r)$ is the maximum volume ellipsoid contained A.*

3.2 Optimality conditions for an ellipsoid being Löwner-John ellipsoid

In this section, we introduce some optimality conditions for proving an ellipsoid is indeed the Löwner-John ellipsoid for a certain set. First, we start with the problem of characterizing when the Euclidean ball centered at the origin is the minimum volume ellipsoid containing a finite set of points. Since ellipsoids are convex sets, an ellipsoid contains a finite set of points if and only if the ellipsoid contains the convex hull of these points. Therefore, the next characterization also applies to polytopes. Let us first state a lemma in matrix theory which will be of use later.

Lemma 3.2.1 (Schur Complement). *Let $X \in \mathbb{S}^n$ and $T \in \mathbb{S}^n_{++}$. Then*

$$M := \begin{pmatrix} T & U^T \\ U & X \end{pmatrix} \succeq 0 \iff X - UT^{-1}U^T \succeq 0.$$

Moreover, $M \succ 0$ if and only if $X - UT^{-1}U^T \succ 0$.

For details and the proof of the lemma please see for instance [48].

Proposition 3.2.2. *Let $p_1, ..., p_r \in \mathbb{E}$ all have unit norm. If there exists $y \in \mathbb{R}^r_{++}$, such that*

$$\sum_{i=1}^r y_i p_i p_i^T = I \text{ and } \sum_{i=1}^r y_i p_i = 0,$$

then the unit Euclidean ball centered at the origin is the minimum volume ellipsoid containing $\text{conv}\{p_i : i \in \{1, ..., r\}\}$.

Let $G := \text{conv}\{p_i : i \in \{1, ..., r\}\}$. Consider the following problem:

$$\min\{-\ln \det(H) : \langle (p_i - x), H(p_i - x) \rangle \leq 1, i \in \{1, \ldots, r\}, H \in \mathbb{S}^n_{++}, x \in \mathbb{R}^n\}. \qquad (3.1)$$

Optimization problem (3.1) is a non-convex problem with variables H and x. However, consider the following convex optimization problem:

$$\min\left\{-\ln \det(A) : \begin{pmatrix} 1 & p_i^T \end{pmatrix} \begin{pmatrix} \alpha & a^T \\ a & A \end{pmatrix} \begin{pmatrix} 1 \\ p_i \end{pmatrix} \leq 1, i \in \{1, ..., r\}, \begin{pmatrix} \alpha & a^T \\ a & A \end{pmatrix} \in \mathbb{S}^{n+1}_+\right\}. \qquad (3.2)$$

It is not hard to see (3.2) is equivalent to (3.1): if we expand the inequality constraint for the point p_i in (3.1) and (3.2) respectively, we obtain:

$$p_i^T H p_i - 2 p_i^T H x + x^T H x \leq 1, \text{ and}$$

$$p_i^T A p_i + 2 p_i^T a + \alpha \leq 1.$$

For a feasible solution pair (H, x) of (3.1), we see

$$\begin{pmatrix} x^T H x & -(Hx)^T \\ -Hx & H \end{pmatrix}$$

is a feasible solution for (3.2) with the same objective value. On the other hand, given any feasible solution

$$\begin{pmatrix} \alpha & a^T \\ a & A \end{pmatrix}$$

of (3.2), we can get a feasible pair $(A, -A^{-1}a)$ for (3.1) with the same objective value. Note that feasibility of $(A, -A^{-1}a)$ comes from the Schur Complement Lemma 3.2.1, which is for every $A \in \mathbb{S}^n_{++}$,

$$\begin{pmatrix} \alpha & a^T \\ a & A \end{pmatrix} \in \mathbb{S}^{n+1}_+ \iff \alpha \geq a^T A^{-1} a.$$

This implies $p_i^T A p_i + 2 p_i^T a + a^T A^{-1} a \leq p_i^T A p_i + 2 p_i^T a + \alpha$, where the right hand side is known to be less than or equal to one. Thus, (3.2) and (3.1) are essentially equivalent problems.

Moreover, by a property of the trace function: for compatible matrices U and V, $\text{Tr}(UV) = \text{Tr}(VU)$, (3.2) can be rewritten as the following optimization problem:

$$\min\left\{-\ln \det(A) : \text{Tr}\left(\begin{pmatrix} 1 & p_i^T \\ p_i & p_i p_i^T \end{pmatrix} \begin{pmatrix} \alpha & a^T \\ a & A \end{pmatrix}\right) \leq 1, i \in \{1, ..., r\}, \begin{pmatrix} \alpha & a^T \\ a & A \end{pmatrix} \in \mathbb{S}^{n+1}_+\right\}. (3.3)$$

Now, we may prove Proposition 3.2.2.

Proof. (of Proposition 3.2.2) By Löwner-John theorem, there exists a unique minimum volume ellipsoid containing $G = \text{conv}\{p_i : i \in \{1, ..., r\}\}$. To find the minimum volume ellipsoid, we formulate the optimization problem as follows:

$$\min\{-\ln\det(H) : \langle(p_i - x), H(p_i - x)\rangle \leq 1, \forall i \in \{1, ..., r\}\}.$$

By the above discussion, we formulate the equivalent minimization problem as:

$$(P') : \min\left\{-\ln\det(A) : \text{Tr}\left(\begin{pmatrix} 1 & p_i^T \\ p_i & p_i p_i^T \end{pmatrix}\begin{pmatrix} \alpha & a^T \\ a & A \end{pmatrix}\right) \leq 1, \forall i \in \{1, ..., r\}, \begin{pmatrix} \alpha & a^T \\ a & A \end{pmatrix} \in \mathbb{S}_+^{n+1}\right\}.$$

We notice the objective function is a convex function on (α, a, A) in $\mathbb{R} \oplus \mathbb{R}^n \oplus \mathbb{S}_{++}^n$, expressed as $\begin{pmatrix} \alpha & a^T \\ a & A \end{pmatrix}$. The constraints are linear inequalities.

This formulation is SDP representable, as it is easy to see this problem is equivalent to the problem of:

$$\min\left\{\det(A)^{-1} : \text{Tr}\left(\begin{pmatrix} 1 & p_i^T \\ p_i & p_i p_i^T \end{pmatrix}\begin{pmatrix} \alpha & a^T \\ a & A \end{pmatrix}\right) \leq 1, \forall i \in \{1, ..., r\}, \begin{pmatrix} \alpha & a^T \\ a & A \end{pmatrix} \in \mathbb{S}_+^{n+1}\right\},$$

and this latter problem is SDP representable. For details please see for instance Ben-Tal and Nemirovski [3], Proposition 4.2.1 and the examples after it. Thus, in principle, we can solve for the minimum volume ellipsoid by any algorithm that solves SDP problems.

Now, we consider a relaxation (P'') of (P'):

$$\min\left\{-\ln\det(A) : \text{Tr}\left(\begin{pmatrix} 1 & p_i^T \\ p_i & p_i p_i^T \end{pmatrix}\begin{pmatrix} \alpha & a^T \\ a & A \end{pmatrix}\right) \leq 1, \forall i \in \{1, ..., r\}, \begin{pmatrix} \alpha & a^T \\ a & A \end{pmatrix} \in \mathbb{S}_+^{n+1}\right\}.$$

Clearly, if we can prove $\begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix}$ is a minimizer of (P''), $\begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix}$ must be a minimizer of (P'). We can apply KKT Theorem to prove $\begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix}$ satisfies the KKT conditions and hence is a minimizer of (P'').

Firstly, we check the stationary condition and dual feasibility of the KKT conditions, which is:

$$\begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix} = \sum_{i=1}^r \lambda_i \begin{pmatrix} 1 & p_i^T \\ p_i & p_i p_i^T \end{pmatrix} - S, \text{ for } \lambda_i \geq 0, i \in \{1, ..., r\} \text{ and } S \in \mathbb{S}_+^{n+1}.$$

33

By the assumptions that $\sum_{i=1}^{r} y_i p_i p_i^T = I$ and $\sum_{i=1}^{r} y_i p_i = 0$, taking

$$\lambda_i = y_i, i \in \{1, ..., r\}, \text{ and } S := \begin{pmatrix} \sum_{i=1}^{r} y_i & 0 \\ 0 & 0 \end{pmatrix},$$

we have the conditions satisfied. Secondly, since all the p_i's have unit length, and $\begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix} \in$
$\mathbb{S}_{+}^{n+1}$, we obtain primal feasibility. Lastly, we check the complementary slackness condition: since all the p_i have unit length implies the linear inequalities are all tight, together with the fact that

$$S^T \begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix},$$

complementary slackness conditions are satisfied.

Hence, $\begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix}$ is a minimizer of (P'') and (P'). By the discussion before the proof, we have that the minimal ellipsoid containing G has the origin as its center and is defined by the identity matrix, which is exactly the Euclidean ball centered at the origin. $\quad\square$

Instead of using a semidefinite programming algorithm to solve the above given convex problem to find the minimum volume ellipsoid containing a polytope which is given as a convex hull of a set of points, in Kumar and Yildirim [30], a $(1 + \epsilon)$ approximation is proposed. The algorithm also produces a "core set" which acts as an approximation to the set of points given as input. This is very useful when the number of points involved is very large, and hence computationally costly. Later, Yildirim [49] extend this first order method to compute a $(1+\epsilon)$ approximation on the problem of finding the minimum volume ellipsoid containing a set of ellipsoids.

Another very efficient computational method for solving the minimum volume ellipsoid containing a given set of points in $\mathbb{R}^n$ is proposed in Sun and Freund [45]. This method is based on interior-point methods. Please refer to the paper and the references therein for more detailed treatment on this subject.

Proposition 3.2.2 can be generalized to cater to an arbitrary ellipsoid:

Proposition 3.2.3. *Given $p_1, ..., p_r \in \mathbb{E}$, if all the p_i satisfy $p_i^T H p_i = 1$ for some positive definite matrix H, and if there exists $y \in \mathbb{R}_{++}^r$, such that*

$$\sum_{i=1}^{r} y_i p_i p_i^T = H^{-1} \text{ and } \sum_{i=1}^{r} y_i p_i = 0,$$

then the ellipsoid $B_H(0,1)$ *is the minimum volume ellipsoid containing* $\mathrm{conv}\{p_i : i \in \{1,...,r\}\}$.

We can obtain this result by using the KKT Theorem as in the proof of the above proposition, but we can also prove it by applying a linear transformation $H^{\frac{1}{2}}$ to the whole space and reducing the problem to the Euclidean ball case:

The original space   The space after applying $H^{\frac{1}{2}}$
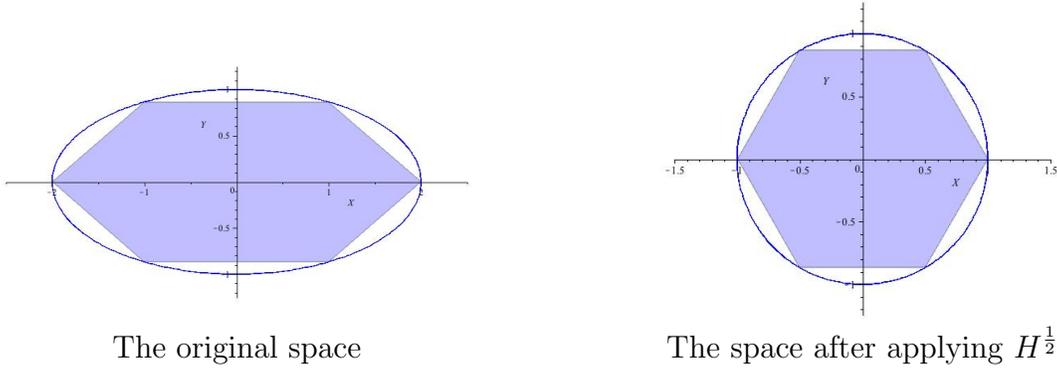
Figure 3.2: Illustration of the affine transformation.

Proof. (of Proposition 3.2.3) Let $G := \mathrm{conv}\{p_i : i \in \{1,...,r\}\}$. We apply a linear transformation $H^{\frac{1}{2}}$ to the whole space. Notice $B_H(0,1)$ becomes $B_I(0,1)$. Let $p_i' := H^{\frac{1}{2}}p_i$ for each $i \in \{1,...,r\}$. By the assumption of $p_i^T H p_i = 1$ for all $i \in \{1,...,r\}$, we have $p_i'^T p_i' = 1$ for all $i \in \{1,...,r\}$. Moreover, by the assumptions that

$$\sum_{i=1}^r y_i p_i p_i^T = H^{-1} \text{ and } \sum_{i=1}^r y_i p_i = 0,$$

we have

$$\sum_{i=1}^r y_i p_i' p_i'^T = I, \text{ and } \sum_{i=1}^r y_i p_i' = 0.$$

With the same set of y_i's, we apply the previous proposition, and get that the Euclidean ball $B_I(0,1)$ is the minimum volume ellipsoid containing $G = \mathrm{conv}\{p_i' : i \in \{1,...,r\}\}$.

Notice that:

$$\mathrm{conv}\{p_i' : i \in \{1,...,r\}\} = \mathrm{conv}\{H^{\frac{1}{2}}p_i : i \in \{1,...,r\}\}$$
$$= H^{\frac{1}{2}}(\mathrm{conv}\{p_i : i \in \{1,...,r\}\}) = H^{\frac{1}{2}}(G).$$

35

The original space   After translation by $\bar{x}$ After transformation of $H^{\frac{1}{2}}$
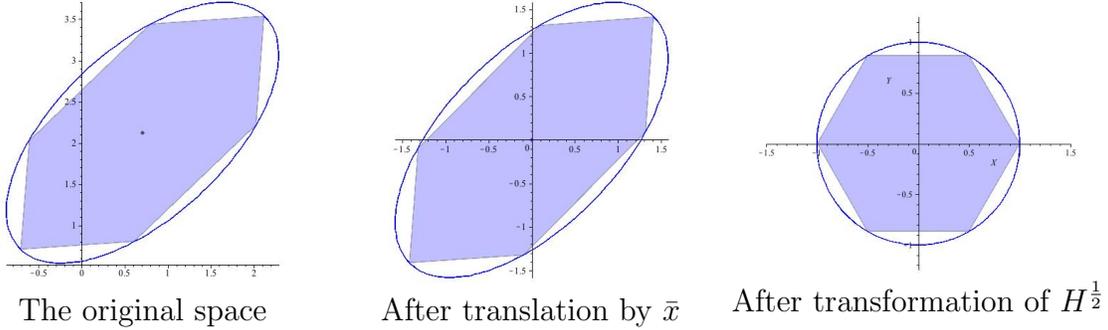
Figure 3.3: Illustration of the translation and affine transformation.

Hence, $B_I(0,1)$ is the minimum volume ellipsoid containing $H^{\frac{1}{2}}(G)$. After applying the inverse transformation $H^{-\frac{1}{2}}$ to the whole space and getting back to the original space, we have $B_H(0,1)$ is the minimum volume ellipsoid containing G. $\square$

We can future generalize this result by allowing the center of the minimum volume ellipsoid to be at any arbitrary point:

Proposition 3.2.4. *Given $p_1, ..., p_r \in \mathbb{E}$, if all the p_i satisfy $\langle p_i - \bar{x}, H(p_i - \bar{x}) \rangle = 1$ for some $\bar{x} \in \mathbb{E}$, H is a positive definite matrix, and if there exists $y \in \mathbb{R}^r_{++}$, such that*

$$H^{-1} = \sum_{i=1}^{r} y_i (p_i - \bar{x})(p_i - \bar{x})^T \ and \ \sum_{i=1}^{r} y_i (p_i - \bar{x}) = 0,$$

then the ellipsoid $B_H(\bar{x}, 1)$ is the minimum volume ellipsoid containing $\mathrm{conv}\{p_i : i \in \{1, ..., r\}\}$.

Notice that the ellipsoid $B_H(\bar{x}, 1)$ is obtained by first applying a linear transformation $H^{-\frac{1}{2}}$ to $B_I(0,1)$, and then applying a translation of $\bar{x}$ to it. Thus, we can do two steps of affine transformations to reduce the problem to the case of an Euclidean ball centered at the origin.

Proof. (of Proposition 3.2.4) Let $G := \mathrm{conv}\{p_i : i \in \{1, ..., r\}\}$, and $p'_i := p_i - \bar{x}$.
Step one: we will first show that all the p'_i satisfy the conditions as in the previous proposition, i.e.,

$$\sum_{i=1}^{r} y_i p'_i p'^T_i = H^{-1}, \ \text{and} \ \sum_{i=1}^{r} y_i p'_i = 0.$$

36

These are straightforward by the assumptions of:

$$H^{-1} = \sum_{i=1}^{r} y_i(p_i - \bar{x})(p_i - \bar{x})^T \text{ and } \sum_{i=1}^{r} y_i(p_i - \bar{x}) = 0.$$

Thus, by Proposition 3.2.3, we have $B_H(0,1)$ is the minimum volume ellipsoid containing $G' = \text{conv}\{p'_i : i \in \{1, ..., r\}\}$.

Step two: since $p'_i = p_i - \bar{x}, \forall i \in \{1, ..., r\}$, $G = G' + \bar{x}$. As $B_H(0,1)$ is the minimum volume ellipsoid containing G', $B_H(\bar{x}, 1)$ is the minimum volume ellipsoid containing $G = \text{conv}\{p_i : i \in \{1, ..., r\}\}$. $\qquad\square$

This result can be extended to general convex bodies by utilizing Fritz John theorem as in Chapter 2. The following characterization of the minimum volume ellipsoids is available in John [25]. To provide a certificate for an ellipsoid to be the maximum or minimum volume ellipsoid for a set, we may use the notion of "contact points". A *contact point* of two sets is a point that belongs to the boundaries of both sets.

Theorem 3.2.5 (Fritz John, 1948 [25])**.** *Let $C \subset \mathbb{R}^n$ be a convex body and $C \subseteq B_H(\bar{x}, 1)$. Then the following statements are equivalent:*

1. *$B_H(\bar{x}, 1)$ is the minimum volume ellipsoid containing C.*

2. *There exist some $r \in \left\{n, n+1, \ldots, \frac{n(n+3)}{2}\right\}$, contact points $p_1, ..., p_r \in \text{bd}(C) \cap \text{bd}(B_H(\bar{x}, 1))$ and $y \in \mathbb{R}^r_{++}$ such that:*

$$H^{-1} = \sum_{i=1}^{r} y_i(p_i - \bar{x})(p_i - \bar{x})^T \text{ and } \sum_{i=1}^{r} y_i(p_i - \bar{x}) = 0.$$

By Remark 2.2.5, the upper bound for the number of contact points needed is the dimension of the variable space, which is $\frac{n(n+1)}{2} + n = \frac{n(n+3)}{2}$ here. Moreover, it was shown in Gruber [16] that for "most" convex bodies in $\mathbb{R}^n$, the numbers of contact points between the convex body and their Löwner-John ellipsoids are exactly $\frac{n(n+3)}{2}$.

Proof. (of Theorem 3.2.5) Applying the Fritz John Theorem 2.2.4 to the optimization problem (3.1), with the indexing set being the boundary points of C, we know there exists

a non-zero vector $(y_0, ..., y_r) \geq 0$, where $r \leq \frac{n(n+3)}{2}$, $y_i > 0$ for $i > 0$, and $p_1, ..., p_r$ are in $\mathrm{bd}(C) \cap \mathrm{bd}(B_H(\bar{x}, 1))$, such that

$$H \sum_{i=1}^{r} y_i(p_i - \bar{x}) = 0 \text{ and } y_0 H^{-1} = \sum_{i=1}^{r} y_i(p_i - \bar{x})(p_i - \bar{x})^T$$

hold. Since H is positive definite, the first equality implies

$$\sum_{i=1}^{r} y_i(p_i - \bar{x}) = 0.$$

Now, suppose $y_0 = 0$, then $0 = \mathrm{Tr}(\sum_{i=1}^{r} y_i(p_i - \bar{x})(p_i - \bar{x})^T) = \sum_{i=1}^{r} y_i \|p_i - \bar{x}\|^2$, which implies $y_i = 0$ for all i. This is a contradiction. Hence, without loss of generality, we may assume $y_0 = 1$. This gives the other equality in (2). We showed (1) implies (2).

The proof of (2) implies (1) can be done by reducing the problem to the polytope case. Notice $\mathrm{conv}(p_1, ..., p_r) \subseteq C$, so if an ellipsoid contains C, it contains $\mathrm{conv}(p_1, ..., p_r)$, and by Proposition 3.1.2, the minimum volume ellipsoid containing $\mathrm{conv}(p_1, ..., p_r)$ while also containing C must be the minimum volume ellipsoid containing C. $\qquad\square$

Please see for instance Güler and Gürtuna [18] for more details.

Here is a similar theorem for the maximum volume ellipsoids:

Theorem 3.2.6. *Let $C \subset \mathbb{R}^n$ be a convex body and $B_H(\bar{x}, 1) \subseteq C$. Then the following statements are equivalent:*

1. *$B_H(\bar{x}, 1)$ is the maximum volume ellipsoid contained in C.*

2. *There exist some $r \in \left\{n, n+1, \ldots, \frac{n(n+3)}{2}\right\}$, contact points $p_1, ..., p_r \in \mathrm{bd}(C) \cap \mathrm{bd}(B_H(\bar{x}, 1))$ and $y \in \mathbb{R}^r_{++}$ such that:*

$$H^{-1} = \sum_{i=1}^{r} y_i(p_i - \bar{x})(p_i - \bar{x})^T \text{ and } \sum_{i=1}^{r} y_i(p_i - \bar{x}) = 0.$$

This characterization can be extended to non-convex sets by taking the convex hulls of the original sets. Recall that the convex hull of a set G is the smallest convex set (in terms of set inclusion) that contains G. In other words, any other convex set containing G must also contain the convex hull of G. Since all ellipsoids are convex sets, if an ellipsoid

contains G, it must contain the convex hull of G. In this way, we reduce the problem back to the convex setting. Here is the corollary in this more general setting for the minimum volume ellipsoid problem. There is an equivalent version for the maximum volume ellipsoid problem as well.

Corollary 3.2.7. *Let $G \subset \mathbb{R}^n$ be a compact set with $\mathrm{aff}(G) = \mathbb{R}^n$, and $G \subseteq B_H(\bar{x}, 1)$. Then the following statements are equivalent:*

1. *$B_H(\bar{x}, 1)$ is the minimum volume ellipsoid containing G.*

2. *There exist some $r \in \left\{ n, n+1, \ldots, \frac{n(n+3)}{2} \right\}$, contact points $p_1, \ldots, p_r \in \mathrm{conv}(G) \cap \mathrm{bd}(B_H(\bar{x}, 1))$ and $y \in \mathbb{R}^r_{++}$ such that:*

$$H^{-1} = \sum_{i=1}^r y_i (p_i - \bar{x})(p_i - \bar{x})^T \ \text{and} \ \sum_{i=1}^r y_i (p_i - \bar{x}) = 0.$$

Here are some related problems and generalizations that have been studied for finding the extremal ellipsoids for compact sets:

In Gürtuna [20], an analysis on the dual optimization problem for finding the minimum volume ellipsoid containing a general compact set is provided. It turns out that a class of design problems in statistics can be precisely formulated as the dual of the minimum volume ellipsoid problem.

A numerical method is proposed in Kojima and Yamashita [28] based on finding "small" ellipsoids or elliptic cylinders (where the determinant of the martix defining the ellipsoid is 0) containing a semi-algebraic set (defined by a set of polynomial inequalities) in $\mathbb{R}^n$ with a prefixed ellipsoidal shape. It also has implications in polynomial optimization. Please refer to the paper for details.

A further non-convex generalization of the problem is discussed In Lasserre [31]. Instead of using an ellipsoid which corresponds to using a homogeneous degree 2 polynomial to approximate the given set, Lasserre considered using a homogeneous polynomial of even degree d to approximate a compact set. In this generalization, both the underlying set and the approximating set defined by the polynomial need not to be convex. This is clearly a non-convex generalization. It turns out this generalized problem also has an optimal solution, and there exists a characterization in terms of contact points similar to what was derived in this section.

Another generalization of this problem is using some other convex body G instead of just ellipsoids to approximate a given convex body C. In Gruber [17] the local extremum properties of the volume on the space of all affine images of G in C is discussed.

3.3 Some general observations on the maximum volume ellipsoids in convex cones

In this section, we introduce some general observations for the problem of finding the maximum volume ellipsoid centered at a specified point, and inscribed in a convex cone. We will discuss the usage of the idea of "line search" in this problem, and also show that this problem can be reduced to the problem of finding the maximum volume ellipsoid contained in a compact convex set.

Intuitively, we can think of the process of finding the maximum radius for an ellipsoid centered at a specified point, with a given ellipsoidal shape H and inscribed in a convex cone as the process of "exploring" the cone from a given center in all directions until we touch the boundaries of the cone. Specifically, we start from the specified center $\bar{x}$, and do a line search for every possible direction d, where $||d||_H = 1$, until one of the search directions first hits the boundary of the cone. Hence, we have the following trivial proposition:

Proposition 3.3.1. *Given an interior point $\bar{x}$ in a cone K, and a symmetric positive definite matrix $H \in \mathbb{S}^n_{++}$, with $\det(H) = 1$, the largest radius r of an ellipsoid centered at $\bar{x}$ with shape defined by H and is contained in the cone K is*

$$r := \min_{d:||d||_H = 1} \{\lambda_d : \bar{x} + \lambda_d \cdot d \in \mathrm{bd}(K)\}.$$

We notice by the above proposition, any ellipsoid $B_H(\bar{x}, r)$ is symmetric around its center $\bar{x}$. Specifically, given any point in the ellipsoid, if we take the point symmetric to it with respect to its center, the resulting point is still in the ellipsoid. Please see Figure 3.4 for illustration. Thus, if the given cone K is a pointed closed convex cone with $\mathrm{int}(K) \neq \emptyset$, for the above problem, we may restrict our attention to a compact convex subset G of K, where $G := K \cap K'$ and K' is the reflection of the cone K around the point $\bar{x}$. Please see Figure 3.5 for illustration.

It is not hard to see that for the same reason, the problem of finding the maximum volume ellipsoid centered at a specified point, and inscribed in a convex cone K can be reduced to the problem of finding the maximum volume ellipsoid contained in a compact convex set G. The proposition is as follows:

Proposition 3.3.2. *Let $K \subset \mathbb{R}^n$ be a pointed closed convex cone with $\mathrm{int}(K) \neq \emptyset$. Let $\bar{x} \in \mathrm{int}(K)$. Then, the maximum volume ellipsoid centered at $\bar{x}$ contained in K is exactly the maximum volume ellipsoid centered at $\bar{x}$ contained in $G := K \cap K'$, where K' is the reflection of the cone K around point $\bar{x}$.*
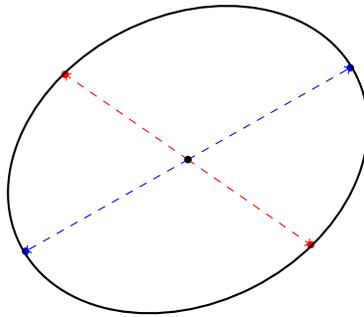
Figure 3.4: Illustration of Proposition 3.3.2 pointwise on an ellipsoid.
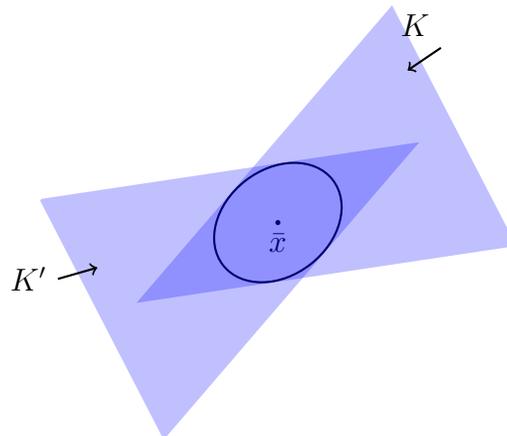
Figure 3.5: Illustration of Proposition 3.3.2.

Proof. First, we notice that the maximum volume ellipsoid contained in G must be centered at $\bar{x}$, since by construction, G is centrally symmetric about $\bar{x}$. Furthermore, by definition, we have $G \subseteq K$, and if we can show the maximum volume ellipsoid B_K centered at $\bar{x}$, contained in K is also contained in G, then by Proposition 3.1.3, we must have the ellipsoid B_K is the maximum volume ellipsoid contained in G. By the symmetric property of the ellipsoid around $\bar{x}$ and by construction of G, we must have that B_K is also contained in G. Hence, by Proposition 3.1.3 the result follows. $\qquad \square$

With this "compactification" method in mind, we can reduce the problem of finding the maximum volume ellipsoid centered at a particular point and inscribed in a convex cone to the problem of finding the maximum volume ellipsoid contained in a compact convex set. For the later problem, we are well equipped with theorems (on contact points and dual certificates as discussed in the previous section of this chapter) for proving optimality. Moreover, we will see in later chapters, the contact points we obtain between an ellipsoid and the compact set generated this way exactly correspond to the contact points of the ellipsoid and the cone together with its reflection.

Chapter 4

Maximum volume ellipsoids with specified centers contained in convex cones

In this chapter, we derive the maximum volume ellipsoids with specified centers contained in various convex cones. We will start from the cones of nonnegative orthant, second order cones and then move to positive semidefinite and homogeneous cones.

Our discussion will start with symmetries of ellipsoids and various convex cones. As we observed at the end of the last chapter: consider a given ellipsoid with center $\bar{x}$, if a line segment with an endpoint $\bar{x}$ is contained in the ellipsoid, then so is the reflection of the line segment around $\bar{x}$. Indeed, this kind of symmetry is expected, since we observed in the first chapter that every ellipsoid is linearly isomorphic to the Euclidean unit ball.

4.1   Background

In mathematics and mathematical sciences, symmetry plays many important roles. For example, characterization of symmetric structures in the problem at hand can help reduce the number of variables and parameters drastically and can lead to a deeper understanding of underlying key structures. Conversely, understanding of a core structure in low dimensions may sometimes be lifted to arbitrarily high dimensions by use of those characterized symmetries. Automorphisms provide a way of formalizing some symmetries in mathematical objects. An automorphism is an isomorphism that maps a mathematical object to itself

while preserving all the structures. If an object has rich symmetry, we can take advantage of it via its automorphism group.

Definition 4.1.1. *An automorphism g of a convex cone $\mathcal{K}$, is a invertible linear map, such that $g(\mathcal{K}) = \mathcal{K}$.*

The collection of all the automorphisms of a convex cone forms a group. It is easy to check:

- Identity exists: the group element that maps every point in the cone to itself acts as the identity element.

- Inverses exist: if $A_1(K) = K$, since A_1 is an invertible linear map, so is A_1^{-1} and $A_1^{-1}(K) = K$.

- Closed under composition: if $A_1(K) = K$, $A_2(K) = K$ and A_1, A_2 are invertible linear maps, then $A_1 A_2(K) = K$, $A_1 A_2(\cdot)$ is linear and invertible (with inverse $A_2^{-1} A_1^{-1}(\cdot)$).

- Associativity: compositions of linear maps are associative.

We denote the automorphism group of a cone K as $\mathrm{Aut}(K)$.

For instance,
$$S := \{G \cdot G^T : G \text{ is an invertible matrix in } \mathbb{R}^{n \times n}\}$$
is the automorphism group of the cone $\mathbb{S}_+^n$. Please see Gowda et al. [15] and the references therein for more details and related results. The linear function $G \cdot G^T$ acts on any given matrix $X \in \mathbb{S}_+^n$ by mapping X to GXG^T.

Note that for every invertible linear operator A, we have that the inverse operation and the adjoint operation commute. Therefore, we are justified in using the notation:

$$A^{-*} := ((A^{-1})^* = (A^*)^{-1}.$$

It is clear that any $G \cdot G^T \in S$ is an automorphism of $\mathbb{S}_+^n$:

1. $G \cdot G^T$ is clearly invertible with the inverse being $G^{-1} \cdot G^{-T}$.

2. Given any $X \in \mathbb{S}_+^n$, we have $GXG^T \in \mathbb{S}_+^n$, and for any $U \in \mathbb{S}_+^n$, $G^{-1}UG^{-T}$ will be mapped to U under $G \cdot G^T$. Hence, $G(\mathbb{S}_+^n)G^T = \mathbb{S}_+^n$.

Hence, we have $S \subseteq \mathrm{Aut}(\mathbb{S}^n_+)$.

Instead of working with the whole set of points a group is acting on, the complexity of a problem might be greatly reduced if we can reduce the problem to one single element of the set, with the other points being somehow "equivalent" to it. This can be achieved if the group acting on a set is "transitive".

The group *orbit* of an element x can be defined as $G(x) := \{gx \in X : g \in G\}$, which intuitively is all the places the element x can be moved to by the group G. A group G is *transitive*, if the group orbit is equal to the entire set X the group is acting on. Intuitively we may think of the transitive property as the group being able to move any given point in the set to any other point in the set. For instance, the group of all diagonal positive definite maps is transitive on the interior of the non-negative orthant, as given any arbitrary point $(x_1, \ldots, x_n)^T$ in the interior of the cone, it can be "moved to" $(1, \ldots, 1)^T$ by the map $\mathrm{Diag}(x_1^{-1}, \ldots, x_n^{-1})$.

A subset of $\mathrm{Aut}(\mathbb{S}^n_+)$ above is

$$\{G \cdot G : G \in \mathbb{S}^n_{++}\},$$

and this subset is transitive on the interior of the cone, i.e., it is transitive on $\mathbb{S}^n_{++}$.

Proposition 4.1.2. $\{G \cdot G : G \in \mathbb{S}^n_{++}\}$ *is a transitive subset of the automorphism group of* $\mathbb{S}^n_{++}$.

Proof. Given any two interior points X and Y in the cone of $\mathbb{S}^n_+$, we know $X, Y \in \mathbb{S}^n_{++}$. Thus, $X^{-\frac{1}{2}}$ and $Y^{\frac{1}{2}}$ exist and both are also in $\mathbb{S}^n_{++}$. Let $G_1 := X^{-\frac{1}{2}}$, clearly $G_1 X G_1 = I$. Let $G_2 := Y^{\frac{1}{2}}$, clearly $G_2 I G_2 = Y$. Hence, for any given X and Y in $\mathbb{S}^n_{++}$, there exists a group element $(G_2 G_1 \cdot G_1 G_2)$ as constructed above that will map X to Y. It follows that the orbit of any element $X \in \mathbb{S}^n_{++}$ is $\mathbb{S}^n_{++}$. By definition, this group is transitive.

This set is a subset of the automorphism group for $\mathbb{S}^n_{++}$ and $\mathbb{S}^n_+$ as for any matrix X in $\mathbb{S}^n_{++}$ or $\mathbb{S}^n_+$, the matrix GXG is still in $\mathbb{S}^n_{++}$ or $\mathbb{S}^n_+$ for $G \in \mathbb{S}^n_{++}$. On the other hand, for any point $Y \in \mathbb{S}^n_{++}$ or $\mathbb{S}^n_+$, the point $G^{-1} Y G^{-1}$ in the cone will be mapped to Y under $G \cdot G$. Hence, $G(\mathbb{S}^n_{++})G = \mathbb{S}^n_{++}$, and $G(\mathbb{S}^n_+)G = \mathbb{S}^n_+$. $\square$

The above proposition shows that the automorphism group of the positive semidefinite cone is rich. This implies that the cone possesses lots of symmetry. Later, we will take advantage of this symmetry and the rich automorphism group of homogeneous cones in general to reduce the complexity of our problem and to work on the problem of a very simple form. Specifically, we can reduce a problem of finding the maximum volume ellipsoid centered at an arbitrary point into the simple problem of finding the maximum volume

ellipsoid centered at identity. In this chapter, we will also see some powerful theorems making use of automorphism groups of various geometric objects to explore their intrinsic symmetries.

Proposition 4.1.3. *Let K be a homogeneous convex cone, and $\bar{x}$, $\bar{v}$ be a pair of interior points in K. An ellipsoid $B_H(\bar{x}, r)$ is the largest volume ellipsoid centered at $\bar{x}$ inscribed in K if and only if*

$$A(B_H(\bar{x}, r)) = B_{A^{-*}HA^{-1}}(A(\bar{x}), r)$$

is the largest volume ellipsoid centered at $\bar{v}$ inscribed in K, where $A \in \mathrm{Aut}(K)$, and $A(\bar{x}) = \bar{v}$.

Proof. Firstly, since K is a homogeneous cone, for any pair of interior points $\bar{x}$, $\bar{v}$ in K, there must exist some $A \in \mathrm{Aut}(K)$, such that $A(\bar{x}) = \bar{v}$.

We also notice that applying an automorphism $A \in \mathrm{Aut}(K)$ to any ellipsoid $B_H(\bar{x}, r)$ in K will result in another ellipsoid $B_{A^{-*}HA^{-1}}(A(\bar{x}), r)$ in K.

Since $A(\bar{x}) = \bar{v}$, the resulting ellipsoid is actually $B_{A^{-*}HA^{-1}}(\bar{v}, r)$. Moreover,

$$\mathrm{vol}(B_{A^{-*}HA^{-1}}(\bar{v}, r)) = \det(A) \cdot \mathrm{vol}(B_H(\bar{x}, r)). \tag{4.1}$$

$B_{A^{-*}HA^{-1}}(\bar{v}, r)$ is the largest volume ellipsoid centered at $\bar{v}$ inscribed in K. Suppose to the contrary that there exists another ellipsoid E having a larger volume, then by equation (4.1), $A(E)$ has a larger volume than $B_H(\bar{x}, r)$, which is a contradiction. The result follows. $\square$

Proposition 4.1.4. *For every convex cone K and $A \in \mathrm{Aut}(K)$, An ellipsoid $B_H(\bar{x}, r)$ is the largest volume ellipsoid centered at $\bar{x}$ inscribed in K if and only if $A(B_H(\bar{x}, r))$ is the largest volume ellipsoid centered at $A(\bar{x})$ inscribed in K.*

Proof. See the proof of Proposition 4.1.3. $\square$

4.2 Maximum volume ellipsoids in the nonnegative orthant

First, we state some propositions which will be helpful for approximating the cone $\mathbb{R}^n_+$ locally by large ellipsoids.

Proposition 4.2.1. *Given an interior point $\bar{x} \in \mathbb{R}_+^n$, and a matrix $H \in \mathbb{S}_{++}^n$ with $\det(H) = 1$, let*

$$\bar{r} := \min_{i \in \{1,...,n\}} \left\{ \frac{\bar{x}_i}{\sqrt{\langle e_i, H^{-1} e_i \rangle}} \right\},$$

then $B_H(\bar{x}, \bar{r})$ is the largest radius ellipsoid that has shape H and center $\bar{x}$ contained in the non-negative orthant.

Proof. We can find the largest radius r as described in the proposition by starting with a small enough value for the radius so that the ellipsoid is in $\mathbb{R}_+^n$, then "grow" the ellipsoid until it touches one of the faces of the non-negative orthant.
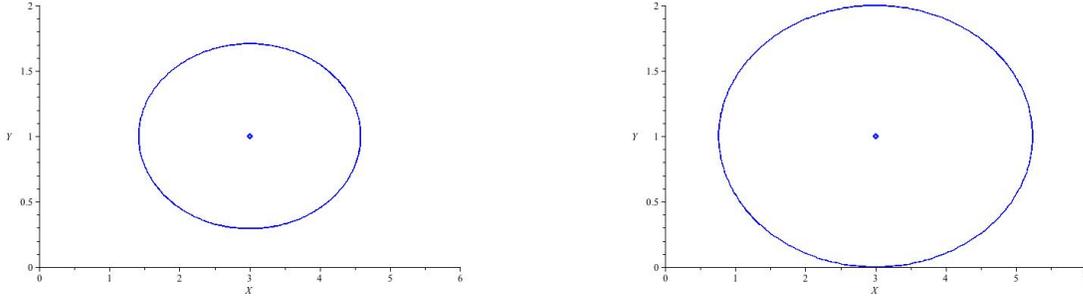
Figure 4.1: Growing the radius given a fixed center and ellipsoidal shape until it hits one of the boundaries of $\mathbb{R}_+^n$.

Notice the n maximal proper faces of the non-negative orthant are exactly:

$$\{x \geq 0 : \langle e_i, x \rangle = 0\}, i \in \{1, ..., n\}.$$

Moreover,

$$\mathbb{R}_+^n = \{x \in \mathbb{R}^n : \min\{\langle e_i, x \rangle, i \in \{1, ..., n\}\} \geq 0\},$$

and a point p is on the boundary of $\mathbb{R}_+^n$ if and only if:

$$\min\{\langle e_i, p \rangle, i \in \{1, ..., n\}\} \geq 0, \forall i \in \{1, ..., n\}, \text{ and } \exists i \in \{1, ..., n\} \text{ such that } \langle e_i, p \rangle = 0.$$

Thus, the problem becomes how much can we grow the radius r of the ellipsoid centered at $\bar{x}$ with shape H such that the minimization problems:

$$\min\{\langle e_i, x \rangle : \langle (x - \bar{x}), H(x - \bar{x}) \rangle \leq r^2\},$$

47

for each e_i, have non-negative optimal values and at least one of the optimal values is 0.

We may find such r by taking the minimum of all $\{r_1, ..., r_n\}$, where each r_i is precisely the radius that makes the optimal value of $\min\{\langle e_i, x \rangle : \langle (x - \bar{x}), H(x - \bar{x}) \rangle \leq r^2\}$ be 0.

We can find such r_i's easily by applying the previous proposition, and obtain

$$r_i = \bar{x}_i / \sqrt{\langle e_i, H^{-1} e_i \rangle},$$

or $r_i = \bar{x}_i / \sqrt{\langle q_i, D^{-1} q_i \rangle}$, where $QDQ^T = H$ is the spectral decomposition of H, and q_i's are the rows of matrix Q. $\qquad \square$

We notice that there are connections to the idea of "line search" and the notion of gauge, implicitly in the above proof.

Now, we derive the maximum volume ellipsoid contained in the non-negative orthant with a specified center. Let us first state a lemma that will be used in the proof of the proposition.

Lemma 4.2.2. *For every matrix $M \in \mathbb{S}^n_+$, $\frac{1}{n} \operatorname{Tr}(M) \geq \det(M)^{\frac{1}{n}}$. Moreover, if $M \in \mathbb{S}^n_{++}$, equality holds if and only if all eigenvalues of M are the same.*

Proof. Notice that for any n-by-n positive semidefinite matrix M, its determinant is the product of all its eigenvalues, and its trace equals to the sum of all its eigenvalues. Thus, the geometric mean of the eigenvalues equals to $\det(M)^{\frac{1}{n}}$, and the arithmetic mean of the eigenvalues equals to $\frac{1}{n} \operatorname{Tr}(M)$. Since the geometric mean of a set of non-negative numbers is smaller or equal to the arithmetic mean of them, and equality holds if and only if all the numbers are the same, the result follows. $\qquad \square$

Proposition 4.2.3. *Given an interior point $\bar{x}$ in the non-negative orthant of $\mathbb{R}^n$, the maximum volume ellipsoid centered at $\bar{x}$ such that the ellipsoid is contained in the non-negative orthant is*

$$\{x \in \mathbb{R}^n : \left\| \bar{X}^{-1}(x - \bar{x}) \right\| \leq 1\}, \text{ where } \bar{X} = \operatorname{Diag}(\bar{x}),$$

or

$$\left\{x \in \mathbb{R}^n : \langle (x - \bar{x}), X'(x - \bar{x}) \rangle \leq \frac{1}{\det(\bar{X}^{-1})^2}\right\}, \text{ where } X' := \frac{\bar{X}^{-2}}{\det(\bar{X}^{-1})^2}.$$

Proof. Applying Proposition 4.2.1, we see that all the r_i values are the same if we fix the center to be $\bar{x}$ and ellipsoidal shape to be as in the proposition statement. Hence, in this case, $r = r_i$ for all i. In order to have an ellipsoidal shape A ($\det(A) = 1$), with an even

48

bigger radius, we would require all the corresponding diagonal entries of A^{-1} to be smaller than the diagonal entries of X'^{-1}, by Proposition 4.2.1. This is impossible given that the determinants of X'^{-1} and A^{-1} are both one.

Suppose to the contrary, there exist some $A \in \mathbb{S}^n_{++}$ such that $\det(A^{-1}) = 1$ and all the corresponding diagonal entries of A^{-1} are smaller than that of X'^{-1}. If we apply $X'^{\frac{1}{2}} \cdot X'^{\frac{1}{2}}$ to both X'^{-1} and A^{-1}, they become I and $M := X'^{\frac{1}{2}} A^{-1} X'^{\frac{1}{2}}$ respectively. Clearly, I has the diagonal entries all equal to one, and easily followed by the assumption, all the diagonal entries of M are strictly smaller than one. This implies

$$\frac{1}{n} \operatorname{Tr}(M) < 1.$$

Since A^{-1}, X'^{-1} are positive definite matrices, so are I and M. By Lemma 4.2.2, together with the fact that $\frac{1}{n} \operatorname{Tr}(M) < 1$, we must have $\det(M)^{\frac{1}{n}}$ is strictly smaller than one. Hence, the determinant of M is smaller than one. This is a contradiction. The result follows. $\square$

We know that $F(x) = -\sum_{j=1}^n \ln(x_j)$ is a self-concordant barrier function for the non-negative orthant. The maximum volume ellipsoid derived in the above proposition, centered at $\bar{x}$ can also be represented as:

$$B_{r^2 F''(\bar{x})}(\bar{x}, r), \text{ where } r := \frac{1}{\sqrt{\det(F''(\bar{x}))}}.$$

4.3 Maximum volume ellipsoids in the 2-by-2 positive semidefinite cone

In this section, we focus on the cone

$$\mathbb{S}^2_+ = \left\{ \begin{pmatrix} x_1 & x_2 \\ x_2 & x_3 \end{pmatrix} : x_1, x_3 \geq 0, x_1 x_3 \geq x_2^2 \right\}.$$

Consider the linear isomorphism on $\mathbb{R}^3$:

$$A := \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{pmatrix}.$$

It is straightforward to check that $A \cdot \text{s2vec}(\mathbb{S}_+^2) = SOC^3$. We see that $\mathbb{S}_+^2$ is a special case of positive semidefinite cones and also a special case of second order cones. We will see the proofs for these generalizations in the later sections.

It was claimed in some papers in the literature that the largest ellipsoids contained in the positive semidefinite cones centering at identity have shape I. We will see in the later Section 4.5, Proposition 4.5.1 that this is true when we restrict the set of self-adjoint positive definite operators to be a subset of automorphisms of the cones: operators of the form "$X \cdot X$" for $X \in \mathbb{S}_+^n$. We will first show that this is indeed a misconception, and then find the maximum volume ellipsoid for $\mathbb{S}_+^2$ over all the self-adjoint positive definite operators.

Remark 4.3.1. The largest volume ellipsoid contained in the positive semidefinite cone $\mathbb{S}_+^2$ centered at identity is not the unit ball centered at I.

Proof. First, recall that $\mathbb{S}^2$ can be embedded in $\mathbb{R}^3$ by the mapping s2vec isometrically. We may reduce the problem of finding the largest ellipsoid centered at identity in $\mathbb{S}_+^2$ to finding the largest ellipsoid centered at $(1, 0, 1)^T$ in s2vec($\mathbb{S}_+^2$).

By Proposition 3.3.2, we can reduce the problem to finding the largest ellipsoid centered at $(1, 0, 1)^T$ in the convex body s2vec($\mathbb{S}_+^2$) $\cap$ s2vec($\mathbb{S}_+^2$)$'$, where s2vec($\mathbb{S}_+^2$)$'$ is the reflection of s2vec($\mathbb{S}_+^2$) about $(1, 0, 1)^T$. Moreover, as translations do not affect the optimal shape or the volume of the maximum volume ellipsoid, we may apply a translation of $(-1, 0, -1)^T$ to the whole space. Hence, the problem becomes finding the largest ellipsoid centered at the origin in the convex body

$$C := (-1, 0, -1)^T + \text{s2vec}(\mathbb{S}_+^2) \cap \text{s2vec}(\mathbb{S}_+^2)'.$$

By Proposition 3.2.6, $B(0, 1)$ is the maximum volume ellipsoid contained in C if and only if $B_I(0, 1) \subseteq C$, and there exist contact points $p_1, ..., p_r \in \text{bd}(C) \cap \text{bd}(B(0, 1))$, $r \geq n$, and $y \in \mathbb{R}_{++}^r$, such that $\sum_{i=1}^r y_i p_i p_i^T = I_n$ and $\sum_{i=1}^r y_i p_i = 0$. We will show it is impossible to obtain p_i's such that $\sum_{i=1}^r y_i p_i p_i^T = I_n$.

In the cone of $\mathbb{S}_+^2$ before translation, the contact points between the unit ball centered at I and $\mathbb{S}_+^2$ should be the matrices on the boundary of $\mathbb{S}_+^2$, i.e., rank one matrices. The contact points should also be on the boundary of the unit ball centered at I, and this implies $||P - I|| = 1$, where P is a contact point of $\mathbb{S}_+^2$ and the unit ball centered at I. Let $P =: Q^T D Q$ be the spectral decomposition of P, then

$$||P - I|| = ||Q^T D Q - I|| = ||Q^T (D - I) Q|| = ||D - I|| = 1.$$

Since D is a rank one diagonal 2-by-2 matrix, we must have the diagonal entries of D being 0 and 1. After the translation of the whole space by $-I$, the eigenvalues of the contact points are 0 and -1. Taking the reflection about the origin, we get the other half of the contact points of the unit ball centered at the origin and $(\mathbb{S}^2_+ - I) \cap (I - \mathbb{S}^2_+)$. Hence, every point on the boundary of C corresponds to a rank one, trace 1 or -1 matrix.

All rank one, trace 1 or -1 matrices in $\mathbb{S}^2$ can be represented as qq^T or $-qq^T$ where $q = (a, \pm\sqrt{1-a^2})^T, -1 \le a \le 1$. Hence, the corresponding contact points p's in C are

$$(a^2, \pm\sqrt{2}a\sqrt{1-a^2}, 1-a^2)^T, (-a^2, \mp\sqrt{2}a\sqrt{1-a^2}, a^2-1)^T, -1 \le a \le 1.$$

Now consider the tensors formed by those p's:

$$\begin{pmatrix} a^4 & \pm\sqrt{2}a^3\sqrt{1-a^2} & a^2(1-a^2) \\ \pm\sqrt{2}a^3\sqrt{1-a^2} & 2a^2(1-a^2) & \pm\sqrt{2}a(1-a^2)^{\frac{3}{2}} \\ a^2(1-a^2) & \pm\sqrt{2}a(1-a^2)^{\frac{3}{2}} & (a^2-1)^2 \end{pmatrix}, -1 \le a \le 1.$$

Clearly, the $(1,3)$ entry of the tensors must be non-negative. In order for any positive sum of the tensors to equal to identity, the $(1,3)$ entries of all the tensors in the positive sum must be 0. This implies a can only be $0, 1$ or -1. And this leaves only two possible tensors in the sum: $e_1 e_1^T$ and $e_3 e_3^T$. Clearly any positive sum of these two tensors cannot add up to I_3. Hence by Proposition 3.2.6, the unit ball centered at the origin is not the maximal volume ellipsoid contained in C.

By construction, C is centrally symmetric around 0, and thus the maximum volume ellipsoid contained in C is centered at 0. It follows that there must be another ellipsoid centered at 0 having a larger volume than the unit ball. Since we can not increase the radius of the unit ball, we know the maximum volume ellipsoid must have a different ellipsoidal shape. The result follows. $\qquad\square$

4.3.1 The maximum volume ellipsoid over all self-adjoint positive definite operators

In this sectiom, we derive the optimal ellipsoid centered at identity contained in $\mathbb{S}^2_+$ over all self-adjoint positive definite operators. We will find the largest ellipsoid in two different ways: the first proof uses geometric and algebraic arguments, and the second proof uses duality theory, contact points arguments and the underlying theory from Chapter 3.

It is tempting to use the exact same techniques as in the proof for the non-negative orthant case. However, by the following proposition, we know a direct analogue may not work in the case of positive semidefinite cones.

Proposition 4.3.2. *There does not exist an orthonormal set in $\mathbb{S}^n_+$ that spans $\mathbb{S}^n$, for $n \geq 2$. The maximum size of an orthonormal set in $\mathbb{S}^n_+$ is n instead of $\frac{n(n+1)}{2}$.*

Proof. Notice that the set of elements in $\mathbb{S}^n$ orthogonal to an element of rank r in $\mathbb{S}^n_+$ is isomorphic to $\mathbb{S}^{n-r}_+$. Thus, by assumption, if there is an element of rank r in the orthonormal set, the size of the set is at most $1 + n - r$. We see that to maximize the cardinality of the set, all the elements in the set must be rank one matrices. Hence, the maximum cardinality of an orthonormal set in $\mathbb{S}^n_+$ is n. $\qquad\square$

Now, let us develop some new techniques and derive the first proof which is geometric in nature. We will first prove that one of the axes of the optimal ellipsoid is along the ray from the origin through the identity.

Danzer et al. [12] has the following well-known theorem on automorphism groups of convex bodies. For more discussion on automorphism groups of convex bodies please see Güler and Gürtuna [19] [18].

Theorem 4.3.3 (Danzer et al. [12]). *Let C be a convex body in $\mathbb{R}^n$. For the maximum volume ellipsoid E_i inscribed in C and the minimum volume ellipsoid E_c containing C, we have their affine automorphism groups satisfy the following properties: $\mathrm{Aut}(C) \subseteq \mathrm{Aut}(E_i)$, $\mathrm{Aut}(C) \subseteq \mathrm{Aut}(E_c)$ and every element in $\mathrm{Aut}(C)$ fixes the centers of E_i and E_c.*

Using the above theorem, we obtain the following proposition:

Proposition 4.3.4. *The maximum volume ellipsoid centered at I inscribed in the cone $\mathbb{S}^2_+$ has one of its axes on the line going through the origin and I, and the other axes all have the same length.*

Proof. Let the convex body $C := \bar{C} \cap \bar{C}'$, where

$$\bar{C} := \{X \in \mathbb{S}^2_+ : \langle I, X \rangle \leq 2\},$$

and $\bar{C}'$ being the reflection of $\bar{C}$ about I. Clearly, all $Q \cdot Q^T$ must be in the automorphism group of C, with Q being any orthogonal matrix. We call this group of elements S. The geometric interpretation of these elements acting on C is the action of spinning around the line l from the origin to I. By the above automorphism group theorem, we know that every $Q \cdot Q^T$ fixes the center of the ellipsoid. The only matrices that satisfy this condition are λI, where $\lambda \in \mathbb{R}$, which is exactly the line l.

By the above theorem, the group of $\mathrm{Aut}(E_i)$, where E_i is the maximum volume ellipsoid contained in C, contains all elements in S as well. Consider any set that is non empty of the following form:

$$C_k := \{X \in E_i : \langle I, X \rangle = k\}, \ k \in \mathbb{R}.$$

Since the operation $Q \cdot Q^T$ preserves the inner product $\langle I, X \rangle$ for any $X \in \mathbb{S}^2_+$, together with the fact that $S \subseteq \mathrm{Aut}(E_i)$, we must have $Q(C_k)Q^T = C_k$ for any orthogonal Q. Hence, C_k is a ball or a point whenever it is not empty.

Let H be the matrix defining the shape of the ellipsoid E_i. The above statement implies that there are two eigenvectors of the matrix H defining the two axes of the ellipsoid on the hyperplane of $h := \{X \in \mathbb{S}^2_+ : \langle I, X \rangle = 0\}$, and their eigenvalues are the same. Clearly, the line of $\{\lambda I, \lambda \in \mathbb{R}\} = l$ is orthogonal to h. Hence, the third eigenvector orthogonal to both of the other two eigenvectors can only be located on the line l.

Moreover, by construction of C, we know C is centrally symmetric about I, so the maximum volume ellipsoid contained in C is centered at I. By similar arguments as before, we must have the maximum volume ellipsoid in C centered at I being the largest volume ellipsoid in $\mathbb{S}^2_+$ centered at I. The result follows. $\qquad \square$

Now, we can derive the maximum volume ellipsoid centered at I in the cone $\mathbb{S}^2_+$.

Proposition 4.3.5. *The maximum volume ellipsoid centered at I inscribed in the cone $\mathbb{S}^2_+$ is $B_H(I, 1)$, where H is a self-adjoint positive definite linear operator on $\mathbb{S}^2$ where $H(I) = \frac{3}{2}I$, and $H(M) = \frac{3}{4}M$ for any matrix M orthogonal to I.*

Notice that $\mathrm{vol}(B_H(I, 1)) = \sqrt{\frac{2^5}{3^3}}\mathrm{vol}(B)$, where $\mathrm{vol}(B)$ is the volume of the unit ball centered at I. Clearly, this ellipsoid has a larger volume than the one described in Remark 4.3.1.

Proof. By Proposition 4.3.4, we know one of the axes of the maximum volume ellipsoid centered at I and contained in $\mathbb{S}^2_+$ is along the line going through the origin and I.

Let us coordinate the space of $\mathbb{S}^2$ by the orthonormal basis $\{id, b_2, b_3\}$, where $id := \frac{I}{||I||}$.

Let $H := \begin{pmatrix} h_1 & h_4 & h_5 \\ h_4 & h_2 & h_6 \\ h_5 & h_6 & h_3 \end{pmatrix}, \det(H) = 1. \ E := \{X \in \mathbb{S}^2 : \langle X - I, H(X - I) \rangle \leq r^2\}.$ We want to maximize r, such that $E \subseteq \mathbb{S}^2_+$.

Since one of the axes of E is on the line going through the origin and I, we have $H(I) = \lambda_1 I$.

Using the above coordination, H becomes $\begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & h_2 & h_6 \\ 0 & h_6 & h_3 \end{pmatrix}$.

The other two axes of E must be in the subspace of $span\{b_2, b_3\}$. By Proposition 4.3.4, we know their eigenvalues are equal to each other. Hence, we may let the other two axes be b_2 and b_3. We get $H(b_2) = \lambda_2 b_2$ and $H(b_3) = \lambda_2 b_3$. Thus, $H = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_2 \end{pmatrix}$.

Let $\lambda_1 =: \lambda > 0$, and $\lambda_2 =: \lambda^{-\frac{1}{2}}$. Clearly, $\det(H) = \lambda_1 \lambda_2^2 = 1$, and

$$H = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda^{-\frac{1}{2}} & 0 \\ 0 & 0 & \lambda^{-\frac{1}{2}} \end{pmatrix}.$$

A point $X = (x_1, x_2, x_3)^T$, by this coordinate system is in E with the above H if and only if:
$$\lambda(x_1 - \sqrt{2})^2 + \lambda^{-\frac{1}{2}}(x_2^2 + x_3^2) \leq r^2. \tag{4.2}$$

We want a pair of λ and r with the largest r value (corresponding to the radius of E) such that any point X that satisfy the inequality above, i.e., every point in E, corresponds to a matrix in the cone of $\mathbb{S}_+^2$. We achieve this by finding a pair of λ and r such that every point X that satisfies (4.2) corresponds to a positive semidefinite matrix. Explicitly, we want to verify that every point in the Ellipsoid satisfies $\sqrt{x_2^2 + x_3^2} \leq x_1$ or

$$x_2^2 + x_3^2 \leq x_1^2 \text{ and } x_1 \geq 0. \tag{4.3}$$

We see that the above inequalities (4.3) is satisfied given X is in E if and only if

$$\lambda(x_1 - \sqrt{2})^2 + \lambda^{-\frac{1}{2}}x_1^2 \geq r^2 \text{ and } x_1 \geq 0. \tag{4.4}$$

Now, we focus on finding the best λ such that

$$\lambda(x_1 - \sqrt{2})^2 + \lambda^{-\frac{1}{2}}x_1^2 \geq r^2 \tag{4.5}$$

holds with the largest r.

Expanding this inequality, we get: $(\lambda + \lambda^{-\frac{1}{2}})x_1^2 - 2\sqrt{2}\lambda x_1 + 2\lambda \geq r^2$. The minimum of the left hand side is obtained at $x_1 = \frac{\sqrt{2}\lambda^{\frac{3}{2}}}{1+\lambda^{\frac{3}{2}}}$.

Plugging in $x_1 = \frac{\sqrt{2}\lambda^{\frac{3}{2}}}{1+\lambda^{\frac{3}{2}}}$ to the inequality, we get:

$$\left(\lambda + \lambda^{-\frac{1}{2}}\right)\left(\frac{\sqrt{2}\lambda^{\frac{3}{2}}}{1+\lambda^{\frac{3}{2}}}\right)^2 - 2\sqrt{2}\lambda\left(\frac{\sqrt{2}\lambda^{\frac{3}{2}}}{1+\lambda^{\frac{3}{2}}}\right) + 2\lambda \geq r^2.$$

After expanding and simplifying the inequality, we obtain:

$$\frac{2}{\sqrt{\lambda} + \frac{1}{\lambda}} \geq r^2.$$

The left hand side is a function with a unique maximum obtained at $\lambda = 4^{\frac{1}{3}}$. Hence, the largest possible r^2 is $\frac{2^{\frac{5}{3}}}{3}$.

We modify our H by multiplying H by a factor of $\frac{3}{2^{\frac{5}{3}}}$ so that the radius of the ellipsoid is 1. It is straightforward to check that the second condition $x_1 \geq 0$ in (4.4) is satisfied for this resulting ellipsoid. This ellipsoid has raduis $r = \frac{2^{\frac{5}{6}}}{\sqrt{3}}$ Thus, this ellipsoid is indeed the largest volume ellipsoid centered at I contained in $\mathbb{S}^2_+$. $\qquad\square$

Now we provide a second proof of this result by utilizing duality theory using contact points. We will use the same embedding of $\mathbb{S}^2$ in $\mathbb{R}^3$ by the isometry s2vec, as before. Hence, we can reduce the problem of finding the largest volume ellipsoid centered at identity in $\mathbb{S}^2_+$ to the problem of finding the largest volume ellipsoid centered at $(1,0,1)^T$ in $\mathbb{R}^3$.

Proposition 4.3.6. *After embedding the cone $\mathbb{S}^2_+$ in $\mathbb{R}^3$ by s2vec, the maximum volume ellipsoid centered at $\bar{x} := (1,0,1)^T$, the point in $\mathbb{R}^3$ corresponding to $I \in \mathbb{S}^2_+$, and inscribed in the cone s2vec$(\mathbb{S}^2_+)$ is $B_H(\bar{x}, 1)$, where*

$$H = \frac{3}{8}\begin{pmatrix} 3 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 3 \end{pmatrix}.$$

Before presenting the proof of the proposition, let us first verify that this ellipsoid is indeed contained in s2vec$(\mathbb{S}^2_+)$.

Proposition 4.3.7. *Given any point $x := (x_1, x_2, x_3)^T$ in $B_H(\bar{x}, 1) \subseteq \mathbb{R}^3$, where $\bar{x} := (1,0,1)^T$, and*

$$H = \frac{3}{8}\begin{pmatrix} 3 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 3 \end{pmatrix},$$

x must also be contained in s2vec$(\mathbb{S}^2_+)$, i.e., $x_1 \geq 0$, $x_3 \geq 0$ and $x_1 x_3 \geq \frac{x_2^2}{2}$.

Proof. Expanding the inequality $\langle (x - \bar{x}), H(x - \bar{x}) \rangle \leq 1$, corresponding to $x \in B_H(\bar{x}, 1)$, we have the polynomial inequality:

$$3(x_1 - 1)^2 + 2(x_1 - 1)(x_3 - 1) + 2x_2^2 + 3(x_3 - 1)^2 \leq \frac{8}{3}. \tag{4.6}$$

Let us do a change of variable by letting $A := x_1 - 1$, $B := x_3 - 1$ and $C = x_2$. The inequality (4.6) becomes:

$$3A^2 + 2AB + 2C^2 + 3B^2 \leq \frac{8}{3}, \text{ or } (A + B)^2 + 2A^2 + 2B^2 + 2C^2 \leq \frac{8}{3}. \tag{4.7}$$

We will first show $x_1, x_3 \geq 0$, when $x \in B_H(\bar{x}, 1)$. Consider the optimization problem of minimizing the left hand side of the equation (4.7), with the constraint $A \leq -1$. This is clearly a convex optimization problem, and applying the KKT Theorem 2.2.1, we verify that the minimum of the problem is achieved when $A = -1$ and $B = \frac{1}{3}$. The minimum value is $\frac{8}{3}$. When $A < -1$, the minimum value of the left hand side of the equation (4.7) is strictly larger than $\frac{8}{3}$, hence violating inequality (4.7). Thus, for $x \in B_H(\bar{x}, 1)$, we must have $A \geq -1$, i.e., $x_1 \geq 0$. Similarly, we can get $x_3 \geq 0$ when $x \in B_H(\bar{x}, 1)$.

Now, we will show $x_1 x_3 \geq \frac{x_2^2}{2}$, when $x \in B_H(\bar{x}, 1)$. Suppose to the contrary that $x_1 x_3 < \frac{x_2^2}{2}$, or $2(A + 1)(B + 1) < C^2$, then together with the inequality (4.7), we have

$$(A + B)^2 + 2A^2 + 2B^2 + 4(A + 1)(B + 1) < \frac{8}{3}.$$

If we denote $A + B =: M$, this inequality becomes:

$$3M^2 + 4M + 4 < \frac{8}{3}.$$

However, the minimum value of $3M^2 + 4M + 4$ for $M \in \mathbb{R}$ is $\frac{8}{3}$. This is a contradiction. Hence, we must have $x_1 x_3 \geq \frac{x_2^2}{2}$, when $x \in B_H(\bar{x}, 1)$. The result follows.

$\square$

Notice this proposition and the proof show the ellipsoid E as in Proposition 4.3.5 is contained in $\mathbb{S}_+^2$ directly. Now, we present the proof for Proposition 4.3.6:

Proof. (of Proposition 4.3.6) Let $C := \text{s2vec}(\mathbb{S}_+^2 \cap (\mathbb{S}_+^2)')$, where $(\mathbb{S}_+^2)'$ is the reflection of $\mathbb{S}_+^2$ around I. Let us try to find the possible contact points between C and the ellipsoid. Since any contact point $X \in \mathbb{S}_+^2$ between the ellipsoid and $\mathbb{S}_+^2$ must be on the boundary of

the cone, it must correspond to rank one matrix. Thus, the embedding $x := (x_1, x_2, x_3)^T$ of $X := \begin{pmatrix} x_1 & \frac{x_2}{\sqrt{2}} \\ \frac{x_2}{\sqrt{2}} & x_3 \end{pmatrix}$ in $\mathbb{R}^3$ must satisfy $x_1 x_3 = \frac{x_2^2}{2}$. Furthermore, we need the contact points to be on the boundary of the ellipsoid, i.e., it satisfies the equality:

$$\langle (x - \bar{x}), H(x - \bar{x}) \rangle = 1. \tag{4.8}$$

By a similar derivation as before, this implies $x_1 + x_3 - 2 = -\frac{2}{3}$, or $x_1 + x_3 = \frac{4}{3}$.

Using these two equalities, we can characterize all the possible contact points between $\text{s2vec}(\mathbb{S}_+^2)$ and the ellipsoid as:

$$\left(x_1, \pm\sqrt{\frac{8}{3}x_1 - 2x_1^2}, \frac{4}{3} - x_1 \right)^T, \quad \text{where } 0 \le x_1 \le \frac{4}{3}. \tag{4.9}$$

Now, we pick a set of contact points between C and the ellipsoid. We choose those with $x_1 = 0, \frac{4}{3}, \frac{2}{3}$ as in (4.9), and their reflection points around $\bar{x}$. We have the following $p_1, \dots, p_8$:

$$\begin{pmatrix} 0 \\ 0 \\ \frac{4}{3} \end{pmatrix}, \begin{pmatrix} \frac{4}{3} \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{2}{3} \\ \frac{2\sqrt{2}}{3} \\ \frac{2}{3} \end{pmatrix}, \begin{pmatrix} \frac{2}{3} \\ -\frac{2\sqrt{2}}{3} \\ \frac{2}{3} \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 2 \end{pmatrix}, \begin{pmatrix} \frac{2}{3} \\ 0 \\ 2 \end{pmatrix}, \begin{pmatrix} \frac{4}{3} \\ -\frac{2\sqrt{2}}{3} \\ \frac{4}{3} \end{pmatrix}, \begin{pmatrix} \frac{4}{3} \\ \frac{2\sqrt{2}}{3} \\ \frac{4}{3} \end{pmatrix}.$$

The corresponding tensor products are:

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \frac{16}{9} \end{pmatrix}, \begin{pmatrix} \frac{16}{9} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} \frac{4}{9} & \frac{4\sqrt{2}}{9} & \frac{4}{9} \\ \frac{4\sqrt{2}}{9} & \frac{8}{9} & \frac{4\sqrt{2}}{9} \\ \frac{4}{9} & \frac{4\sqrt{2}}{9} & \frac{4}{9} \end{pmatrix}, \begin{pmatrix} \frac{4}{9} & -\frac{4\sqrt{2}}{9} & \frac{4}{9} \\ -\frac{4\sqrt{2}}{9} & \frac{8}{9} & -\frac{4\sqrt{2}}{9} \\ \frac{4}{9} & -\frac{4\sqrt{2}}{9} & \frac{4}{9} \end{pmatrix},$$

$$\begin{pmatrix} 4 & 0 & \frac{4}{3} \\ 0 & 0 & 0 \\ \frac{4}{3} & 0 & \frac{4}{9} \end{pmatrix}, \begin{pmatrix} \frac{4}{9} & 0 & \frac{4}{3} \\ 0 & 0 & 0 \\ \frac{4}{3} & 0 & 4 \end{pmatrix}, \begin{pmatrix} \frac{16}{9} & -\frac{8\sqrt{2}}{9} & \frac{16}{9} \\ -\frac{8\sqrt{2}}{9} & \frac{8}{9} & -\frac{8\sqrt{2}}{9} \\ \frac{16}{9} & -\frac{8\sqrt{2}}{9} & \frac{16}{9} \end{pmatrix}, \begin{pmatrix} \frac{16}{9} & \frac{8\sqrt{2}}{9} & \frac{16}{9} \\ \frac{8\sqrt{2}}{9} & \frac{8}{9} & \frac{8\sqrt{2}}{9} \\ \frac{16}{9} & \frac{8\sqrt{2}}{9} & \frac{16}{9} \end{pmatrix}.$$

Now, let us take all the dual variables $y_1, \dots, y_8$ to be $\frac{3}{8}$, and it is easy to check that the two equalities

$$H^{-1} = \frac{1}{3} \begin{pmatrix} 3 & 0 & -1 \\ 0 & 4 & 0 \\ -1 & 0 & 3 \end{pmatrix} = \sum_{i=1}^{r} y_i (p_i - \bar{x})(p_i - \bar{x})^T \text{ and } \sum_{i=1}^{r} y_i (p_i - \bar{x}) = 0$$

are satisfied.

Hence, applying the Theorem 3.2.6, together with Proposition 4.3.7, we conclude that this ellipsoid is the maximum volume ellipsoid centered at $\bar{x}$ contained in C. By similar arguments as before, this ellipsoid is the largest volume ellipsoid in s2vec($\mathbb{S}^2_+$) centered at $\bar{x}$. The result follows. $\qquad\square$

Note that in the proof of Proposition 4.3.6, eight contact points (which is less than $\frac{n(n+3)}{2} = 9$) were sufficient.

By Proposition 4.1.3, and the fact that $\mathbb{S}^2_+$ is a homogeneous cone, we may reduce the problem of finding the largest volume ellipsoid centered at an arbitrary point to the problem of finding the largest volume ellipsoid centered at I inscribed in $\mathbb{S}^2_+$. We obtain this general result as a corollary.

Corollary 4.3.8. *The maximum volume ellipsoid centered at an interior point $\bar{X}$ inscribed in the cone $\mathbb{S}^2_+$ is $B_{(\bar{X}^{-\frac{1}{2}} \overset{s}{\otimes} \bar{X}^{-\frac{1}{2}}) H (\bar{X}^{-\frac{1}{2}} \overset{s}{\otimes} \bar{X}^{-\frac{1}{2}})}(\bar{X}, 1)$, where H is a self-adjoint positive definite operator on $\mathbb{S}^2$ where $H(I) = \frac{3}{2}I$, and $H(M) = \frac{3}{4}M$ for every matrix M orthogonal to I.*

Proof. By Proposition 4.3.5, 4.1.3 and Proposition 4.1.2 and its proof, the result follows immediately. $\qquad\square$

We know the positive semidefinite cone $\mathbb{S}^2_+$ is isometric to the second order cone in $\mathbb{R}^3$. Using similar techniques, we obtain the following results for second order cones in general.

4.4 Maximum volume ellipsoids in second order cones

In this section, we derive the maximum volume ellipsoid centered at a given point contained in general second order cone. This problem for the second order cones has been solved by Güler and Gürtuna [19]. In this section, we will use different techniques to derive the solutions. We will still do that in two ways: the first proof is geometric and algebraic in nature, and the second proof uses John's theorem based on duality theory and contact point technique. We will first present the geometric proof.

Proposition 4.4.1. *The maximum volume ellipsoid centered at $\bar{x} := e_1$ contained in second order cone $K := \left\{ \begin{pmatrix} t \\ x \end{pmatrix} \in \mathbb{R} \oplus \mathbb{R}^n : t \geq ||x||_2 \right\}$ is $B_H(\bar{x}, 1)$, where*

$$
H = \begin{pmatrix} n+1 & 0 & \cdots & 0 \\ 0 & \frac{n+1}{n} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{n+1}{n} \end{pmatrix}.
$$

Proof. Using similar techniques as in Proposition 4.3.4 for the cone of $\mathbb{S}^2_+$, we can derive that one of the axes of the maximum volume ellipsoid centered at e_1 is on the line l going through the origin and e_1 for second order cones.

This is due to again the rich symmetries of the second order cones about the line l, and it is captured mathematically by the automorphism group of the cones. Specifically, in this case, we may prove the statement analogously as in Proposition 4.3.4 by using the set of matrices acting on $\mathbb{R}^{n+1}$ who maps e_1 to itself and when restricted to the subspace of $span\{e_2, ..., e_n\}$, it is an orthogonal matrix. It is easy to check this set of matrices is a subset of the automorphism group of the cones. Again by similar techniques as before, we can prove that the other axes are on the lines of $\{\alpha e_2 : \alpha \in \mathbb{R}\}, \ldots,$ and $\{\alpha e_{n+1} : \alpha \in \mathbb{R}\}$ with the same length.

Note that any ellipsoid centered at $\bar{x}$ can be represented as:

$$
B_H(\bar{x}, r) = \left\{ \begin{pmatrix} t \\ x \end{pmatrix} \in \mathbb{R}^n : \left\langle \begin{pmatrix} t-1 \\ x \end{pmatrix}, H \begin{pmatrix} t-1 \\ x \end{pmatrix} \right\rangle \leq r^2 \right\}.
$$

If we fix $\det(H) = 1$, the problem becomes finding an ellipsoid with the greatest r and contained in the cone K.

By information on the axes and their lengths as described above, we obtain the following equalities for H: $He_1 = \frac{1}{\lambda^n} e_1$, $He_2 = \lambda e_2,...,$ and $He_{n+1} = \lambda e_{n+1}$, with $\lambda > 0$. Thus,

$$
H = \begin{pmatrix} \frac{1}{\lambda^n} & 0 & \cdots & 0 \\ 0 & \lambda & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda \end{pmatrix}.
$$

The inequality $\left\langle \begin{pmatrix} t-1 \\ x \end{pmatrix}, H \begin{pmatrix} t-1 \\ x \end{pmatrix} \right\rangle \leq r^2$ simply becomes $\frac{(t-1)^2}{\lambda^n} + \lambda ||x||_2^2 \leq r^2$, or

$$
\lambda^{n+1} ||x||_2^2 \leq \lambda^n r^2 - (t-1)^2. \tag{4.10}
$$

59

Hence, our question becomes finding a best λ such that the above inequality (4.10) implies

$$||x||_2 \leq t \tag{4.11}$$

with the largest r. We notice that (4.10) implies (4.11) if and only if

$$\lambda^n r^2 - (t-1)^2 \leq \lambda^{n+1} t^2 \text{ and } t \geq 0. \tag{4.12}$$

Now, we focus on finding the best λ and r pair such that

$$\lambda^n r^2 - (t-1)^2 \leq \lambda^{n+1} t^2 \tag{4.13}$$

holds with the largest r. We rewrite (4.13) as:

$$r^2 \leq \frac{(\lambda^{n+1}+1)t^2 - 2t + 1}{\lambda^n}.$$

The minimum of the right hand side is obtained at $t = \frac{1}{\lambda^{n+1}+1}$. Hence,

$$r^2 \leq \frac{\lambda}{\lambda^{n+1}+1}.$$

It is easy to check that the right hand side achieves its maximum at $\lambda = \sqrt[n+1]{\frac{1}{n}}$, and we get the maximum of r^2 being $\frac{n^{\frac{n}{n+1}}}{n+1}$.

Plugging in the value of λ and r^2, and multiplying both sides by $\frac{n+1}{n^{\frac{n}{n+1}}}$, we obtain our ellipsoid in the form as in the proposition above: $B_H(\bar{x}, 1)$, where

$$H = \begin{pmatrix} n+1 & 0 & \cdots & 0 \\ 0 & \frac{n+1}{n} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{n+1}{n} \end{pmatrix}.$$

It is straightforward to check that the second condition $t \geq 0$ in (4.12) is satisfied as well. Thus, this is indeed the maximum volume ellipsoid centered at $\bar{x}$ contained in the second order cone. $\qquad\square$

Now, let us use duality theory and contact points technique to provide a second proof for this result.

Proposition 4.4.2. *The maximum volume ellipsoid centered at* $\bar{x} := e_1$ *contained in second order cone* $K := \left\{ \begin{pmatrix} t \\ x \end{pmatrix} \in \mathbb{R} \oplus \mathbb{R}^n : t \geq ||x||_2 \right\}$ *is* $B_H(\bar{x}, 1)$, *where*

$$H = \begin{pmatrix} n+1 & 0 & \cdots & 0 \\ 0 & \frac{n+1}{n} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{n+1}{n} \end{pmatrix}.$$

Proof. Firstly, using similar arguments as in Proposition 4.4.1, we have that this ellipsoid is contained in the cone K. Let $C := K \cap K' - \bar{x}$, where K' is the reflection of K around $\bar{x}$. We can again reduce the problem to the problem of finding the maximum volume ellipsoid centered at the origin contained in C. By reducing the problem to the convex compact and centrally symmetric (around the origin) setting, we can apply John's theorem as in Theorem 3.2.6 to show optimality.

The contact points of the ellipsoid and the cone K are those points that satisfy (4.10) and (4.13) with equality. After some derivation, we get all contact points should satisfy $t = \frac{n}{n+1}$, and $||x||_2 = \frac{n}{n+1}$. Hence, the contact points of the ellipsoid and the cone K are in the form of:

$$\left(\frac{n}{n+1}, x_1, \ldots, x_{n-1}, \pm\sqrt{\frac{n^2}{(n+1)^2} - x_1^2 - \ldots - x_{n-1}^2} \right)^T, \tag{4.14}$$

where $-\frac{n}{n+1} \leq x_i \leq \frac{n}{n+1}, i \in \{1, \ldots, n-1\}$.

With the contact points of the ellipsoid and the cone K as above, we can pick the contact points of the ellipsoid and C by first picking the points as in (4.14) where exactly one $x_i = -\frac{n}{n+1}$ or $\frac{n}{n+1}$, for $i \in \{1, \ldots, n\}$, and all the other entries are 0, then preform a translation of $-e_1$. We also include their reflected counterparts around the origin.

Now, if we write out all the $4n$ contact points $p_1, \ldots, p_{4n}$, and their corresponding tensor products $p_1 p_1^T, \ldots, p_{4n} p_{4n}^T$, it is not hard to see that by taking all

$$y_i = \frac{n+1}{4n} > 0,$$

the two equality conditions

$$\sum_{i=1}^{r} y_i p_i p_i^T = H^{-1} \text{ and } \sum_{i=1}^{r} y_i p_i = 0$$

61

as in Theorem 3.2.6 hold. Thus, going back to the original cone setting using the same arguments as before, we conclude that the ellipsoid is indeed the maximum volume ellipsoid contained in the cone K centered at $\bar{x}$. $\qquad\square$

Corollary 4.4.3. *The volume of the maximum volume ellipsoid centered at $\bar{x} := e_1$ contained in second order cone $K := \left\{ \begin{pmatrix} t \\ x \end{pmatrix} \in \mathbb{R} \oplus \mathbb{R}^n : t \geq ||x||_2 \right\}$ is*

$$\frac{\operatorname{vol}(B_{n+1})}{\sqrt{(n+1)(1+\frac{1}{n})^n}},$$

with B_{n+1} being the unit Euclidean ball in $\mathbb{R}^{n+1}$. In particular, when n gets very large, the volume approaches $\frac{\operatorname{vol}(B_{n+1})}{\sqrt{(n+1)\mathbf{e}}}$.

Proof. The maximum volume ellipsoid $B_H(\bar{x}, 1)$ as in the last proposition has volume:

$$\det(H^{-\frac{1}{2}})\operatorname{vol}(B_{n+1}) = \frac{\operatorname{vol}(B_{n+1})}{\sqrt{(n+1)(1+\frac{1}{n})^n}}.$$

Moreover, as n approaches infinity, we notice that $(1+\frac{1}{n})^n$ goes to $\mathbf{e}$. $\qquad\square$

Since the second order cones are homogeneous as well, by Proposition 4.1.3, we obtain the general result of finding the largest volume ellipsoid inscribed in a second order cone centered at an arbitrary point as a corollary.

Corollary 4.4.4. *The maximum volume ellipsoid centered at an interior point x contained in second order cone $K := \left\{ \begin{pmatrix} t \\ x \end{pmatrix} \in \mathbb{R} \oplus \mathbb{R}^n : t \geq ||x||_2 \right\}$ is $B_{A^{-*}HA^{-1}}(x, 1)$, where $A : \mathbb{R}^{n+1} \to \mathbb{R}^{n+1}$, $A \in \operatorname{Aut}(SOC^n)$ such that $A(e_1) = x$ and*

$$H = \begin{pmatrix} n+1 & 0 & \cdots & 0 \\ 0 & \frac{n+1}{n} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{n+1}{n} \end{pmatrix}.$$

Proof. By Proposition 4.4.2 and 4.1.3, the result follows immediately. $\qquad\square$

This result can be easily adapted to find the maximum volume ellipsoid contained in any general second order cones $\hat{K}$ that are not canonical as K in Corollary 4.4.4, centered at an arbitrary point x. To achieve this, we can first find an invertible linear map L for which $L(K) = \hat{K}$, and then we find the maximum volume ellipsoid B contained in K centered at $K^{-1}(x)$. Clearly, $E := L(B)$ is the maximum volume ellipsoid contained in $\hat{K}$ and centered at x. Moreover, the contact points of B and K correspond to that of E and $\hat{K}$ by the linear map L.

A "dual version" of this problem: finding the minimum volume ellipsoid containing a truncated affine image of the second order cone:

$$C := \{(x, y) \in \mathbb{R}^{n+1} : ||B(x - c)|| \leq y, a \leq y \leq b\},$$

where B is an invertible matrix in $\mathbb{R}^{n \times n}$, $c \in \mathbb{R}^n$ and $0 \leq a < b$ are constants, is discussed in Güler and Gürtuna [19]. An explicit expression of the center and the positive definite form of the minimum volume ellipsoid is given as well. Thus, we can obtain yet another proof for the second order cone cases using their result:

First, we let $a = 0$ and pick $b > 0$ such that the center of the minimum volume ellipsoid containing the truncated second order cone with $0 \leq y \leq b$ is e_1. Using the formulas as in [19], we can find b that satisfy this requirement explicitly. We then apply a translation of $-e_1$ to the whole space so that e_1 is mapped to the origin. We can now find the the minimum volume ellipsoid E of the translated truncated cone C by the formulas in [19]. By construction, E is centered at the origin.

We take the polars of C and the ellipsoid E. Notice that the polar of a truncated second order cone (containing the origin) is still a truncated second order cone. By the discussion after Proposition 2.1.7, we know the polar of the ellipsoid E is the maximum volume ellipsoid contained in C°. Since E is centered at the origin by construction, so is E°. By Proposition 3.1.2, the symmetry properties of ellipsoids and the fact that E is the minimum volume ellipsoid containing C, E must also be the minimum volume ellipsoid containing $\text{conv}(C \cup C')$, where C' is the reflection of C around the origin. Hence, E° is also the maximum volume ellipsoid contained in $\bar{C} := C^\circ \cap C'^\circ$, which is the intersection of two truncated second order cones.

We can apply an affine transformation A to make $A(\bar{C})$ the intersection of the canonical second order cone and its reflection around a certain point p. By similar arguments as before, we know $A(E^\circ)$ is the maximum volume ellipsoid centered at p and contained in the second order cone. The result for an arbitrary center can be obtained with an additional automorphism as before.

4.5 Maximum volume ellipsoids in positive semidefinite cones

In this section, we will find the maximum volume ellipsoid in general positive semidefinite cones over the set of automorphism operators and also over all the self-adjoint positive definite operators. We will provide two proofs for the later problem. First proof is geometric and the second proof uses contact points and duality theory as in Chapter 3. These are original results of the thesis.

4.5.1 The maximum volume ellipsoids over the set of automorphism operators

Proposition 4.5.1. *Given an interior point $\bar{X}$ in the cone $\mathbb{S}^n_+$, the unique largest volume ellipsoid centered at $\bar{X}$ and with its shape defined by an operator in*

$$S = \{T \in \mathbb{L}(\mathbb{S}^n, \mathbb{S}^n) : T \succ 0, T(\cdot) = U \cdot U \text{ for some } U \in \mathbb{S}^n_{++}\}, \tag{4.15}$$

and being contained in $\mathbb{S}^n_+$ is:

$$\{X \in \mathbb{S}^n : \left|\left|\bar{X}^{-1/2}(X - \bar{X})\bar{X}^{-1/2}\right|\right| \leq 1\} \text{ or } \{X \in \mathbb{S}^n : \left\langle (X - \bar{X}), \bar{X}^{-1}(X - \bar{X})\bar{X}^{-1} \right\rangle \leq 1\}.$$

Proof. By Proposition 4.1.3, we know we can reduce the problem of finding the maximum volume ellipsoid centered at any interior point $\bar{X}$ contained in $\mathbb{S}^n_+$ to finding such an ellipsoid centered at I, as the cone $\mathbb{S}^n_+$ is homogeneous. The automorphism in S that will map $\bar{X}$ to I is $\bar{X}^{-\frac{1}{2}} \cdot \bar{X}^{-\frac{1}{2}}$. An ellipsoid $B_H(\bar{X}, r)$ is the largest volume ellipsoid centered at $\bar{X}$ inscribed in $\mathbb{S}^n_+$ if and only if $\bar{X}^{-\frac{1}{2}}(B_H(\bar{X}, r))\bar{X}^{-\frac{1}{2}}$ is the largest volume ellipsoid centered at I inscribed in $\mathbb{S}^n_+$.

Now, we only need to prove that the largest volume ellipsoid centered at I inscribed in $\mathbb{S}^n_+$ with the self-adjoint positive definite operator being in S is

$$B(I, 1) := \{X \in \mathbb{S}^n : ||X - I|| \leq 1\}.$$

It is straightforward to verify that if $X \in \mathbb{S}^n$ has a negative eigenvalue, $||X - I||$ will be strictly larger than 1. Hence, $B(I, 1)$ is contained in $\mathbb{S}^n_+$.

We are left to show $B(I, 1)$ is the largest ellipsoid inscribed in $\mathbb{S}^n_+$ centered at I under the assumption. Suppose to the contrary,

$$B_{V \cdot V}(I, r) = \{X \in \mathbb{S}^n : \left|\left|V^{-1/2}(X - \bar{X})V^{-1/2}\right|\right| \leq r\},$$

where $V \in \mathbb{S}_{++}^n \setminus \{I\}$, $\det(H) = 1$, and $r \geq 1$ is also inscribed in $\mathbb{S}_+^n$. Since $\det(V) = 1$, $\det(V^{-1}) = 1$. Since $V \neq I$, $V^{-1} \neq I$. Thus, V^{-1} must have some positive eigenvalue λ that is smaller than 1. Let X be in $\mathbb{S}^n$ such that $V^{-1/2}XV^{-1/2}$ has the same eigenvectors as V^{-1} and all the eigenvalues are the same as well except that we replace the eigenvalue of λ by $\lambda - r$. Since $\lambda < 1$ and $r \geq 1$, $\lambda - r < 0$. Clearly, this X is in the ellipsoid by construction, but it is not in $\mathbb{S}_{++}^n$. This is a contradiction.

Hence, the unique maximum volume ellipsoid centered at I inscribed in $\mathbb{S}_+^n$ with the self-adjoint positive definite operator being in S is

$$\{X \in \mathbb{S}^n : \left|\left|\bar{X}^{-1/2}(X - \bar{X})\bar{X}^{-1/2}\right|\right| \leq 1\} \text{ or } \{X \in \mathbb{S}^n : \left\langle (X - \bar{X}), \bar{X}^{-1}(X - \bar{X})\bar{X}^{-1} \right\rangle \leq 1\}.$$

$\square$

Later, in Section 5.4, we will see a similar result from Lim [32]. It shows that the minimum distance from an interior point A in a positive semidefinite cone to its boundary under the local norm defined by the positive definite matrix X, i.e., $||Y||_X = \left|\left|X^{-\frac{1}{2}}YX^{-\frac{1}{2}}\right|\right|$ is equal to the minimum eigenvalue of $X^{-1}A$.

This result exactly coincides with our observation above. When $A = I$, the minimum eigenvalue of $X^{-1}A$ is the minimum eigenvalue of X^{-1}. Since we assume X has determinant equal to 1, the determinant of X^{-1} is also equal to 1. Hence, the minimum eigenvalue of X^{-1} is 1, obtained only when X is the identity matrix.

4.5.2 The maximum volume ellipsoids over all self-adjoint positive definite operators by geometry

Now, we want to consider all the possible ellipsoidal shapes, that is, we will consider all the self-adjoint positive definite forms.

Let us first consider the problem of finding the largest radius with a given center and ellipsoidal shape so that the ellipsoid is contained in the cone.

Proposition 4.5.2. *Given an interior point $\bar{X} \in \mathbb{S}_{++}^n$, and a self-adjoint positive definite linear operator $H : \mathbb{S}^n \to \mathbb{S}^n$, with $\det(H) = 1$, the largest radius r of an ellipsoid centered at $\bar{X}$ with shape defined by H and is contained in $\mathbb{S}_+^n$ is:*

$$\min \left\{ \frac{\langle S, \bar{X} \rangle}{\sqrt{\langle S, H^{-1}(S) \rangle}} : \forall S \in \mathbb{S}_+^n, \text{rank}(S) = 1, \text{Tr}(S) = 1 \right\}. \tag{4.16}$$

Proof. We use similar techniques as in the non-negative orthant case. Notice now for positive semidefinite cone $\mathbb{S}^n_+$ we have infinitely many proper faces. We focus on the maximal proper faces defined by rank one matrices, i.e., the extreme rays of its dual cone.

We may think of $\mathbb{S}^n_+$ as defined by the extreme rays of its dual cone as follows:

$$\mathbb{S}^n_+ = \{X \in \mathbb{S}^n : \langle X, S \rangle \geq 0, \forall S \in \mathbb{S}^n_+, \operatorname{rank}(S) = 1, \operatorname{Tr}(S) = 1\}.$$

Moreover, a point P is on the boundary of $\mathbb{S}^n_+$ if and only if it satisfy the above condition and there exist some $S \in \mathbb{S}^n_+$ so that:

$$\operatorname{rank}(S) = 1, \ \operatorname{Tr}(S) = 1 \text{ and } \langle S, P \rangle = 0.$$

Thus, the problem now becomes how much can we grow the radius r of $B_H(\bar{X}, r)$ such that the minimization problems:

$$\min\{\langle X, S \rangle : \langle (X - \bar{X}), H(X - \bar{X}) \rangle \leq r^2\}$$

for each $S \in \mathbb{S}^n_+$ with $\operatorname{rank}(S) = 1$ and $\operatorname{Tr}(S) = 1$ have non-negative optimal values and at least one of the optimal values is 0.

Such r can be found by taking the minimum of all r_S's, for each S as above, where each r_S is the radius that makes the optimal value of the following optimization problem 0:

$$\min\{\langle S, X \rangle : \langle (X - I), H(X - I) \rangle \leq r^2\}.$$

After embedding $\mathbb{S}^n$ in $\mathbb{R}^{\frac{n(n+1)}{2}}$ by the isometry s2vec, we can use Proposition 2.3.5 to obtain the above result. □

For ellipsoidal shapes from the set (4.15), we can find the largest radius explicitly in terms of the minimum eigenvalue of an expression of the point $\bar{X}$ and the matrix U defining the ellipsoidal shape. We will see this result by Lim [32] in Section 5.4. Explicitly, given an interior point $\bar{X} \in \mathbb{S}^n_{++}$, and a self adjoint positive definite linear operator of the form $U \cdot U$ for some $U \in \mathbb{S}^n_{++}$, the largest radius r of the ellipsoid centered at $\bar{X}$ with the above shape and is contained in $\mathbb{S}^n_+$ is the minimum eigenvalue of $U\bar{X}$.

As discussed before, we know by Proposition 4.1.3, the problem of finding the maximum volume ellipsoid centered at an arbitrary interior point in the positive semidefinite cone can be reduced to the single case where the point is the identity. This is because any positive semidefinite cone is homogeneous. We now focus on the problem of finding the maximal volume ellipsoid centered at identity inscribed in the cone of $\mathbb{S}^n_+$ over all the self-adjoint positive definite forms.

Before presenting the first proof, we state a proposition exposing the geometric properties of the positive semidefinite cones:

Proposition 4.5.3. *The distance between I and any positive semidefinite matrix on the boundary of $\mathbb{S}^n_+$ with rank less than or equal to k and on the affine subspace of*

$$\{H \in \mathbb{S}^n : \langle I, H \rangle = n\}$$

is less than or equal to $\sqrt{n(n-1)}$ and greater than or equal to $\sqrt{\frac{n(n-k)}{k}}$. Moreover, they are obtained by rank one matrices and rank k matrices respectively.

Proof. For the matrix H as described above, we must have $\mathrm{Tr}(H) = n$, as it is in the affine subspace. The square of the norm in terms of the trace inner product between I and H is:

$$\|I - H\|^2 = \langle I - H, I - H \rangle = \mathrm{Tr}(I) + \mathrm{Tr}(H^2) - 2\,\mathrm{Tr}(H) = n - 2n + \mathrm{Tr}(H^2) = \mathrm{Tr}(H^2) - n.$$

Suppose $\lambda_1, ..., \lambda_n$ are the eigenvalues of H, then $\mathrm{Tr}(H^2) - n = (\sum_{i=1}^n \lambda_i^2) - n$.

To find the minimum and the maximum distance between identity and any positive semidefinite matrix with rank less or equal to k on the above affine subspace, we consider the following optimization problems:

$$\min\{(\sum_{i=1}^k \lambda_i^2) - n : \sum_{i=1}^k \lambda_i = n, \lambda_i \geq 0, \forall\{1, ..., k\}\},$$

$$\max\{(\sum_{i=1}^k \lambda_i^2) - n : \sum_{i=1}^k \lambda_i = n, \lambda_i \geq 0, \forall\{1, ..., k\}\}.$$

For the first minimization problem, we apply the KKT Theorem, and get the minimal value is obtained when $\lambda_1 = ... = \lambda_k = \frac{n}{k}$. This implies the minimum value is obtained at any matrix with k of the eigenvalues being $\frac{n}{k}$, and the rest being 0. It is straightforward to check the minimum distance in this case is $\sqrt{\frac{n(n-k)}{k}}$.

The second optimization problem is maximizing a convex function over a convex set. We know by the property of convex functions, the optimal value is achieved at an extreme point. In this case, clearly all the extreme points correspond to rank one matrices, so the maximum is achieved at some rank one matrices. However, all the rank one matrices clearly have the same objective value. Thus, any rank one matrix is optimal, and the maximum distance is $\sqrt{n(n-1)}$. $\qquad\square$

In particular, the distance between I and any positive semidefinite matrix on the boundary of $\mathbb{S}^n_+$ and on the affine subspace of $\{H \in \mathbb{S}^n : \langle I, H \rangle = n\}$ is less or equal to

$\sqrt{n(n-1)}$, and greater than or equal to $\sqrt{\frac{n}{n-1}}$. They are obtained by rank one matrices and rank $n-1$ matrices respectively.

With this observation, we can now derive the maximum volume ellipsoid contained in the general positive semidefinite cone centered at I. We will first compute the maximum volume ellipsoid in a suitable second order cone, which will be used later in the proof for the positive semidefinite cone. In fact, we will prove that this ellipsoid is also the maximum volume ellipsoid contained in $\mathbb{S}^n_+$ centered at I.

Proposition 4.5.4. *The maximum volume ellipsoid centered at $\bar{x} := (\sqrt{n}, 0, ..., 0)^T$ contained in the scaled second order cone $K := \left\{ \begin{pmatrix} t \\ x \end{pmatrix} \in \mathbb{R} \oplus \mathbb{R}^{N-1} : t \geq \sqrt{n-1} \, ||x||_2 \right\}$ is $B_H(\bar{x}, 1)$, where*

$$H = \begin{pmatrix} \frac{n+1}{2} & 0 & \cdots & 0 \\ 0 & \frac{n+1}{n+2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{n+1}{n+2} \end{pmatrix}.$$

Proof. Using similar techniques as in Proposition 4.3.4 for the cone of $\mathbb{S}^2_+$, we can derive that one of the axes of the maximum volume ellipsoid centered at $\bar{x}$ is on the line l going through the origin and $\bar{x}$ for the scaled second order cone.

This is due to again the rich symmetries of the second order cones about the line l, and it is captured mathematically by the automorphism group of the cones. Specifically, in this case, we may prove the statement analogously as in Proposition 4.3.4 by using the set of matrices acting on $\mathbb{R}^{n+1}$ who maps $\bar{x}$ to itself and when restricted to the subspace of $span\{e_2, ..., e_n\}$, it is an orthogonal matrix. It is easy to check this set of matrices is a subset of the automorphism group of the cones. Again by similar techniques as before, we can prove that the other axes are on the lines of $\{\alpha e_2 : \alpha \in \mathbb{R}\}, \ldots,$ and $\{\alpha e_{n+1} : \alpha \in \mathbb{R}\}$ with the same length.

Note that any ellipsoid centered at $\bar{x}$ can be represented as:

$$B_H(\bar{x}, r) = \left\{ \begin{pmatrix} t \\ x \end{pmatrix} \in \mathbb{R}^N : \left\langle \begin{pmatrix} t - \sqrt{n} \\ x \end{pmatrix}, H \begin{pmatrix} t - \sqrt{n} \\ x \end{pmatrix} \right\rangle \leq r^2 \right\},$$

where $N := \frac{n(n+1)}{2}$. If we fix $\det(H) = 1$, the problem becomes finding an ellipsoid with the greatest r and contained in the cone K.

By information on the axes and their lengths as described above, we obtain the following equalities for H: $He_1 = \frac{1}{\lambda^{N-1}}e_1$, $He_2 = \lambda e_2$,..., and $He_N = \lambda e_N$, with $\lambda > 0$. Thus,

$$H = \begin{pmatrix} \frac{1}{\lambda^{N-1}} & 0 & \cdots & 0 \\ 0 & \lambda & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda \end{pmatrix}.$$

The inequality $\left\langle \begin{pmatrix} t - \sqrt{n} \\ x \end{pmatrix}, H \begin{pmatrix} t - \sqrt{n} \\ x \end{pmatrix} \right\rangle \leq r^2$ simply becomes $\frac{(t-\sqrt{n})^2}{\lambda^{N-1}} + \lambda \, ||x||_2^2 \leq r^2$, or

$$\lambda^N \, ||x||_2^2 \leq \lambda^{N-1} r^2 - (t - \sqrt{n})^2. \tag{4.17}$$

Hence, our question becomes finding the best λ such that the above inequality (4.17) implies

$$\sqrt{n-1} \, ||x||_2 \leq t. \tag{4.18}$$

We notice that (4.17) implies (4.18) if and only if

$$\lambda^{N-1} r^2 - (t - \sqrt{n})^2 \leq \frac{\lambda^N t^2}{n-1} \text{ and } t \geq 0. \tag{4.19}$$

Now, we focus on finding the best λ and r pair such that

$$\lambda^{N-1} r^2 - (t - \sqrt{n})^2 \leq \frac{\lambda^N t^2}{n-1} \tag{4.20}$$

holds with the largest r.

We rewrite (4.20) as:

$$r^2 \leq \frac{(\frac{\lambda^N}{n-1} + 1)t^2 - 2\sqrt{n}t + n}{\lambda^{N-1}}.$$

The minimum of the right hand side is obtained at $t = \frac{\sqrt{n}}{\frac{\lambda^N}{n-1}+1}$. Hence,

$$r^2 \leq \frac{n\lambda}{\lambda^N + n - 1}.$$

It is easy to check that the right hand side achieves its maximum at $\lambda = \sqrt[N]{\frac{n-1}{N-1}}$, and we get the maximum of r^2 being $\frac{n(N-1)^{\frac{N-1}{N}}(n-1)^{\frac{1-N}{N}}}{N} = \left(\frac{2}{n+2}\right)^{\frac{1}{N}}\left(\frac{n+2}{n+1}\right)$.

69

Plugging in the value of λ and r^2, and multiplying both sides by $\left(\frac{n+2}{2}\right)^{\frac{1}{N}}\left(\frac{n+1}{n+2}\right)$ we obtain our ellipsoid in the form as in the proposition above: $B_H(\bar{x}, 1)$, where

$$
H = \begin{pmatrix}
\frac{n+1}{2} & 0 & \cdots & 0 \\
0 & \frac{n+1}{n+2} & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \frac{n+1}{n+2}
\end{pmatrix}.
$$

It is straightforward to check that the second condition $t \geq 0$ in (4.19) is satisfied as well. Thus, this is indeed the maximum volume ellipsoid centered at $\bar{x}$ contained in the scaled second order cone. $\qquad\square$

Now, let us present the first proof for finding the maximum volume ellipsoids contained in general positive semidefinite cones with the center being I.

Theorem 4.5.5. *The maximum volume ellipsoid centered at I inscribed in the cone $\mathbb{S}^n_+$ is $B_H(I, 1)$, where H is a self-adjoint positive definite operator on $\mathbb{S}^n$ where $H(I) = \frac{n+1}{2}I$, and $H(A) = \frac{n+1}{n+2}A$ for any matrix A orthogonal to I.*

Proof. First we embed the space $\mathbb{S}^n_+$ in $\mathbb{R}^N$ by another isometry s2vec$'$, where s2vec$'$ is the composition of first applying the map s2vec, and then applying a permutation matrix to the resulting vector so that the new vector has its first n entries correspond to the n diagonal entries in the original matrix. Recall that $N = \frac{n(n+1)}{2}$.

Then, we apply a linear map M to the whole space so that the identity gets mapped to $(\sqrt{n}, 0, ..., 0)^T$. The linear map can be simply a rotation map on the subspace of $span\{(1, ..., 1, 0, ..., 0)^T, (1, 0, ..., 0)^T\}$, and identity map when restricted to the subspace orthogonal to it. We call the resulting cone K_S.

Given Proposition 4.5.3, we try to find the maximum volume ellipsoid centered at $\bar{x} = (\sqrt{n}, 0, ..., 0)^T$ in the transformed second order cone:

$$
K := \left\{ \begin{pmatrix} t \\ x \end{pmatrix} \in \mathbb{R} \oplus \mathbb{R}^{N-1} : t \geq \sqrt{n-1}\,\|x\|_2 \right\}.
$$

By construction, K is contained in K_S.

By Proposition 4.5.4, we have the maximum volume ellipsoid in K centered at $\bar{x}$ is $B_{\bar{H}}(\bar{x}, 1)$, with

$$
\bar{H} = \begin{pmatrix}
\frac{n+1}{2} & 0 & \cdots & 0 \\
0 & \frac{n+1}{n+2} & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \frac{n+1}{n+2}
\end{pmatrix}.
$$

Since this ellipsoid is contained in K, we must have it is contained in K_S as well.

Now, let us prove its optimality by utilizing Prop 4.3.3.

Let the optimal ellipsoid of the positive semidefinite cone be E. Consider the slice of the cone:

$$C := \{X \in \mathbb{S}_+^n : \langle I, X \rangle = n\},$$

and the slice of the optimal ellipsoid:

$$B := \{X \in E : \langle I, X \rangle = n\}.$$

By Prop 4.3.3, any linear map $Q \cdot Q^T$ with $Q \in O^n$, where O^n represents the set of orthognal n-by-n matrices, is in the automorphism group of the optimal ellipsoid and the compact set $P := \mathbb{S}_+^n \cap (\mathbb{S}_+^n)'$, where $(\mathbb{S}_+^n)'$ is the reflection of the cone around I. We see that this implies all the $Q \cdot Q^T$ are also in the automorphism group of C and B.

Since B is a slice of the ellipsoid, it must be an ellipsoid of one smaller dimension. Suppose U is one of the farthest points to I in B. Then, all of QUQ^T must be in the set B as well. Moreover, they are all at the same distance to I. Since $Q \cdot Q^T$ will generate an infinite full dimensional set of points in C, B must be a Euclidean ball.

By previous discussion, we know $K \subseteq K_S$. This implies $B \subseteq C$. Moreover, let $\bar{K}$ be the image of cone K when we apply a rotation back. Since B is a Euclidean ball centered at I, and by construction of the cone K, $B \subseteq \{X \in \bar{K} : \langle I, X \rangle = n\}$. This argument works for all non-empty slices of the ellipsoid E. Hence, we know the optimal ellipsoid must be contained in $\bar{K}$ which is contained in the positive semidefinite cone.

As a result the problem of finding the maximum volume ellipsoid centered at I in the positive semidefinite cone can be reduced to the problem of finding the maximum volume ellipsoid centered at I in the scaled second order cone. This shows the ellipsoid obtained above is indeed the maximum volume ellipsoid in $\mathbb{S}_+^n$ centered at I. The result follows. $\square$

4.5.3 The maximum volume ellipsoids over all self-adjoint positive definite operators by contact points and duality

We can also prove the optimality of the maximum volume ellipsoid in $\mathbb{S}_+^n$ using duality theory by finding a proper set of contact points. We use the same notations as in Proposition 4.5.5.

The contact points of the ellipsoid and the cone K are

$$\left(\frac{\sqrt{n}(N-1)}{N}, x_1, \ldots, x_{N-2}, \pm \sqrt{\frac{n(N-1)^2}{(n-1)N^2} - x_1^2 - \ldots - x_{N-2}^2} \right)^T,$$

where

$$-\frac{\sqrt{n}(N-1)}{\sqrt{n-1}N} \leq x_i \leq \frac{\sqrt{n}(N-1)}{\sqrt{n-1}N}, i \in \{1, \ldots, N-2\}.$$

We know these contact points correspond to rank $(n-1)$ matrices in $\mathbb{S}_+^n$. The reflection of these contact points around $\sqrt{2}e_1$ correspond to the full rank matrices in $\mathbb{S}_+^n$ which are the reflections of those rank $(n-1)$ matrices around identity. They also correspond to contact points of the ellipsoid and the compact set P. We want to select a set of contact points that corresponds to those rank $(n-1)$ matrices.

Proposition 4.5.6. *The contact points of $\mathbb{S}_+^n$ and $B_H(I,1)$: the maximum volume ellipsoid contained in $\mathbb{S}_+^n$ with center I is:*

$$\left\{ X : X = Q \begin{pmatrix} \frac{n(N-1)}{(n-1)N} & 0 & \ldots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \frac{n(N-1)}{(n-1)N} & \vdots \\ 0 & \ldots & \ldots & 0 \end{pmatrix} Q^T, Q \in O^n \right\}. \tag{4.21}$$

Proof. By the characterization and formula of the contact points of $B_{\bar{H}}(\bar{x},1)$ and K given above, we know the contact points of $B_H(I,1)$ (corresponding to $M^{-1}(K)$) and the cone $K' := \text{sMat}(M^{-1}(K))$ (corresponding to $M^{-1}(B_{\bar{H}}(\bar{x},1))$) are all on the hyperplane of trace equal to some constant C. By this formula, we know $\frac{N-1}{N}I$ is in the hyperplane. Thus, $C = \frac{n(N-1)}{N}$. Hence, the contact points of $B_H(I,1)$ and K' are exactly those located on the boundary of the cone K' and on the hyperplane of trace equal to $\frac{n(N-1)}{N}$.

By construction of K', we enforced that on any hyperplane of trace equal to some constant C, the only contact points of K' and $\mathbb{S}_+^n$ are the ones described in Proposition 4.5.3, i.e., the set of matrices with rank $n-1$, and every eigenvalue equal to $\frac{C}{n-1}$. Hence, the contact points of K' and $\mathbb{S}_+^n$ on the hyperplane of trace equal to $\frac{n(N-1)}{N}$ are exactly the set (4.21).

Putting these two arguments together, we conclude that the contact points of $\mathbb{S}_+^n$ and $B_H(I,1)$ are the set (4.21). $\square$

With the above characterization of the contact points of the cone $\mathbb{S}_+^n$ and the maximum volume ellipsoid, we provide a second proof for finding the maximum volume ellipsoids in $\mathbb{S}_+^n$ centered at I using duality theory.

By similar arguments as before, we know the maximum volume ellipsoid in $\mathbb{S}_+^n$ centered at I must be contained in the intersection of $\mathbb{S}_+^n$ and its reflection $(\mathbb{S}_+^n)'$ around I. We denote this resulting convex compact set by S'. We see the original problem is now reduced to finding the maximum volume ellipsoid contained in $S' \subseteq \mathbb{S}^n$.

Theorem 4.5.7. *The maximum volume ellipsoid centered at identity contained in S' is $B_H(I,1)$, where H is a self-adjoint positive definite operator on $\mathbb{S}^n$ where $H(I) = \frac{n+1}{2}I$, and $H(M) = \frac{n+1}{n+2}M$ for any matrix M orthogonal to I in $\mathbb{S}^n$.*

Proof. To prove this, we will use the Corollary 4.1.3. We will verify that all the conditions are satisfied for the Euclidean embedding of $B_H(I,1)$ and S'.

We first check that $B_H(I,1) \subseteq S'$. This can be proven in a similar fashion as in Proposition 4.5.5, i.e., we can verify the containment for any given hyperplane with a constant trace.

We take matrices described below as the contact points of S' and $B_H(I,1)$ using Proposition 4.5.6, by choosing a "balanced" set of orthogonal matrices in O^n:

We take the subset of matrices described in Proposition 4.5.6, where the 0 eigenvectors of the matrices correspond to vectors in $\mathbb{R}^n$ of the form:

$$(0, \ldots, \pm 1, \ldots, 0)^T \text{ or } \frac{1}{\sqrt{n}}(\pm 1, \ldots, \pm 1)^T.$$

As taking the 0 eigenvector to be v or $-v$ corresponds to the same matrix, we will only take one of the them. As proved in Proposition 4.5.6, these matrices are indeed contact points of S' and $B_H(I,1)$. For each contact point M, we will also include $2I - M$, the reflection around identity, as the corresponding contact point.

There are $n + n$ contact points generated from the first type, and $2^{n-1} + 2^{n-1} = 2^n$ contact points generated from the second type. These 2 sets of contact points when embedded in Euclidean space $\mathbb{R}^N$ ($N = \frac{n(n+1)}{2}$), are of the following forms:

Contact points p_i's of the first type are of the form $\left(\frac{n(N-1)}{N(n-1)}, \ldots, 0, \ldots, \frac{n(N-1)}{N(n-1)}, 0, \ldots, 0\right)^T$ and $\left(2 - \frac{n(N-1)}{N(n-1)}, \ldots, 2, \ldots, 2 - \frac{n(N-1)}{N(n-1)}, 0, \ldots, 0\right)^T$. After simplification, they are of the form:

$$\left(\frac{n+2}{n+1}, \ldots, 0, \ldots, \frac{n+2}{n+1}, 0, \ldots, 0\right)^T$$

73

and
$$\left(\frac{n}{n+1}, \ldots, 2, \ldots, \frac{n}{n+1}, 0, \ldots, 0\right)^T.$$

Only the first n entries of these vectors may be non-zero, corresponding to the diagonal entries of the matrices. The rest of the entries are 0, corresponding to the off-diagonal entries of the matrices. The number 0 appears once in each of the first n entries of the n vectors, and The number 2 also appears once in each of the first n entries of the other n vectors.

Let p_i be a vector whose i's entry is either 0 or 2, and let s2vec$'(I)$ be the embedding of I in $\mathbb{R}^N$. The N-by-N tensor $(p_i - \text{s2vec}'(I))(p_i - \text{s2vec}'(I))^T$ is of the form $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$. $A \in \mathbb{S}^n$, $B \in \mathbb{R}^{n \times (N-n)}$, $C \in \mathbb{R}^{(N-n) \times n}$, $D \in \mathbb{S}^{N-n}$. B, C and D are zero matrices, and

$$A = \frac{1}{(n+1)^2}\{ee^T + [(n+1)^2 + 2(n+1)]e_i e_i^T - (n+1)(e_i e^T + e e_i^T)\}$$

$$= \begin{pmatrix} \frac{1}{(n+1)^2} & \cdots & -\frac{1}{n+1} & \cdots & \frac{1}{(n+1)^2} \\ \vdots & \ddots & \vdots & & \vdots \\ -\frac{1}{n+1} & \cdots & 1 & \cdots & -\frac{1}{n+1} \\ \vdots & & \vdots & \ddots & \vdots \\ \frac{1}{(n+1)^2} & \cdots & -\frac{1}{n+1} & \cdots & \frac{1}{(n+1)^2} \end{pmatrix}.$$

All entries of A are equal to $\frac{1}{(n+1)^2}$, except the i-th row and column. A is a rank two matrix.

Summing the $2 \times n$ tensors of the first type, we obtain a matrix of the form $2 \times \begin{pmatrix} A & B \\ C & D \end{pmatrix}$. $A \in \mathbb{S}^n$, $B \in \mathbb{R}^{n \times (N-n)}$, $C \in \mathbb{R}^{(N-n) \times n}$, $D \in \mathbb{S}^{N-n}$. B, C and D are zero matrices, and

$$A = \begin{pmatrix} 1 + \frac{n-1}{(n+1)^2} & \cdots & -\frac{n+4}{(n+1)^2} \\ \vdots & \ddots & \vdots \\ -\frac{n+4}{(n+1)^2} & \cdots & 1 + \frac{n-1}{(n+1)^2} \end{pmatrix} = \begin{pmatrix} \frac{n(n+3)}{(n+1)^2} & \cdots & -\frac{n+4}{(n+1)^2} \\ \vdots & \ddots & \vdots \\ -\frac{n+4}{(n+1)^2} & \cdots & \frac{n(n+3)}{(n+1)^2} \end{pmatrix}$$

$$= \frac{1}{(n+1)^2}[(n+2)^2 I - (n+4)ee^T].$$

Entries of A have only two values. One value for diagonal entries, and another for off-diagonal entries.

Contact points p_j's of the second type are of the form

$$\left(\frac{N-1}{N},\ldots,\frac{N-1}{N},\pm\frac{(N-1)\sqrt{2}}{N(n-1)},\ldots,\pm\frac{(N-1)\sqrt{2}}{N(n-1)}\right)^T \quad (4.22)$$

and

$$\left(2-\frac{N-1}{N},\ldots,2-\frac{N-1}{N},\pm\frac{(N-1)\sqrt{2}}{N(n-1)},\ldots,\pm\frac{(N-1)\sqrt{2}}{N(n-1)}\right)^T. \quad (4.23)$$

The first n entries of these vectors are all equal to $\frac{N-1}{N}$ or $2-\frac{N-1}{N}$, and the rest of the entries are all equal to $\pm\frac{(N-1)\sqrt{2}}{N(n-1)}$, corresponding to the off-diagonal entries of the matrices. The sign pattern of the last $\frac{n(n-1)}{2}$ entries for each vector in (4.22) is dependent on the sign pattern of the corresponding 0 eigenvector v. Specifically, the signs of these vectors correspond to the signs of the column entries in the strict lower triangular part of the corresponding matrix, and the signs of these column entries are as follows:

For the i-th column c, if v_i is 1, then the signs of $c_{i+1},\ldots,c_n$ will be the same as that of $-v_{i+1},\ldots,-v_n$. If v_i is -1, then the signs of $c_{i+1},\ldots,c_n$ will be the same as that of $v_{i+1},\ldots,v_n$.

The vectors in (4.23) are the reflections of the vectors in (4.22) around s2vec$'(I)$, hence the signs of the last $\frac{n(n-1)}{2}$ entries of these vectors are the opposite of their counterparts in (4.22).

For any such contact point p_j, the N-by-N tensor $(p_j - \text{s2vec}'(I))(p_j - \text{s2vec}'(I))^T$ is of the form $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$, where $A \in \mathbb{S}^n$, $B \in \mathbb{R}^{n\times(N-n)}$, $C \in \mathbb{R}^{(N-n)\times n}$, $D \in \mathbb{S}^{N-n}$. Every entry of matrix A has value $\frac{1}{N^2}$, every entry of matrix B and C has value $\pm\frac{(N-1)\sqrt{2}}{N^2(n-1)}$, and

$$D = \begin{pmatrix} \frac{2(N-1)^2}{N^2(n-1)^2} & \cdots & \pm\frac{2(N-1)^2}{N^2(n-1)^2} \\ \vdots & \ddots & \vdots \\ \pm\frac{2(N-1)^2}{N^2(n-1)^2} & \cdots & \frac{2(N-1)^2}{N^2(n-1)^2} \end{pmatrix}.$$

Entries of D take one of two values: one value for diagonal entries and another for off-diagonal entries.

Summing the $2 \times 2^{n-1}$ tensors of the second type, we obtain a matrix of the form $2 \times \begin{pmatrix} A & B \\ C & D \end{pmatrix}$, where $A \in \mathbb{S}^n$, $B \in \mathbb{R}^{n\times(N-n)}$, $C \in \mathbb{R}^{(N-n)\times n}$, $D \in \mathbb{S}^{N-n}$. Every entry of matrix

A has value $\frac{2^{n-1}}{N^2}$, and B and C are zero matrices. This is because the 0 eigenvectors we generated of the second type include all the possible combinations of $+1$ and -1. Hence, for any given entry of matrix B and C, there are an equal number of $+\frac{(N-1)\sqrt{2}}{N^2(n-1)}$ and $-\frac{(N-1)\sqrt{2}}{N^2(n-1)}$ in the sum.

$$D = \text{Diag}(\frac{2^n(n+2)^2}{n^2(n+1)^2}, \ldots, \frac{2^n(n+2)^2}{n^2(n+1)^2}).$$

The off-diagonal entries sum to be 0 for the same reason as B and C being zero matrices.

Notice $H^{-1} = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$, where $A \in \mathbb{S}^n$, $B \in \mathbb{R}^{n \times (N-n)}$, $C \in \mathbb{R}^{(N-n) \times n}$, $D \in \mathbb{S}^{N-n}$. B and C are zero matrices,

$$A = \begin{pmatrix} 1 & \cdots & -\frac{1}{n+1} \\ \vdots & \ddots & \vdots \\ -\frac{1}{n+1} & \cdots & 1 \end{pmatrix} = \frac{1}{n+1}\left((n+2)I - ee^T\right),$$

and

$$D = \text{Diag}(\frac{n+2}{n+1}, \ldots, \frac{n+2}{n+1}).$$

Let the dual variables of the first type to be:

$$y_1 = \cdots = y_{2n} = \frac{n+1}{2(n+2)},$$

and the dual variables of the second type to be:

$$y_{2n+1} = \cdots = y_{2n+2^n} = \frac{n^2(n+1)}{2^{n+1}(n+2)}.$$

It is straightforward to check that

$$H^{-1} = \sum_{i=1}^{2n+2^n} y_i(p_i - \text{s2vec}'(I))(p_i - \text{s2vec}'(I))^T$$

and

$$\sum_{i=1}^{2n+2^n} y_i(p_i - \text{s2vec}'(I)) = \text{s2vec}'(I).$$

By Corollary 4.1.3, and similar arguments as before, the result follows. $\square$

Since the positive semidefinite cones are homogeneous, by Proposition 4.1.3, we obtain the general result of finding the largest volume ellipsoid inscribed in a positive semidefinite cone centered at an arbitrary point as a corollary.

Corollary 4.5.8. *The maximum volume ellipsoid centered at at an interior point $\bar{X}$ contained in $\mathbb{S}_+^n$ is*

$$B_{(\bar{X}^{-\frac{1}{2}} \overset{s}{\otimes} \bar{X}^{-\frac{1}{2}})H(\bar{X}^{-\frac{1}{2}} \overset{s}{\otimes} \bar{X}^{-\frac{1}{2}})}(\bar{X}, 1),$$

where H is a positive operator on $\mathbb{S}^n$ where $H(I) = \frac{n+1}{2}I$, and $H(M) = \frac{n+1}{n+2}M$ for any matrix M orthogonal to I in $\mathbb{S}^n$.

Proof. By Proposition 4.5.7, 4.1.3 and Proposition 4.1.2 and its proof, the result follows immediately. $\qquad\square$

4.6 Maximum volume ellipsoids in special homogeneous cones

We now consider a further generalization of the positive semidefinite cone. Consider the following cone K obtained from intersecting a positive semidefinite cone with a linear subspace:

$$K := \left\{ (t, U, X) \in \mathbb{R} \oplus \mathbb{R}^{m \times k} \oplus \mathbb{S}^m : \begin{pmatrix} tI & U^T \\ U & X \end{pmatrix} \succeq 0 \right\}.$$

It is clear that $K = \mathbb{S}_+^{m+k} \cap \mathbb{L}$, where

$$\mathbb{L} := \left\{ M \in \mathbb{S}^{m+k} : M = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix}, M_{11} = tI, \text{ for some } t \in \mathbb{R} \right\}.$$

It is easy to check that this cone is a homogeneous cone, since it is the Siegel cone constructed from $\mathbb{S}_+^m$ (Please refer to Truong and Tunçel [47] and the references therein). We can also easily find an automorphism that maps a given point in the cone K to another point in the cone.

Let us consider some geometric structures of this cone K:

Proposition 4.6.1 (Truong and Tunçel [47])**.** *The extreme rays of cone K are either of the form:*

$$\left\{ r \begin{pmatrix} tI & U^T \\ U & \frac{1}{t}UU^T \end{pmatrix}, r \geq 0, t > 0 \right\},$$

or of the form:

$$\left\{ r \begin{pmatrix} 0 & 0 \\ 0 & X \end{pmatrix}, r \geq 0, \; X \text{ on the extreme ray of } \mathbb{S}^m_+ \right\}.$$

Proof. For any matrix $M = \begin{pmatrix} tI & U^T \\ U & X \end{pmatrix}$ in K, we want to write it as a non-negative sum of the matrices in the proposition.

First, we consider the case of $t = 0$. Since M is positive semidefinite, we know in this case $U = 0$ as well. Hence, $M = \begin{pmatrix} 0 & 0 \\ 0 & X \end{pmatrix}$. The problem reduces to finding the extreme rays of $\mathbb{S}^m_+$, which we know already. Thus, we get the extreme rays of the second form.

Now, we consider the case when $t > 0$. Since M is positive semidefinite, by the Schur complement lemma, we know $X - \frac{1}{t} U U^T \succeq 0$. Hence, M can be written as a sum of two other matrices in the cone:

$$\begin{pmatrix} tI & U^T \\ U & \frac{1}{t} U U^T \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & X - \frac{1}{t} U U^T \end{pmatrix}.$$

We already know how to decompose the second matrix, and the first matrix corresponds to the first form in the proposition.

Hence, we are able to write any M as a non-negative sum of the matrices in the proposition.

It is clear that matrices of neither form can be represented by a positive sum of other matrices in the cone K. Thus, the result follows. $\square$

By Truong and Tunçel [47], We also know that any face of the cone K is of one of the two forms below:

$$\left\{ \begin{pmatrix} 0 & 0 \\ 0 & X \end{pmatrix} : X \in P, P \text{ is a face of } \mathbb{S}^m_+ \right\},$$

or

$$\left\{ r \begin{pmatrix} tI & U^T \\ U & X \end{pmatrix} : r, t \geq 0, tX - UU^T \succeq 0, \text{for some specific } X \in \text{relint}(P), P \text{ is a face of } \mathbb{S}^m_+ \right\}.$$

We see that every face in the lower order cone uniquely defines or relates to an infinite set of faces in the higher order cone.

Following a similar line of thought, as in the positive semidefinite cone case, we consider the distances from the identity, which corresponds to $(1, 0, I)^T$ in cone K, to the faces of

the cone on the slice $S := \{M \in K : \mathrm{Tr}(M) = m + k\}$. We know the minimum distance from $(1, 0, I)^T$, corresponding to $I_{m+k} \in \mathbb{S}_+^{m+k}$, to all the proper face of $\mathbb{S}_+^{m+k}$ with the trace being $m + k$ is obtained by

$$
D = \begin{pmatrix} \frac{m+k}{m+k-1} & & & \\ & \ddots & & \\ & & \frac{m+k}{m+k-1} & \\ & & & 0 \end{pmatrix},
$$

by Prop 4.5.3. Cone K is defined as $\mathbb{S}_+^{m+k} \cap \mathbb{L}$, hence S is a subset of the points on $\mathbb{S}_+^{m+k}$ with the trace being $m+k$. Hence, the minimum distance between $(1, 0, I)^T$ and any proper face of the cone on S is at least the distance from $(1, 0, I)^T$ to D. It is straightforward to verify that D is also in K. Thus, the minimum distance between $(1, 0, I)^T$ and all the proper face of the cone on S is obtained by D.

By Proposition 4.3.3, we know for any $g \in \mathrm{Aut}(K)$, $g(D)$ is also in the cone K. In particular, since D is on the boundary of the cone, $g(D)$ will be on the boundary of the cone as well. It is easy to see that any $Q \cdot Q$, where

$$
Q \in O^{m+k}, \text{ with } Q_{m+k,1} = \cdots = Q_{m+k,k} = 0
$$

are all valid automorphisms for K. Moreover, all such QDQ^T are of the same distance to $(1, 0, I)^T$. These matrices span a subspace of $\mathbb{R} \oplus \mathbb{R}^{m \times k} \oplus \mathbb{S}^m$. We denote this subspace as J. When restricting our attention to this subspace, the optimal ellipsoid is a ball of fixed radius. However, $J \neq \mathbb{R} \oplus \mathbb{R}^{m \times k} \oplus \mathbb{S}^m$ in this case. Hence, the intersection of the optimal ellipsoid and S may not be a unit ball any more. We need to consider the geometry of the optimal ellipsoid "outside" this subspace.

Since the optimal ellipsoid may not be a ball anymore, it will be very useful if we can characterize the distances from the identity $(1, 0, I)^T$ to the proper faces of the cone on S. Using the characterization of the face structures as above, we know the distances to the faces where $t = 0$ can be easily calculated using similar arguments as in Proposition 4.5.3, as any face of the first form is isometric to a face in the cone $\mathbb{S}_+^m$. For faces of the second type, we can again separate them in two sub-cases: P is a proper face of $\mathbb{S}_+^m$, or P is a face of full dimension in $\mathbb{S}_+^m$.

To find even more points of minimum distance from the identity to the boundaries of the cone on S and in general to explore the boundary structures of the cone K, the following proposition due to Truong and Tunçel [47] can be of great help. The proposition provides a description of a transitive subset of the automorphism group of the cone K.

Proposition 4.6.2 (Truong and Tunçel [47])**.** *A transitive subgroup of the automorphism group of K can be generated from the following linear maps for each $(t, u, x) \in$ interior of K:*

1. *$T_1(t, u, x) := (\alpha t, \sqrt{\alpha} u, x), \alpha > 0$,*

2. *$T_2(t, u, x) := (t, u + tv, x + 2B(u, v) + tB(v, v)), B(u, v) := \frac{1}{2}(uv^T + vu^T)$,*

3. *$T_3(t, u, x) := (t, Pu, PxP^T)$, where P is any (symmetric) non-singular matrix.*

This is a direct application of Lemma 1 in the paper and the fact that:
$\{G \cdot G^T : G \text{ is invertible}\}$ is the automorphism group of $\mathbb{S}_+^m$, and $\{G \cdot G^T : G \in \mathbb{S}_{++}^m\}$ is a transitive subset of it.

We can use this theorem to find all the points of minimum distance from the identity to the boundaries of the cone on S and hence obtain the subspace spanned by them. Moreover, the boundary structures of the cone K can be explored if we map some boundary points of interest by the maps in this transitive group.

Specifically, it may be worthwhile to look at boundary points with high rank, as it is likely that the contact points between the maximum volume ellipsoid and the cone are still obtained by matrices that correspond to a sum of extreme points as in the positive semidefinite cone case. We can also try to characterize the generators and the subset of transitive automorphisms of the cone as in the above proposition in matrix form. With the characterization, we can get a more concrete idea of where the automorphisms will map a matrix on the boundary of the cone to. Knowing the boundary structures of the cone will help to determine the possible axes and lengths of the maximum volume ellipsoid, and hence is the fundamental building block of finding the maximum volume ellipsoid in cone K.

With Proposition 4.6.2, we can also determine whether one of the axes of the maximum volume ellipsoid lies on the line going through the origin and $(1, 0, I)^T$ as in the positive semidefinite cone case.

Furthermore, we can consider coordinating this space differently so that the representation of the optimal ellipsoid is in a "nice" form. For instance, we may coordinate the space so that the matrix defining the ellipsoidal shape of the maximum volume ellipsoid is in the diagonal form. The most natural coordination of the space of $\mathbb{R} \oplus \mathbb{R}^{m \times k} \oplus \mathbb{S}^m$ may not reflect the geometric structure of the ellipsoid.

Chapter 5

Largest primal-dual pairs of ellipsoids with specified centers in convex cones

In this chapter, we consider the primal-dual version of the maximum ellipsoid problem. Given a pair of primal-dual cones and a pair of interior points in the cones respectively, we would like to maximize the product of the volumes of the pair of primal-dual ellipsoids centered at the given points inscribed in the cones over all possible ellipsoidal shapes.

Recall that given a self-adjoint positive definite operator H, we have the (local) norm being defined as $||x||_H := \langle x, Hx \rangle^{\frac{1}{2}}$, and its dual norm defined as $||x||_H^* := \langle x, H^{-1}x \rangle^{\frac{1}{2}}$. The problem of finding the largest primal-dual pairs of ellipsoids is related to dual norms as well, as the positive definite forms of the primal and dual ellipsoids induces a pair of norms dual to each other. In particular, the largest radii r_1, r_2 of a pair of ellipsoids dual to each other with positive form H and H^{-1} centered and inscribed in a pair of primal-dual cones is exactly the smallest distances in local norms defined by H and H^{-1} from the two centers to the boundaries of the cones.

5.1 Largest primal-dual pairs of ellipsoids in positive semidefinite cones

We first consider the special symmetric cone of $\mathbb{S}_+^n$ with its dual cone also equal to $\mathbb{S}_+^n$. We know for any $\bar{X} \in \text{int}(\mathbb{S}_+^n)$ and $\bar{S} \in \text{int}(\mathbb{S}_+^n)$, there exist an automorphism "$W \cdot W$", $W \in \mathbb{S}_{++}^n$, such that $\bar{X}$ is mapped to $\bar{S}$, $\text{int}(\mathbb{S}_+^n)$ is mapped to $\text{int}(\mathbb{S}_+^n)$ and $\text{bd}(\mathbb{S}_+^n)$ is mapped to $\text{bd}(\mathbb{S}_+^n)$ under the automorphism.

It is proved in Todd [46] that this automorphism induces a pair of ellipsoidal shapes $W \cdot W$ and $W^{-1} \cdot W^{-1}$ that maximizes the product of the volumes of the ellipsoids D and D^*, where D and D^* are ellipsoids centered at $\bar{X}$ and $\bar{S}$ with ellipsoidal shapes H and H^{-1} for any given self-adjoint and positive definite H and inscribed in the primal cone $\mathbb{S}_+^n$ and dual cone $\mathbb{S}_+^n$.

Given a pair of points $\bar{X}$ in the interior of the positive definite cone $\mathbb{S}_+^n$ and $\bar{S}$ in the interior of $\mathbb{S}_+^n$, the corresponding dual cone, we define:

$$\alpha_{\bar{X}}(H) := \max\{\alpha_{\bar{X}} : B_H(\bar{X}; \alpha_{\bar{X}}) \subseteq \mathbb{S}_+^n\},$$

$$\alpha_{\bar{S}}(H^{-1}) := \max\{\alpha_{\bar{S}} : B_{H^{-1}}(\bar{S}; \alpha_{\bar{S}}) \subseteq \mathbb{S}_+^n\}.$$

The main theorem for the positive semidefinite cone setting is as follows:

Theorem 5.1.1 (Todd [46]). *Let $\bar{X}$ be in the interior of the positive definite cone $\mathbb{S}_+^n$ and $\bar{S}$ be in the interior of the dual cone $\mathbb{S}_+^n$. Then,*

$$\max_{H \succ 0} \alpha_{\bar{X}}(H)\alpha_{\bar{S}}(H^{-1}) = \alpha_{\bar{X}}(W \cdot W)\alpha_{\bar{S}}(W^{-1} \cdot W^{-1}),$$

where $W \in \mathbb{S}_{++}^n$ is the unique matrix that satisfies the equality $W\bar{X}W = \bar{S}$.

Proof. It is easy to check that $W \cdot W$, where

$$W = \bar{X}^{-\frac{1}{2}}(\bar{X}^{\frac{1}{2}}\bar{S}\bar{X}^{\frac{1}{2}})^{\frac{1}{2}}\bar{X}^{-\frac{1}{2}}$$

is an automorphism of $\mathbb{S}_+^n$, and $W\bar{X}W = \bar{S}$. There must exist a $\bar{Z} \in \mathrm{bd}(\mathbb{S}_+^n)$, such that

$$\left\|\bar{Z} - \bar{X}\right\|_{W \cdot W} = \alpha_{\bar{X}}(W \cdot W),$$

i.e., the distance between $\bar{Z}$ and $\bar{X}$ is exactly $\alpha_{\bar{X}}(W \cdot W)$ with respect to the local norm. We can also find a non-zero $\bar{U} \in \mathrm{bd}(\mathbb{S}_+^n)$ such that $\bar{U}\bar{Z} = 0$ by eigenvalue decomposition of $\bar{Z}$. Thus, $\langle \bar{U}, \bar{Z} \rangle = 0$ (we can obtain the same result by supporting hyperplane theorem).

Since $\bar{U} \in \mathrm{bd}(\mathbb{S}_+^n)$ (in the dual cone), and $B_{W \cdot W}(\bar{X}; \alpha_{\bar{X}}(W \cdot W)) \subseteq \mathbb{S}_+^n$ (ellipsoid contained entirely in the primal cone) we have:

$$\min\{\langle \bar{U}, P \rangle : P \in B_{W \cdot W}(\bar{X}; \alpha_{\bar{X}}(W \cdot W))\} = 0 = \langle \bar{U}, \bar{Z} \rangle.$$

$W \cdot W$ is an automorphism that maps $\bar{X}$ to $\bar{S}$, $\mathrm{int}(\mathbb{S}_+^n)$ to $\mathrm{int}(\mathbb{S}_+^n)$, and $\mathrm{bd}(\mathbb{S}_+^n)$ to $\mathrm{bd}(\mathbb{S}_+^n)$. Moreover, by simple algebra we can show $W \cdot W$ maps an ellipsoid of shape $W \cdot W$ to

an ellipsoid of shape $W^{-1} \cdot W^{-1}$. Hence, we have $W \cdot W$ maps $B_{W \cdot W}(\bar{X}; \alpha_{\bar{X}}(W \cdot W))$ to $B_{W^{-1} \cdot W^{-1}}(\bar{S}; \alpha_{\bar{S}}(W^{-1} \cdot W^{-1}))$, i.e., $W \cdot W$ maps the corresponding maximal primal ellipsoid to its maximal dual ellipsoid.

Since $W \cdot W$ maps the boundary of $\mathbb{S}_+^n$ to the boundary of $\mathbb{S}_+^n$, we can take $W \bar{Z} W$ (the image of the boundary point $\bar{Z}$ in the dual cone, which is on the boundary of the dual $\mathbb{S}_+^n$) to be $\bar{T}$, and let $\bar{V}$ to be $W^{-1} \bar{U} W^{-1}$. This image is in primal $\mathbb{S}_+^n$, since $W^{-1} \cdot W^{-1}$ is an automorphism as well. In this way,

$$\langle \bar{V}, \bar{T} \rangle = \langle W^{-1} \bar{U} W^{-1}, W \bar{Z} W \rangle = \langle \bar{U}, \bar{Z} \rangle = 0.$$

Similarly, with $\bar{T} \in \mathrm{bd}(\mathbb{S}_+^n)$ we get $\left\| \bar{T} - \bar{S} \right\|_{W \cdot W}^* = \alpha_{\bar{S}}(W^{-1} \cdot W^{-1})$, and for $\bar{V} \in \mathbb{S}_+^n \setminus \{0\}$,

$$\min\{\langle \bar{V}, Q \rangle : Q \in B_{W^{-1} \cdot W^{-1}}(\bar{S}; \alpha_{\bar{S}}(W^{-1} \cdot W^{-1}))\} = 0 = \langle \bar{V}, \bar{T} \rangle .$$

The following results where $\bar{U}$ and $\bar{V}$ are nonzero can be proven in a similar way as in Proposition 2.3.5:

$$\min\{\langle \bar{U}, P \rangle : P \in B_H(\bar{X}; \alpha_{\bar{X}})\} = \langle \bar{U}, \bar{X} \rangle - \alpha_{\bar{X}} \left\| \bar{U} \right\|_H^* ,$$

and the minimum value is attained by $P := \bar{X} - (\alpha_{\bar{X}} / \left\| \bar{U} \right\|_H^*) H^{-1} \bar{U}$.

$$\min\{\langle \bar{V}, Q \rangle : Q \in B_{H^{-1}}(\bar{S}; \alpha_{\bar{S}})\} = \langle \bar{V}, \bar{S} \rangle - \alpha_{\bar{S}} \left\| \bar{V} \right\|_H ,$$

and the minimum value is attained by $Q := \bar{S} - (\alpha_{\bar{S}} / \left\| \bar{V} \right\|_H) H \bar{V}$.

Moreover, we may suppose $\left\| \bar{U} \right\|_{W \cdot W}^* = 1$. Since $\bar{V} = W^{-1} \bar{U} W^{-1}$, we have $\left\| \bar{V} \right\|_{W \cdot W} = 1$ as well.

Hence,

$$\langle \bar{U}, \bar{X} \rangle = \alpha_{\bar{X}}(W \cdot W) \left\| \bar{U} \right\|_{W \cdot W}^* = \alpha_{\bar{X}}(W \cdot W), \tag{5.1}$$

and

$$\langle \bar{V}, \bar{S} \rangle = \alpha_{\bar{S}}(W^{-1} \cdot W^{-1}) \left\| \bar{V} \right\|_{W \cdot W} = \alpha_{\bar{S}}(W^{-1} \cdot W^{-1}). \tag{5.2}$$

Here we use the inner product as a measure of the radius, and will use it as an intermediate step to compare the radius of the optimal shape with that of the other shapes.

By $B_H(\bar{X}; \alpha_{\bar{X}}(H)) \subseteq \mathbb{S}_+^n$, and the argument above, we have:

$$\min\{\langle \bar{U}, P \rangle : P \in B_H(\bar{X}; \alpha_{\bar{X}})\} = \langle \bar{U}, \bar{X} \rangle - \alpha_{\bar{X}} ||\bar{U}||_H^* \geq 0$$

for any arbitrary self-adjoint and positive-definite H from $\mathbb{S}^n$ to $\mathbb{S}^n$. Thus,

$$\langle \bar{U}, \bar{X} \rangle \geq \alpha_{\bar{X}}(H) ||\bar{U}||_H^* . \tag{5.3}$$

Similarly, we can obtain

$$\langle \bar{V}, \bar{S} \rangle \geq \alpha_{\bar{S}}(H^{-1}) ||\bar{V}||_H . \tag{5.4}$$

Now, we are going to use them as an intermediate step to compare the radius of the optimal shape with that of the other shapes:

$$
\begin{aligned}
\alpha_{\bar{X}}(W \cdot W)\alpha_{\bar{S}}(W^{-1} \cdot W^{-1}) \quad &= \langle \bar{U}, \bar{X} \rangle \langle \bar{V}, \bar{S} \rangle \\
&\geq \alpha_{\bar{X}}(H)\alpha_{\bar{S}}(H^{-1}) ||\bar{U}||_H^* ||\bar{V}||_H \\
&\geq \alpha_{\bar{X}}(H)\alpha_{\bar{S}}(H^{-1}) \langle \bar{U}, \bar{V} \rangle \\
&= \alpha_{\bar{X}}(H)\alpha_{\bar{S}}(H^{-1}).
\end{aligned}
$$

In the above, the first equality follows by equalities (5.1) and (5.2), the first inequality follows from inequalities (5.3) and (5.4), the second inequality uses Cauchy-Schwarz inequality, and the last equation follows from the fact that

$$\langle \bar{U}, \bar{V} \rangle = \langle \bar{U}, W^{-1}\bar{U}W^{-1} \rangle = (||\bar{U}||_{W \cdot W}^*)^2 = 1.$$

This concludes the proof. $\qquad\square$

It follows that the largest ellipsoid pair centered at identity for a pair of primal-dual semidefinite cones is a pair of unit Euclidean balls. The product of the radii of the ellipsoids is 1.

However, we know the radii of the largest inscribed ellipsoid centered at identity in a single semidefinite cone is greater than 1. In particular,

$$r = \left(\frac{2}{n+2}\right)^{\frac{1}{N}} \left(\frac{n+2}{n+1}\right).$$

If both of the semidefinite cones take their largest ellipsoids centered at identity separately, the product of the radii of the ellipsoids will be $\left(\frac{2}{n+2}\right)^{\frac{2}{N}} \left(\frac{n+2}{n+1}\right)^2$, which is greater than 1.

Nevertheless, we must note that as n approaches infinity, this value gets close to one. For example, for $n = 2$, 10, 100 and 1000, this product of the radii is less than 1.059, 1.056, 1.01 and 1.002, respectively.

However, this pair of ellipsoids are not dual to each other. To make them a primal-dual pair, we need to change the ellipsoidal shape of one of the ellipsoids. Geometrically, we will change the ellipsoid by stretching the axis along identity of the ellipsoid. The same arguments as in Proposition 4.5.4 can be used to obtain the largest radius of the new ellipsoid so that it is still contained in the cone. Its radius is:

$$\sqrt{\left(n \sqrt[N]{\frac{N-1}{n-1}} \right) \bigg/ \left(\frac{N-1}{n-1} + n - 1 \right)} = \sqrt{\frac{2}{3} \left(\frac{n+2}{2} \right)^{\frac{1}{N}}}.$$

Now, if we calculate the product of the radii of the primal-dual pair of ellipsoids obtained this way, we get:

$$\sqrt{\frac{2}{3} \left(\frac{n+2}{n+1} \right)}.$$

Clearly, this product is smaller than 1 for all $n \geq 2$, and tends to $\sqrt{\frac{2}{3}}$ as n goes to infinity.

Geometrically speaking, taking a unit ball and squeezing it on the axis along identity potentially allows the inscribed ellipsoid to be larger, and stretching it along the same axis makes it smaller. There is some trade-off between them. It turns out, in this way, no matter how much larger we can make one of the ellipsoids, the other ellipsoid will become even smaller according to the measure of the product of the radii. Hence, to maximize the product of the radii, we shall choose a pair of unit balls.

Proposition 5.1.2. *For an arbitrary pair of interior points $\bar{X}$ and $\bar{S}$ in $\mathbb{S}^n_+$, the product of the radius of the largest volume ellipsoid centered at $\bar{X}$ contained in the positive semidefinite cone and the radius of its dual ellipsoid centered at $\bar{S}$ contained in the dual cone is* $\det(\bar{X})^{-\frac{2}{n}} \sqrt{\left(\frac{2}{n+2} \right)^{\frac{1}{N}} \left(\frac{n+2}{n+1} \right)} \cdot \min \left\{ \frac{\langle S, \bar{S} \rangle}{\sqrt{\langle S, \hat{H}S \rangle}} : \forall S \in \mathbb{S}^n_+, \operatorname{rank}(S) = 1, \operatorname{Tr}(S) = 1 \right\}$, *where* $\hat{H} := \frac{(\bar{X}^{-\frac{1}{2}} \overset{s}{\otimes} \bar{X}^{-\frac{1}{2}}) H (\bar{X}^{-\frac{1}{2}} \overset{s}{\otimes} \bar{X}^{-\frac{1}{2}})}{\left\| (\bar{X}^{-\frac{1}{2}} \overset{s}{\otimes} \bar{X}^{-\frac{1}{2}}) H (\bar{X}^{-\frac{1}{2}} \overset{s}{\otimes} \bar{X}^{-\frac{1}{2}}) \right\|}.$

Proof. According to the automorphism theorem 4.1.3 and Theorem 4.5.7, the largest ellipsoid centered at $\bar{X}$ and contained in the positive semidefinite cone has radius

$$\det(\bar{X})^{-\frac{2}{n}} \sqrt{\left(\frac{2}{n+2} \right)^{\frac{1}{N}} \left(\frac{n+2}{n+1} \right)}$$

with shape $\hat{H} := \frac{(\bar{X}^{-\frac{1}{2}} \overset{s}{\otimes} \bar{X}^{-\frac{1}{2}}) H (\bar{X}^{-\frac{1}{2}} \overset{s}{\otimes} \bar{X}^{-\frac{1}{2}})}{\left\| (\bar{X}^{-\frac{1}{2}} \overset{s}{\otimes} \bar{X}^{-\frac{1}{2}}) H (\bar{X}^{-\frac{1}{2}} \overset{s}{\otimes} \bar{X}^{-\frac{1}{2}}) \right\|}$. H is the same positive definite form as in Theorem 4.5.7.

The largest radius of the ellipsoid centered at $\bar{S}$ with the dual ellipsoidal shape inscribed in the dual cone (which is also a positive semidefinite cone) can be expressed in a similar way as in (4.16) of Proposition 4.5.2, which is:

$$\min \left\{ \frac{\langle S, \bar{S} \rangle}{\sqrt{\left\langle S, \hat{H} S \right\rangle}} : \forall S \in \mathbb{S}_+^n, \operatorname{rank}(S) = 1, \operatorname{Tr}(S) = 1 \right\}.$$

The result follows. $\qquad\square$

5.2 Largest primal-dual pairs of ellipsoids in self-scaled (symmetric) cones

Let K be a symmetric cone, and F be a self-scaled barrier for K. Pick any $\bar{x} \in \operatorname{int}(K)$ and $\bar{s} \in \operatorname{int}(K^*)$. It is proved in Todd (2009) that the unique $w \in \operatorname{int}(K)$ such that $F''(w)(K) = K^*$ and $F''(w)\bar{x} = \bar{s}$ induces a pair of ellipsoid shapes $F''(w)$ and $[F''(w)]^{-1}$ that maximizes the product of the volumes of the ellipsoids D and D^*. D and D^* are ellipsoids inscribed in K and K^* respectively with centers $\bar{x}$ and $\bar{s}$, and the shapes are defined by H and H^{-1} for some self-adjoint and positive definite H.

We denote
$$\alpha_{\bar{x}}(H) := \max\{\alpha_{\bar{x}} : B_H(\bar{x}; \alpha_{\bar{x}}) \subseteq K\},$$
and
$$\alpha_{\bar{s}}(H^{-1}) := \max\{\alpha_{\bar{s}} : B_{H^{-1}}(\bar{s}; \alpha_{\bar{s}}) \subseteq K^*\}.$$

We will abbreviate the notation of the positive definite form $F''(v)$ by v in subscript, and denote $B_{H^{-1}}(\bar{s}; \alpha_{\bar{s}}) := B_H^*(\bar{s}; \alpha_{\bar{s}})$.

The theorem by Todd [46] for general symmetric cones is as follows.

Theorem 5.2.1 (Todd [46]). *Let K be a symmetric cone, F be a self-scaled barrier for K, $\bar{x} \in \operatorname{int}(K)$ and $\bar{s} \in \operatorname{int}(K^*)$. Then,*
$$\max_{H \succ 0} \alpha_{\bar{x}}(H)\alpha_{\bar{s}}(H^{-1}) = \alpha_{\bar{x}}(F''(w))\alpha_{\bar{s}}([F''(w)]^{-1}),$$
where $w \in \operatorname{int}(K)$ is the unique scaling point so that $F''(w)(K) = K^*$ and $F''(w)\bar{x} = \bar{s}$.

Proof. Given $\bar{x}, \bar{s}$ and their scaling point w as above, and by the above notations, there must exist a $\bar{z} \in \mathrm{bd}(K)$, such that $||\bar{z} - \bar{x}||_w = \alpha_{\bar{x}}(F''(w))$. By supporting hyperplane theorem, there must exist a non-zero $\bar{u} \in K^*$ such that $\langle \bar{u}, \bar{z} \rangle = 0$.

Since $B_w(\bar{x}; \alpha_{\bar{x}}(F''(w))) \subseteq K$, we have:

$$0 = \langle \bar{u}, \bar{z} \rangle = \min\{\langle \bar{u}, p \rangle : p \in B_w(\bar{x}; \alpha_{\bar{x}}(F''(w)))\}.$$

We can also prove that $F''(w)$ maps $B_w(\bar{x}; \alpha_{\bar{x}}(F''(w)))$ to $B_w^*(\bar{s}; \alpha_{\bar{s}}([F''(w)]^{-1}))$. And since $F''(w)$ maps K to K^*, $\mathrm{int}(K)$ to $\mathrm{int}(K^*)$, it must map $\mathrm{bd}(K)$ to $\mathrm{bd}(K^*)$.

We can take $F''(w)\bar{z}$ (on the boundary of K^* by above argument) to be $\bar{t}$, and let $\bar{v}$ to be $[F''(w)]^{-1}\bar{u} \in K$. Thus,

$$\langle \bar{v}, \bar{t} \rangle = \langle [F''(w)]^{-1}\bar{u}, F''(w)\bar{z} \rangle = \langle \bar{u}, \bar{z} \rangle = 0.$$

Similarly, with $\bar{t} \in \mathrm{bd}(K^*)$ we get $||t - s||_w^* = \alpha_s(w)$, and for $\bar{v} \in K \setminus \{0\}$,

$$0 = \langle \bar{v}, \bar{t} \rangle = \min\{\langle \bar{v}, q \rangle : q \in B_w^*(\bar{s}; \alpha_{\bar{s}}([F''(w)]^{-1}))\}.$$

The following results where $\bar{u}$ and $\bar{v}$ are nonzero can be proven in a similar way as in Proposition 2.3.5:

$\min\{\langle \bar{u}, p \rangle : p \in B_H(\bar{x}; \alpha_{\bar{x}})\} = \langle \bar{u}, \bar{x} \rangle - \alpha_{\bar{x}} ||\bar{u}||_H^*$, attained by $p = \bar{x} - (\alpha_{\bar{x}}/||\bar{u}||_H^*)H^{-1}\bar{u}$.

$\min\{\langle \bar{v}, q \rangle : q \in B_H^*(\bar{s}; \alpha_{\bar{s}})\} = \langle \bar{v}, \bar{s} \rangle - \alpha_{\bar{s}} ||\bar{v}||_H$, attained by $q = \bar{s} - (\alpha_{\bar{s}}/||\bar{v}||_H)H\bar{v}$.

We may suppose $||\bar{u}||_w^* = 1$. Since $\bar{v} = [F''(w)]^{-1}\bar{u}$, we have $||\bar{v}||_w = 1$ as well.

Together with the above arguments, we get:

$$\langle \bar{u}, \bar{x} \rangle = \alpha_{\bar{x}}(F''(w)) ||\bar{u}||_w^* = \alpha_{\bar{x}}(F''(w)), \tag{5.5}$$

and

$$\langle \bar{v}, \bar{s} \rangle = \alpha_{\bar{s}}([F''(w)]^{-1}) ||\bar{v}||_w = \alpha_{\bar{s}}([F''(w)]^{-1}). \tag{5.6}$$

Since $B_H(\bar{x}; \alpha_{\bar{x}}(H)) \subseteq K$, so for an arbitrary self-adjoint and positive-definite H from $\mathbb{E}$ to $\mathbb{E}^*$, we have

$$\langle \bar{u}, \bar{x} \rangle \geq \alpha_{\bar{x}}(H) ||\bar{u}||_H^*, \tag{5.7}$$

and similarly

$$\langle \bar{v}, \bar{s} \rangle \geq \alpha_{\bar{s}}(H^{-1}) \, \|\bar{v}\|_H \, . \tag{5.8}$$

Thus, we have:

$$\begin{aligned}
\alpha_{\bar{x}}(F''(w))\alpha_{\bar{s}}([F''(w)]^{-1}) = & \quad \langle \bar{u}, \bar{x} \rangle \, \langle \bar{v}, \bar{s} \rangle \\
\geq & \quad \alpha_{\bar{x}}(H)\alpha_{\bar{s}}(H^{-1}) \, \|\bar{u}\|_H^* \, \|\bar{v}\|_H \\
\geq & \quad \alpha_{\bar{x}}(H)\alpha_{\bar{s}}(H^{-1}) \, \langle \bar{u}, \bar{v} \rangle \\
= & \quad \alpha_{\bar{x}}(H)\alpha_{\bar{s}}(H^{-1}).
\end{aligned}$$

In the above, the first equality follows by equalities (5.5) and (5.6), the first inequality follows from inequalities (5.7) and (5.8), the second inequality uses Cauchy-Schwarz inequality, and the last equation follows from the fact that

$$\langle \bar{u}, \bar{v} \rangle = \left\langle \bar{u}, [F''(w)]^{-1}\bar{u} \right\rangle = (\|\bar{u}\|_w^*)^2 = 1$$

This concludes the proof. □

Notice that in the above proof, we heavily rely on the fact that barrier function F maps boundary points in the primal cone to boundary points in the dual cone, which comes automatically for any symmetric cone K and its self-scaled barrier. This proof uses $\langle \bar{u}, \bar{x} \rangle$ and $\langle \bar{v}, \bar{s} \rangle$ as a "bridge" to compare the largest possible radii of ellipsoids defined by $F''(w)$ and ones defined by other H.

The problem of finding the largest pair of primal-dual ellipsoid for symmetric cones using Jordan Algebra (for more detailed treatment on Jordan Algebra, please refer to Faraut and Korányi [14]) is considered by Lim [33]. An explicit formula in terms of the minimum eigenvalues of the centers are given.

5.3 Largest primal-dual pairs of ellipsoids for homogeneous and hyperbolicity cones

In this section, we try to explore on the largest primal-dual pairs of ellipsoids in a pair of homogeneous or hyperbolicity primal-dual cones. We would like to see how the result of the largest pair of primal-dual ellipsoids in symmetric cones may change in these more general setting.

Definition 5.3.1. *A homogeneous polynomial* $p(x) \in \mathbb{R}[x_1, ..., x_n]$ *is* hyperbolic *with respect to a vector* $d \in \mathbb{R}^n$ *if* $p(d) \neq 0$, *and if for all* $x \in \mathbb{R}^n$ *the univariate polynomial* $t \to p(x - \lambda d)$ *has only real roots.*

Definition 5.3.2. *Suppose polynomial* $p(x)$ *is hyperbolic with respect to d. The* hyperbolicity cone *is the set:* $\{x \in \mathbb{R}^n : \text{ all the roots of } p(x - \lambda d) \text{ is nonnegative}\}$.

We notice that all the cones we defined so far are special cases of hyperbolicity cones:

1. The nonnegative orthant is a hyperbolicity cone with hyperbolic function $x_1 \times \cdots \times x_n$ and the directional vector $(1, ..., 1)^T$, the vector of all ones. We may check that the roots of the polynomial $p(t) = (x_1 - t) \ldots (x_n - t)$ are all nonnegative indeed happens when all the $x_i \geq 0$. This is exactly the nonnegative orthant.

2. The second order cone is also a hyperbolicity cone with hyperbolic function $y^2 - ||x||_2^2$ and the directional vector $(0, ..., 0, 1)^T$, which is all $x_i = 0$, and $y = 1$. We may check that the roots of the polynomial $p(t) = (y - t)^2 - ||x||_2^2$ are all nonnegative indeed happens when $y^2 \geq ||x||_2^2$. This is exactly the second order cone.

3. The positive semidefinite cone is another special case of hyperbolicity cone with hyperbolic function $\det(X)$ and the directional vector I, i.e., the identity matrix. We may check that the roots of the polynomial $p(t) = \det(X - tI)$ are all nonnegative indeed happens when the eigenvalues of the matrix X are all nonnegative. This is exactly the positive semidefinite cone.

In fact, the set of homogeneous cones is a proper subset of the set of hyperbolicity cones.

In the homogeneous and hyperbolicity cone settings, we may still find the Hessian of the scaling point, however the Hessian may map the primal cone into a proper subset of the dual cone in this case. As a result, the object $\bar{v}$ in the proof of Theorem 5.2.1 involving the inverse of the Hessian can be mapped outside the primal cone. The same method does not seem to work immediately for homogeneous and hyperbolicity cones.

One possible remedy is to map $\bar{v}$ back to the cone by projection with respect to the Euclidean norm. Since the cone is a convex set, the point projected is the unique point that has the minimum distance to $\bar{v}$. Another possible fix for this is to find a point of minimum distance to $\bar{v}$ on the boundary of the cone under a suitable local norm. We can then normalize the point on the boundary of the cone and follow the rest of the proof to obtain an approximation ratio of $\langle \bar{u}, \bar{v} \rangle$.

5.4 Other generalizations of primal-dual pairs of ellipsoids

The version of finding the largest primal-dual pairs of ellipsoids by Todd when applied to the positive semidefinite cone setting or positive semidefinite Hermitian setting can be formulated as follows:

Given $A, B \succ 0$, find a self-adjoint positive definite form H on the space, so that it achieves the maximum of the following optimization problem:

$$\max_{H \succ 0} \alpha_A(H) \alpha_B(H^{-1}).$$

In Lim [32], the result of Todd in the setting of positive semidefinite Hermitian cone is generalized by an additional symmetric gauge norm on the eigenvalues of the matrices. We will see that Todd's result corresponds to the case where the gauge norm is the standard 2-norm in $\mathbb{R}^n$.

A norm ϕ on the Euclidean space $\mathbb{R}^n$ is a *symmetric gauge function* if it is invariant under permutations, sign changes of coordinates, and $\phi((1, 0, \ldots, 0)^T) = 1$. Every symmetric gauge function induces a unitarily invariant norm on the whole matrix space, and in particular for hermitian or symmetric matrices we have the following definition:

Let $\mathbb{H}^n$ be the space of n-by-n Hermitian matrices, and let ϕ be a symmetric gauge norm on $\mathbb{R}^n$. $||\cdot||_\phi$ is the corresponding unitarily invariant norm on the space of $\mathbb{H}^n$ by $||X||_\phi = \phi(\lambda(X))$, where $\lambda(X) = (\lambda_1(X), \ldots, \lambda_n(X))^T \in \mathbb{R}^n$, and $\lambda_i(X)$ are the eigenvalues of the matrix $X \in \mathbb{H}^n$. The eigenvalues $\lambda_i(X)$ are ordered in non-increasing order. For instance, when we take the symmetric gauge norms to be the 1-norm and 2-norm on $\mathbb{R}^n$, the corresponding unitarily invariant norms on the space of $\mathbb{H}^n$ are the nuclear norm and the operator 2-norm respectively.

For any given $X \succ 0$, we define a norm on $\mathbb{H}^n$ as:

$$||Y||_{\phi,X} := \left|\left| X^{-\frac{1}{2}} Y X^{-\frac{1}{2}} \right|\right|_\phi.$$

We will use similar notations as in the last section,

$$B_{\phi,X}(A, t) := \{U \in \mathbb{H}^n : ||A - U||_{\phi,X} \leq t\},$$

and

$$\alpha_A(\phi, X) := \max\{t > 0 : B_{\phi,X}(A, t) \subseteq \mathbb{S}^n_+\}.$$

We first consider the problem of finding the largest volume ellipsoid contained in the positive semidefinite cone given a specific center A and an ellipsoidal shape $||\cdot||_{\phi,X}$, i.e., to find $\alpha_A(\phi, X)$. Then, we consider a similar problem as in Todd's largest primal-dual ellipsoids section, i.e., given two interior point A, B and a gauge norm ϕ, to find

$$\max_{X \succ 0} \alpha_A(\phi, X)\alpha_B(\phi, X^{-1}).$$

We see that Todd's result is the special case where the gauge norm is the 2-norm in the Euclidean space $\mathbb{R}^n$.

The first result below by Lim is a direct generalization of our Proposition 4.5.1. Proposition 4.5.1 deals with the special case of the gauge norm being the standard 2-norm. We include a modification of the proof here to illustrate the analogy and difference between the proof techniques for the more general gauge norm and the special case of it being the 2-norm in the Euclidean space $\mathbb{R}^n$.

Theorem 5.4.1 (Lim [32]). *Let ϕ be any symmetric gauge norm and let $A, X \succ 0$. Then*

$$\alpha_A(\phi, X) = \lambda_{min}(X^{-1}A).$$

Proof. Given a symmetric gauge norm ϕ, we first notice that for any $x \in \mathbb{R}^n$,

$$|x_i| \leq \phi(x), i = 1, 2, ..., n. \tag{5.9}$$

This is because of the monotonicity property of ϕ: $|x_i| = \phi(x_i e_i) \leq \phi(x)$.

Let us first show the largest radius is at least $\lambda_{min}(A)$, i.e.,

$$||A - Y||_\phi \leq \lambda_{min}(A), Y \succeq 0.$$

When ϕ is 2-norm, this statement corresponds to the second part of the proof of Proposition 4.5.1. When we fix $A = I$ and $X = I$ the ball $B_I(I, 1)$ is entirely contained in the positive semidefinite cone. This implies the radius of the ellipsoid must be at least $1 = \lambda_{min}(I \cdot I)$.

Using Weyl's perturbation theorem (Please see Theorem III.4.4 in Bhatia [5]), the statement for a general ϕ can be proved:

$$\phi(\lambda(A) - \lambda(Y)) \leq (\phi \circ \lambda)(A - Y) = ||A - Y||_\phi,$$

and suppose Y has a negative eigenvalue, i.e., $\lambda_i(Y) < 0$ for some i, then

$$\lambda_i(A) < \lambda_i(A) - \lambda_i(Y) \leq \phi(\lambda(A) - \lambda(Y)) \leq ||A - Y||_\phi \leq \lambda_{min}(A).$$

This is clearly a contradiction. Thus, if $||A - Y||_\phi \leq \lambda_{min}(A)$, $Y \succeq 0$.

Secondly, we will show $\alpha_A(\phi, X) = \alpha_{X^{-\frac{1}{2}}AX^{-\frac{1}{2}}}(\phi, I)$.

Notice $X^{-\frac{1}{2}} \cdot X^{-\frac{1}{2}}$ is an automorphism of the cone, and an ellipsoid is maximal if and only if it is maximal after applying an automorphism. Moreover, after some simple algebraic calculation we have:

$$X^{-\frac{1}{2}}(B_{\phi,X}(A, \alpha_A(\phi, X)))X^{-\frac{1}{2}} = B_{\phi,I}(X^{-\frac{1}{2}}AX^{-\frac{1}{2}}, \alpha_A(\phi, X)).$$

Hence, $\alpha_{X^{-\frac{1}{2}}AX^{-\frac{1}{2}}}(\phi, I) = \alpha_A(\phi, X)$.

Lastly, we show $\alpha_A(\phi, I) = \lambda_{min}(A)$. Since we have shown $\lambda_{min}(A)$ is a lower bound, we just need to show it is an upper bound as well.

This is again obvious when ϕ is the 2-norm. To see it is also an upper bound, we just need to notice the minimum distance from A to the boundary of the cone under the local norm defined by I is $\lambda_{min}(A)$. The boundary point that achieves the minimum distance from A is $\bar{A} = A - \lambda_{min}(A)vv^T$, where v is the normalized smallest eigenvector of A.

For general ϕ, the arguments are very similar. We just need to notice that

$$||vv^T||_\phi = ||\text{diag}(0, \ldots, 0, 1)||_\phi = \phi(0, \ldots, 0, 1) = 1.$$

Hence,

$$||A - \bar{A}||_\phi = ||\lambda_{min}(A)vv^T||_\phi = \lambda_{min}(A) ||vv^T||_\phi = \lambda_{min}(A)$$

This shows there exists a point $\bar{A}$ on the boundary of the cone that is $\lambda_{min}(A)$ away from A with the norm $||\cdot||_{\phi,I}$. As a result, $\alpha_A(\phi, I) \leq \lambda_{min}(A)$, and it implies $\alpha_A(\phi, I) = \lambda_{min}(A)$.

Together with the previous argument, we must have

$$\alpha_A(\phi, X) = \alpha_{X^{-\frac{1}{2}}AX^{-\frac{1}{2}}}(\phi, I) = \lambda_{min}(X^{-\frac{1}{2}}AX^{-\frac{1}{2}}) = \lambda_{min}(X^{-1}A).$$

$\square$

As a result, the minimum distance in terms of the standard norm from an interior point A in a positive semidefinite cone to its boundary under the local norm defined by the positive definite matrix X, i.e., $||Y||_X = \left|\left|X^{-\frac{1}{2}}YX^{-\frac{1}{2}}\right|\right|$, is equal to the minimum eigenvalue of $X^{-1}A$. Also, given an interior point $A \in \mathbb{S}^n_{++}$, and a self adjoint positive definite linear operator of the form $U \cdot U$ for some $U \in \mathbb{S}^n_{++}$, the largest radius r of the ellipsoid centered at A with the above shape and is contained in $\mathbb{S}^n_+$ is the minimum eigenvalue of UA.

The following result can be derived from the Gel'fand-Naimark theorem on singular values (Please see Theorem III.4.5 in Bhatia [5]):

Theorem 5.4.2. *Let $A, B \succ 0$, then*

$$\max_{X \succ 0} \lambda_{min}(AX^{-1})\lambda_{min}(XB) = \lambda_{min}(AB).$$

Theorem 5.4.3 (Lim [32])**.** *Let ϕ be any symmetric gauge norm and let $A, B \succ 0$. Then*

$$\max_{X \succ 0} \alpha_A(\phi, X)\alpha_B(\phi, X^{-1}) = \lambda_{min}(AB).$$

Proof. The result follows immediately from Theorem 5.4.1 and Theorem 5.4.2. □

Notice that for both result, the choice of the symmetric gauge norm does not affect the results at all.

It also turns out that the set:

$$
\begin{aligned}
T(A, B) \quad &:= \operatorname{argmax}_{X \succ 0} \alpha_A(\phi, X)\alpha_B(\phi, X^{-1}) \\
&= \operatorname{argmax}_{X \succ 0} \lambda_{min}(AX^{-1})\lambda_{min}(XB) \\
&= \{X \succ 0 : \lambda_{min}(AX^{-1})\lambda_{min}(XB) = \lambda_{min}(AB)\}
\end{aligned}
$$

of maximizers is a closed convex cone in the Hermitian matrix space independent of the symmetric gauge norms, and every point on the geodesic line with respect to the Finsler distance (see [6]) between A and B^{-1} achieves the maximum.

Chapter 6

Conclusion and future research

We studied inner approximations to convex cones in Euclidean spaces by ellipsoids. We considered two fundamental problems in this area:

1. Given a convex cone K, and an interior point x of K, what is the maximum volume ellipsoid centered at x and contained in K?

2. Given a convex cone K, a scalar product, and a pair of points $\bar{x} \in \text{int}(K)$, $\bar{s} \in \text{int}(K^*)$, what is the largest pair (in terms of product of their volumes) of dual ellipsoids that are centered at $\bar{x}$ and $\bar{s}$, and contained in K and K^* respectively?

The second fundamental problem above was proposed and solved by Todd [46] for all symmetric cones K.

In this thesis, we solve problem 1. above for second order cones (which was first solved by Güler and Gürtuna [19]) and cones of n-by-n symmetric positive semidefinite matrices. Our solutions are easily extendible to symmetric cones via employing the Euclidean Jordan algebra techniques.

However, generalizations of our results to more general classes of cones, such as homogeneous cones or hyperbolic cones seem to require yet more new techniques. These are left for future work.

Similarly, solution of problem 2. for classes of convex cones beyond the symmetric cones seem to require new techniques, and are also left for future work.

There are a few more questions relating to these two fundamental problems which could be of interest for future research:

Questions relating to the first problem:

Let $\mathbb{L} \subset \mathbb{S}^n$ be a linear subspace such that $\mathbb{S}^n_{++} \cap \mathbb{L} \neq \emptyset$. For which subspaces $\mathbb{L}$, can we "easily" utilize our knowledge of the maximum volume ellipsoid for $\mathbb{S}^n_+$ to obtain the maximum volume ellipsoid for $\mathbb{S}^n_+ \cap \mathbb{L}$?

For the more general version of this question, assume that K is an arbitrary closed convex cone with non-empty interior and $\mathbb{L}$ is a linear subspace such that $\text{int}(K) \cap \mathbb{L} \neq \emptyset$. Then, for which subspaces $\mathbb{L}$ can we "easily" utilize our knowledge of the maximum volume ellipsoid for K to obtain the maximum volume ellipsoid for $K \cap \mathbb{L}$?

Questions relating to the second problem:

In this thesis, we used the product of the radii or the product of the volumes of the pair of primal-dual ellipsoids as our measure of "largeness" for the pair. This measure corresponds to the geometric mean of the radii $\alpha_{\bar{x}}$ and $\alpha_{\bar{s}}$. If we were to use some other notions of mean instead, how would the solution change as a result?

How does the product of the volumes of the primal-dual pair of ellipsoids compare to that of the volumes of two ellipsoids taken individually as the largest volume ellipsoids inscribed in the primal and dual cones respectively?

Let F be the standard self-concordant barrier function for a positive semidefinite cone. By Lim [32], $F''(g)$ and $[F''(g)]^{-1}$, where g is on the geodesic line with respect to the Finsler distance between $\bar{x}$ and $z = -F'_*(\bar{s})$, define the largest pair of primal-dual ellipsoids centered at $\bar{x}$ and $\bar{s}$ contained in the respective cones. Would this result be true in an analogous sense in the more general setting of homogeneous or hyperbolic cones?

References

[1] Keith Ball. Ellipsoids of maximal volume in convex bodies. *Geom. Dedicata*, 41(2):241–250, 1992.

[2] Aharon Ben-Tal, Laurent El Ghaoui, and Arkadi Nemirovski. *Robust optimization.* Princeton University Press, 2009.

[3] Aharon Ben-Tal and Arkadi Nemirovski. *Lectures on modern convex optimization.* MPS/SIAM Series on Optimization. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA; Mathematical Programming Society (MPS), Philadelphia, PA, 2001. Analysis, algorithms, and engineering applications.

[4] Aharon Ben-Tal and Arkadi Nemirovski. Robust optimization–methodology and applications. *Mathematical Programming*, 92(3):453–480, 2002.

[5] Rajendra Bhatia. *Matrix analysis*, volume 169 of *Graduate Texts in Mathematics.* Springer-Verlag, New York, 1997.

[6] Rajendra Bhatia. On the exponential metric increasing property. *Linear Algebra Appl.*, 375:211–220, 2003.

[7] Stephen Boyd, Laurent El Ghaoui, Eric Feron, and Venkataramanan Balakrishnan. *Linear matrix inequalities in system and control theory*, volume 15 of *SIAM Studies in Applied Mathematics.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1994.

[8] Stephen Boyd and Lieven Vandenberghe. *Convex optimization.* Cambridge University Press, Cambridge, 2004.

[9] Chek Beng Chua. The primal-dual second-order cone approximations algorithm for symmetric cone programming. *Found. Comput. Math.*, 7(3):271–302, 2007.

[10] Andrew R. Conn, Nicholas I. M. Gould, and Philippe L. Toint. *Trust-region methods.* MPS/SIAM Series on Optimization. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA; Mathematical Programming Society (MPS), Philadelphia, PA, 2000.

[11] Christophe Croux, Gentiane Haesbroeck, and Peter J. Rousseeuw. Location adjustment for the minimum volume ellipsoid estimator. *Stat. Comput.*, 12(3):191–200, 2002.

[12] Ludwig Danzer, Detlef Laugwitz, and Hanfried Lenz. Über das Löwnersche Ellipsoid und sein Analogon unter den einem Eikörper einbeschriebenen Ellipsoiden. *Arch. Math. (Basel)*, 8:214–219, 1957.

[13] Alan Edelman and N. Raj Rao. Random matrix theory. *Acta Numer.*, 14:233–297, 2005.

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

[15] M. Seetharama Gowda, Roman Sznajder, and Jiyuan Tao. The automorphism group of a completely positive cone and its Lie algebra. *Linear Algebra Appl.*, 438(10):3862–3871, 2013.

[16] Peter M. Gruber. Minimal ellipsoids and their duals. *Rend. Circ. Mat. Palermo (2)*, 37(1):35–64, 1988.

[17] Peter M. Gruber. John and Loewner ellipsoids. *Discrete Comput. Geom.*, 46(4):776–788, 2011.

[18] Osman Güler and Filiz Gürtuna. The extremal volume ellipsoids of convex bodies, their symmetry properties, and their determination in some special cases. *arXiv preprint arXiv:0709.0707*, 2007.

[19] Osman Güler and Filiz Gürtuna. Symmetry of convex sets and its applications to the extremal ellipsoids of convex bodies. *Optim. Methods Softw.*, 27(4-5):735–759, 2012.

[20] Filiz Gürtuna. Duality of ellipsoidal approximations via semi-infinite programming. *SIAM J. Optim.*, 20(3):1421–1438, 2009.

[21] L. G. Hačijan. A polynomial algorithm in linear programming. *Dokl. Akad. Nauk SSSR*, 244(5):1093–1096, 1979.

[22] L. G. Hačijan. Polynomial algorithms in linear programming. *Zh. Vychisl. Mat. i Mat. Fiz.*, 20(1):51–68, 260, 1980.

[23] Martin Henk. Löwner-John ellipsoids. *Doc. Math.*, (Extra volume: Optimization stories):95–106, 2012.

[24] Jean-Baptiste Hiriart-Urruty and Claude Lemaréchal. *Fundamentals of convex analysis.* Grundlehren Text Editions. Springer-Verlag, Berlin, 2001. Abridged version of ıt Convex analysis and minimization algorithms. I [Springer, Berlin, 1993; MR1261420 (95m:90001)] and ıt II [ibid.; MR1295240 (95m:90002)].

[25] Fritz John. Extremum problems with inequalities as subsidiary conditions. In *Studies and Essays Presented to R. Courant on his 60th Birthday, January 8, 1948*, pages 187–204. Interscience Publishers, Inc., New York, N. Y., 1948.

[26] D. B. Judin and A. S. Nemirovskiĭ. Informational complexity and effective methods for the solution of convex extremal problems. *Èkonom. i Mat. Metody*, 12(2):357–369, 1976.

[27] William Karush. *MINIMA OF FUNCTIONS OF SEVERAL VARIABLES WITH INEQUALITIES AS SIDE CONDITIONS*. ProQuest LLC, Ann Arbor, MI, 1939. Thesis (SM)–The University of Chicago.

[28] Masakazu Kojima and Makoto Yamashita. Enclosing ellipsoids and elliptic cylinders of semialgebraic sets and their application to error bounds in polynomial optimization. *Math. Program.*, 138(1-2, Ser. A):333–364, 2013.

[29] H. W. Kuhn and A. W. Tucker. Nonlinear programming. In *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, 1950*, pages 481–492. University of California Press, Berkeley and Los Angeles, 1951.

[30] P. Kumar and E. A. Yildirim. Minimum-volume enclosing ellipsoids and core sets. *J. Optim. Theory Appl.*, 126(1):1–21, 2005.

[31] Jean B. Lasserre. A generalization of Löwner-John's ellipsoid theorem. *Math. Program.*, 152(1-2, Ser. A):559–591, 2015.

[32] Yongdo Lim. Maximum-volume symmetric gauge ball problem on the convex cone of positive definite matrices and convexity of optimal sets. *SIAM J. Optim.*, 21(4):1275–1288, 2011.

[33] Yongdo Lim. Todd's maximum-volume ellipsoid problem on symmetric cones. *Forum Math.*, 23(2):413–426, 2011.

[34] Igor S. Litvinchev. A circular cone relaxation primal interior point algorithm for LP. *Optimization*, 52(4-5):529–540, 2003. Theory, methods and applications of optimization.

[35] Yurii Nesterov and Arkadii Nemirovskii. *Interior-point polynomial algorithms in convex programming*, volume 13 of *SIAM Studies in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1994.

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

[37] Yurii Nesterov and Michael J. Todd. Primal-dual interior-point methods for self-scaled cones. *SIAM J. Optim.*, 8(2):324–364, 1998.

[38] Yurii Nesterov and Michael J. Todd. On the Riemannian geometry defined by self-concordant barriers and interior-point methods. *Found. Comput. Math.*, 2(4):333–361, 2002.

[39] Ting Kei Pong and Henry Wolkowicz. The generalized trust region subproblem. *Comput. Optim. Appl.*, 58(2):273–322, 2014.

[40] James Renegar and Mutiara Sondjaja. A polynomial-time affine-scaling method for semidefinite and hyperbolic programming. *arXiv preprint arXiv:1410.6734*, 2014.

[41] R. Tyrrell Rockafellar. *Convex analysis*. Princeton Mathematical Series, No. 28. Princeton University Press, Princeton, N.J., 1970.

[42] J. Saunderson, P. A. Parrilo, and A. S. Willsky. Semidefinite descriptions of the convex hull of rotation matrices. *SIAM J. Optim.*, 25(3):1314–1343, 2015.

[43] Romy Shioda and Levent Tunçel. Clustering via minimum volume ellipsoids. *Comput. Optim. Appl.*, 37(3):247–295, 2007.

[44] N. Z. Šor. Use of the space expansion operation in problems of convex function minimalization. *Kibernetika (Kiev)*, (1):6–12, 1970.

[45] Peng Sun and Robert M. Freund. Computation of minimum-volume covering ellipsoids. *Oper. Res.*, 52(5):690–706, 2004.

[46] M. J. Todd. Largest dual ellipsoids inscribed in dual cones. *Math. Program.*, 117(1-2, Ser. B):425–434, 2009.

[47] Van Anh Truong and Levent Tunçel. Geometry of homogeneous convex cones, duality mapping, and optimal self-concordant barriers. *Math. Program.*, 100(2, Ser. A):295–316, 2004.

[48] Levent Tunçel. *Polyhedral and semidefinite programming methods in combinatorial optimization*, volume 27 of *Fields Institute Monographs*. American Mathematical Society, Providence, RI; Fields Institute for Research in Mathematical Sciences, Toronto, ON, 2010.

[49] E. Alper Yildirim. On the minimum volume covering ellipsoid of ellipsoids. *SIAM J. Optim.*, 17(3):621–641, 2006.

[50] Yin Zhang and Liyan Gao. On numerical solution of the maximum volume ellipsoid problem. *SIAM J. Optim.*, 14(1):53–76 (electronic), 2003.