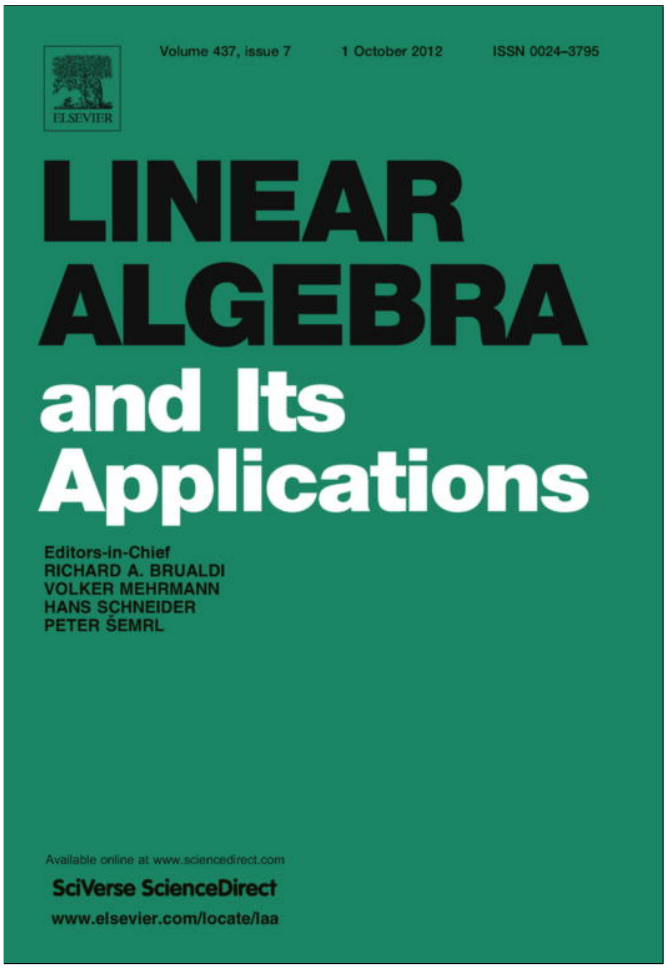
Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

http://www.elsevier.com/copyright

Author's personal copy

Linear Algebra and its Applications 437 (2012) 1562–1578

Contents lists available at SciVerse ScienceDirect

Linear Algebra and its Applications

journal homepage: www.elsevier.com/locate/laa

Computational geometry of positive definiteness

Marko Huhtanen [*,1], Otto Seiskari [2]

Department of Mathematics and Systems Analysis, Aalto University, Box 1100, FIN-02015, Finland

ARTICLE INFO

Article history:
Received 13 April 2011
Accepted 24 April 2012
Available online 28 May 2012

Submitted by R.A. Brualdi

AMS classifications:
15B48
47L25

Keywords:
Hermitian matrix subspace
Positive definiteness
Joint numerical range
Computational geometry
Eigenvalue optimization
Convex analysis

ABSTRACT

In matrix computations, such as in factoring matrices, Hermitian and, preferably, positive definite elements are occasionally required. Related problems can often be cast as those of existence of respective elements in a matrix subspace. For two dimensional matrix subspaces, first results in this regard are due to Finsler. To assess positive definiteness in larger dimensional cases, the task becomes computational geometric for the joint numerical range in a natural way. The Hermitian element of the Frobenius norm one with the maximal least eigenvalue is found. To this end, extreme eigenvalue computations are combined with ellipsoid and perceptron algorithms.

© 2012 Elsevier Inc. All rights reserved.

1. Introduction

Appearing in diverse applications, positive definiteness is a central notion for square matrices and operators; see, e.g. [19,5]. For related computational matters, see [15, Chapter 4.2]. For matrix subspaces, the concept of positive definiteness is a far more delicate issue. Matrix subspaces with Hermitian and, preferably, positive definite elements arise in factoring problems and in large scale numerical linear algebra of preconditioning [21,9]. [3] In both cases, the existence of such elements reflects fundamental aspects of operators. The challenge with matrix subspaces lies, not least computationally, in the fact that the subset of positive definite matrices can be a tiny, needle-like set.

* Corresponding author.
E-mail addresses: Marko.Huhtanen@hut.fi (M. Huhtanen), oseiskar@cc.hut.fi (O. Seiskari).

[1] Supported by the Academy of Finland.
[2] Partially supported by the Academy of Finland.

[3] The study of matrix subspaces can be classified as being part of operator space theory [11,26].

0024-3795/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.laa.2012.05.002

Author's personal copy

This paper is concerned with making this computational geometrically more quantitative. Ways to locate positive definite elements are devised. In a certain sense the most positive definite element is found.

Focusing on two dimensional matrix subspaces, first results regarding the existence of positive definite elements are due to Finsler [12]. (For related computations, see [10,18].) The three dimensional case is related with the investigations of Binding [6]. Later, in semidefinite programming, a similar task defines the feasibility problem of semidefinite programs [27]. Quantitatively, for inclusion regions, we employ strictly positive maps of the simplest possible type. Denote by $\mathcal{V}$ a matrix subspace of $\mathbb{C}^{n\times n}$ over $\mathbb{R}$ whose elements are Hermitian. In terms of an orthonormal basis $V_1, \ldots, V_k$ of $\mathcal{V}$, this leads us to consider the map

$$x \longmapsto (x^*V_1x, \ldots, x^*V_kx) \quad \text{with } ||x|| = 1 \tag{1.1}$$

whose image, i.e., the joint numerical range, is seemingly the most tangible object to study positive definiteness of $\mathcal{V}$. Traditionally, its convexity has been an object of interest; see [17] and references therein. It is noteworthy that the convexity of the image in a basis of $\mathcal{V}$ implies convexity in any of its basis. Thereby we are primarily dealing with a property of the matrix subspace $\mathcal{V}$ rather than that of the map (1.1).

In this paper we devise methods to approximate the joint numerical range with a small number of half-spaces.[4] In this sense the problem becomes computational geometric. In particular, $\mathcal{V}$ possessing positive definite elements is information expressible in terms of a single half-space whose boundary contains the origin. This interpretation leads to the notion of most positive definite element of $\mathcal{V}$ expressed by the hyperplane containing the origin which is farthest away from the joint numerical range. To generate half-spaces, we use the fact that the structure of maps of the form (1.1) is invariant under orthogonal transformations. This combined with extreme eigenvalue computations to locate boundary points of the convex hull of the image yields relatively sharp information on the location of the image.

It is a natural task to find the distance of the joint numerical range from the origin, yielding an orthogonal invariant of $\mathcal{V}$. We solve the problem for the convex hull of the image. Equivalently, we look for a positive definite element of $\mathcal{V}$ of the Frobenius norm one having the maximal least eigenvalue. Two algorithms proposed to solve the problem are based on the ellipsoid algorithm used in convex optimization. These methods allow us to locate the most positive definite element in the prescribed sense and to compute the distance of the convex hull of the joint numerical range from the origin. For the easier feasibility problem of locating a positive definite element in $\mathcal{V}$, the perceptron algorithm is suggested as a simpler alternative.

The paper is organized as follows. In Section 2 fundamentals of Hermitian matrix subspaces are presented, including examples involving positive definiteness. In Section 3 geometric aspects of locating positive definite elements of a Hermitian matrix subspace are developed. Algorithms to solve the maximal least eigenvalue problem for locating positive definite elements are devised in Section 4. In Section 5 numerical experiments are presented to illustrate the performance of the algorithms. In Appendix A the classical case of $\dim \mathcal{V} = 2$ and Finsler's result are covered. Related problems involving positive definite matrices are discussed in Appendix B.

2. Hermitian matrix subspaces and positive definiteness

Denote by $\mathcal{V}$ a matrix subspace of $\mathbb{C}^{n\times n}$ over $\mathbb{R}$.[5] To avoid confusion, throughout the paper also $\mathbb{C}^{n\times n}$ is regarded as a vector space over $\mathbb{R}$. Then in $\mathcal{V}$ the standard inner product

$$(V, W) = \operatorname{Re} \operatorname{tr} W^*V \tag{2.1}$$

[4] A half-space (open) in $\mathbb{R}^k$ is defined as $\{(t_1, \ldots, t_k) : u_1t_1 + \cdots + u_kt_k > c\}$ for a fixed $(u_1, \ldots, u_k) \in \mathbb{R}^k$ and $c \in \mathbb{R}$.

[5] If $\mathcal{V}$ is initially a matrix subspace of $\mathbb{C}^{n\times n}$ over $\mathbb{C}$, then it can be treated as a matrix subspace over $\mathbb{R}$ by doubling the dimension.

Author's personal copy

is used. The respective Frobenius norm is denoted by $||\cdot||_F$. Regarding our interests, the so-called nonsingular matrix subspaces are of central relevance [21,9]. A matrix subspace is said to be nonsingular if it contains invertible elements. Among matrix subspaces, nonsingularity is a generic property [22].

For additional properties, the set of Hermitian matrices $\mathcal{H}$ is of dimension n^2 in $\mathbb{C}^{n\times n}$. The notion of Hermitian matrix subspace is defined in a natural way as follows.

Definition 2.1. A matrix subspace $\mathcal{V}$ of $\mathbb{C}^{n\times n}$ over $\mathbb{R}$ is Hermitian if all its elements are Hermitian.

The Hermitian elements of a matrix subspace $\mathcal{V}$ can be readily recovered by computing the nullspace of the linear map

$$V \longmapsto V - V^* \tag{2.2}$$

from $\mathcal{V}$ to $\mathbb{C}^{n\times n}$. We call this nullspace the Hermitian matrix subspace of $\mathcal{V}$.[6]

Equivalence is a fundamental notion for matrix subspaces which can be regarded as a relaxation. Namely, matrix subspaces $\mathcal{V}$ and $\mathcal{W}$ are said to be equivalent if there exist invertible matrices $X, Y \in \mathbb{C}^{n\times n}$ such that

$$\mathcal{W} = X\mathcal{V}Y^{-1}.$$

Clearly, the Hermitian structure is preserved in congruence, i.e., when $Y^{-1} = X^*$. For a necessary and sufficient condition on a matrix subspace $\mathcal{V}$ to be equivalent to a Hermitian matrix subspace, suppose $V_1, \ldots, V_k$ is its basis. Consider the problem of finding out, whether the matrices $XV_1Y^{-1}, \ldots, XV_kY^{-1}$ are Hermitian for some invertible matrices X and Y. To solve this, compute the intersection of the nullspaces of the linear maps

$$M \longmapsto V_jM - M^*V_j^*$$

on $\mathbb{C}^{n\times n}$, for $j = 1, \ldots, k$. If there exists an invertible element M in the intersection, then X and Y are determined by the condition $Y^{-1}X^{-*} = M$. When $k = 2$, this problem is of interest in the generalized eigenvalue problem.

Denote by $\mathcal{S}_{++}$ the convex cone of positive definite matrices in $\mathbb{C}^{n\times n}$. (See, e.g. [2, II Sec. 12–15] for the convexity of $\mathcal{S}_{++}$.) As is well-known, $\mathcal{S}_{++}$ and its closure are of importance in convex optimization, see, e.g. [7].

Definition 2.2. A Hermitian matrix subspace $\mathcal{V}$ is said to possess positive definite elements if $\mathcal{V} \cap \mathcal{S}_{++} \neq \emptyset$.

For a classical two dimensional example, consider the generalized eigenvalue problem. Then it is of central relevance to know if the associated matrix subspace possesses positive definite elements; see [25, Chapter 15.3]. Further examples follow.

Example 1. Denote by $\mathcal{H}$ the set of Hermitian matrices. An invertible matrix $A \in \mathbb{C}^{n\times n}$ is the product of a Hermitian matrix and a positive definite matrix if and only if the Hermitian subspace of $\mathcal{V} = A^{-1}\mathcal{H}$ contains positive definite elements. (Of course, A is Hermitian if and only if $\mathcal{V}$ contains the identity.) This is a classical notion, such a matrix is said to be symmetrizable [4, p. 67].

Example 2. In view of preconditioning very large linear systems, assume having an invertible sparse matrix $A \in \mathbb{C}^{n\times n}$. Consider the homogeneous linear system

$$AW - W^*A^* = 0. \tag{2.3}$$

[6] To fully measure how much $\mathcal{V}$ deviates from being a Hermitian matrix subspace, inspect the singular values of the linear map (2.2).

Author's personal copy

We are interested in those solutions $W \in \mathbb{C}^{n\times n}$ which are sparse.[7] Let $\mathcal{W}$ denote their span. Then one looks for the positive definite elements of the matrix subspace $A\mathcal{W}$ by the fact that with the respective products, the conjugate gradient [15, Chapter 10.2] method can be executed.

It is noteworthy that Hermitian matrix subspaces possessing no positive definite matrices can be high dimensional. For example, consider a matrix subspace whose every member has the $(1, 1)$-entry equaling zero.

Whenever nonempty, $\mathcal{V} \cap \mathcal{S}_{++}$ is an open subset of $\mathcal{V}$ by the fact that if $V \in \mathcal{V} \cap \mathcal{S}_{++}$, then $V + E \in \mathcal{V} \cap \mathcal{S}_{++}$ for $E \in \mathcal{V}$ small enough in norm.[8] Hence the convex cone $\mathcal{V} \cap \mathcal{S}_{++}$ is a submanifold of $\mathcal{V}$ of the same dimension. This is useful, although hardly completely satisfactory information in practice.

Example 3. The set of diagonal Hermitian matrices in $\mathbb{C}^{3\times 3}$ is isometrically isomorphic to $\mathbb{R}^3$ in a natural way. The positive definite elements correspond to

$$\{(d_1, d_2, d_3) \in \mathbb{R}^3 \,:\, d_j > 0,\ j = 1, 2, 3\}. \tag{2.4}$$

Let $\mathcal{V}$ be a two dimensional subspace of $\mathbb{R}^3$ (i.e., a plane through the origin) whose intersection with (2.4) is a sharp needle-like set.

To express this in terms of a volume, on the tangent spaces of $\mathcal{V} \cap \mathcal{S}_{++}$ we employ the standard inner product (2.1). Thereby $\mathcal{V} \cap \mathcal{S}_{++}$ is regarded as a Riemannian submanifold of $\mathbb{C}^{n\times n}$.[9] Because the intersection can be a very small set, any purely random process to decide whether $\mathcal{V}$ possesses positive definite elements is highly unlikely to be successful. It is informative, for comparison, to bear in mind that the set of Hermitian matrices is of dimension n^2 in $\mathbb{C}^{n\times n}$ of which $\mathcal{S}_{++}$ occupies just a $\frac{1}{2^n}$ portion.

The question of existence of positive semidefinite elements and estimating their volume can be turned into a problem in real algebraic geometry. For the minimum dimension of the underlying space, denote by $\mathcal{L}$ the set of lower triangular matrices with real diagonal entries, regarded as a subspace of $\mathbb{C}^{n\times n}$ over $\mathbb{R}$ of dimension n^2.

Theorem 2.3. *To the set of positive semidefinite elements of a Hermitian matrix subspace $\mathcal{V} \subset \mathbb{C}^{n\times n}$ corresponds a real homogeneous variety of $\mathcal{L} \subset \mathbb{C}^{n\times n}$.*

Proof. By the Cholesky factorization, a Hermitian matrix H is positive definite if and only if $H = LL^*$ for a lower triangular matrix with a positive diagonal. Moreover, if L is lower triangular, it readily seen that LL^* positive definite if and only if L has nonzero diagonal entries. Otherwise LL^* is positive semidefinite.

For the construction, with respect to the inner product (2.1), denote by $\mathbf{P}$ the orthogonal projector on $\mathcal{H}$ onto $\mathcal{V}$. To characterize the positive semidefinite elements of $\mathcal{V}$, define

$$L \longmapsto (I - \mathbf{P})LL^* \tag{2.5}$$

from $\mathcal{L}$ to $\mathcal{H}$. This equals zero if and only if LL^*, which is positive semidefinite, belongs to $\mathcal{V}$. Let $M_1, \ldots, M_l$ be an orthonormal basis of the orthogonal complement of $\mathcal{V}$ in $\mathcal{H}$. Then L is mapped to zero by (2.5) if and only if

$$(LL^*, M_j) = 0 \quad \text{for } j = 1, \ldots, l. \tag{2.6}$$

[7] Clearly, there are sparse solutions as $W = A^*$ illustrates. Then AA^* leads to the normal equations which is not attractive because of the squaring of the condition number. Therefore other sparse solutions are more of interest.

[8] For not too large n, to test whether $V \in \mathcal{V} \cap \mathcal{S}_{++}$, it is advisable to attempt to compute a Cholesky factorization [15, p. 146].

[9] It is somewhat exceptional to use the inner product (2.1) with the manifold of positive definite matrices. For the usual Riemannian geometry of nonpositive curvature, see [5, Chapter 6].

Author's personal copy

Since these are homogeneous polynomial maps of degree two in the entries of L separated into real and imaginary parts, to the positive semi-definite elements of $\mathcal{V}$ corresponds a homogeneous variety of $\mathcal{L}$. $\square$

There are $n^2 - \dim \mathcal{V}$ equations. (For computational aspects of real algebraic geometry, see [3].) As an extreme, the corresponding variety is the whole $\mathcal{L}$ if and only if $\mathcal{V} = \mathcal{H}$. Involving n^2 real variables, solving (2.6) does not appear very realistic unless n is small. On the positive side, though, the degrees of the polynomial equations are just two.

The construction has the advantage that the elements mapped from this variety according to $L \mapsto LL^*$ to the set of positive definite elements are immediately recovered.

Corollary 2.4. *$\mathcal{V}$ does not possess positive definite elements if and only if the variety contains only singular elements.*

3. Locating positive definite elements geometrically

There are several necessary and sufficient conditions guaranteeing positive definiteness of a Hermitian matrix [19, Chapter 7]. For a Hermitian matrix subspace $\mathcal{V}$, an analogous problem consists of locating positive definite elements, if any. (If $\mathcal{V}$ is not Hermitian, then start by computing its Hermitian matrix subspace.) As just described, with matrix subspaces the challenge lies in the fact that the subset of positive definite matrices can be needle-like.

3.1. Positive definiteness and the joint numerical range

To locate possible positive definite elements for $k > 2$, an approach can be based on polynomial inequalities. (For $k = 2$, see Appendix A.) To this end, suppose $V_1, \ldots, V_k$ is a basis of a Hermitian matrix subspace $\mathcal{V}$ and set

$$V \equiv V(t_1, \ldots, t_k) = t_1 V_1 + \cdots + t_k V_k \tag{3.1}$$

with $t_j \in \mathbb{R}$ for $j = 1, \ldots, k$. A Hermitian matrix is positive definite if and only if all its leading principal minors are positive; see, e.g. [19, p. 404]. [10] This gives rise to n polynomial inequalities in the parameters $t_1, \ldots, t_k$ for determining $\mathcal{V} \cap \mathcal{S}_{++}$.

Clearly, even for moderate n, dealing with the determinants of large leading principal submatrices is computationally very unappealing. In particular, it is not the simplest way to inspect the structure of $\mathcal{V} \cap \mathcal{S}_{++}$.

Regions including $\mathcal{V} \cap \mathcal{S}_{++}$ can be determined more economically with the help of strictly positive maps. For matrix analysis of positive maps, see [5, Chapter 2] and references therein.

Definition 3.1. A linear map $\Phi : \mathbb{C}^{n\times n} \to \mathbb{C}^{l\times l}$ is strictly positive if $\Phi(A)$ is positive definite whenever A is.

For example, the linear map on $\mathbb{C}^{n\times n}$ to any leading principal submatrix is strictly positive.

Combine (3.1) with a strictly positive linear map Φ and set

$$(t_1, \ldots, t_k) \longmapsto \det \Phi(V(t_1, \ldots, t_k)). \tag{3.2}$$

This is a homogeneous polynomial of degree l. For the volumes, it is natural to choose the basis $V_1, \ldots, V_k$ to be orthonormal, so that the linear map (3.1) yields an isometric isomorphism between the parameter space $\mathbb{R}^k$ and $\mathcal{V}$. This is assumed in what follows. To have regions including $\mathcal{V} \cap \mathcal{S}_{++}$, we are interested in those parameter values for which the function (3.2) is positive.

For linear inequalities, inexpensive to generate, consider the strictly positive linear map $\Phi_x(A) = x^*Ax$ for any fixed $x \in \mathbb{C}^n$. This inserted into (3.2) gives rise to the open half-space in $\mathbb{R}^k$ (through the origin) defined as

[10] Also called Sylvester's criterion.

Author's personal copy

$$\{(t_1, \ldots, t_k) \ : \ \sum_{j=1}^{k} (x^* V_j x)\, t_j > 0\}. \tag{3.3}$$

For a necessary condition the parameters must satisfy, any intersection of such half-planes yields an unbounded convex polytope in the parameter space (as long as there are no contradicting inequalities).

What is the best we can do with a small number of half-spaces? With an orthonormal basis $V_1, \ldots, V_k$ of $\mathcal{V}$, define

$$\mathbf{V}(x) = (x^* V_1 x, \ldots, x^* V_k x) \tag{3.4}$$

for $x \in \mathbb{S}^{2n-1} = \{x \in \mathbb{C}^n \ : \ ||x|| = 1\}$. Clearly, $\mathbf{V}$ is smooth having values in $\mathbb{R}^k$. Without any loss of generality, we allow only orthonormal bases of $\mathcal{V}$. Then we have orthogonal invariance in the sense that $U\mathbf{V}(x)$, where $U \in \mathbb{R}^{k\times k}$ is an orthogonal matrix, is of the same form as (3.4) but in another orthonormal basis of $\mathcal{V}$. And conversely, to any orthonormal basis corresponds such an orthogonal matrix U.

Example 4. Let V_j be simultaneously unitarily diagonalizable. (That is, UV_jU^* are diagonal for a unitary matrix U, for $j = 1, \ldots, k$.) Then the image of (3.4) is a convex polytope.

The image of $\mathbf{V}$ is called the joint numerical range of the matrices $V_1, \ldots, V_k$. With respect to the image, a single half-space determines the existence of positive definite elements as follows.

Proposition 3.2. *A Hermitian matrix subspace $\mathcal{V}$ possesses positive definite elements if and only if the image of* (3.4) *is contained in an open half-space whose boundary contains the origin.*

Proof. Suppose there is a positive definite linear combination $\sum_{j=1}^{k} u_j V_j$. Assume $\sum_{j=1}^{k} u_j^2 = 1$. Then consider $U\mathbf{V}(x)$ where $U \in \mathbb{R}^{k\times k}$ is an orthogonal matrix having $(u_1, \ldots, u_k)$ as its first row. This is just (3.4) represented in another orthonormal basis. By construction, its first component is strictly positive, so that the image is contained in an open half-space.

For the converse, if the image is contained in an open half-space, then for some unit vector $(u_1, \ldots, u_k)$ and for every nonzero x holds $\sum_{j=1}^{k} u_j\, x^* V_j x > 0$. Thereby the linear combination $\sum_{j=1}^{k} u_j V_j$ is positive definite. □

For a unit vector $(u_1, \ldots, u_k)$, suppose the open half-space

$$\{(t_1, \ldots, t_k) \ : \ u_1 t_1 + \cdots + u_k t_k > 0\}$$

contains the image of $\mathbf{V}$. Then the positive definite matrix $V = \sum_{j=1}^{k} u_j V_j$ is said to correspond to the hyperplane $\{(t_1, \ldots, t_k) \ : \ u_1 t_1 + \cdots + u_k t_k = 0\}$. This interpretation yields a way to define the most positive definite element in terms of the hyperplane farthest away from the joint numerical range.

Definition 3.3. Assume a Hermitian matrix subspace $\mathcal{V}$ possesses positive definite elements. The most positive definite element corresponds to the hyperplane containing the origin having the maximum distance from the image of $\mathbf{V}$.[11]

By the fact that the image of $\mathbf{V}$ is connected, the hyperplane in question is outside the convex hull of the image. Thereby the notion is well-defined. Interpreted in terms of the corresponding positive definite elements, the most positive definite element of the definition is hence defined as being the one solving the minimization problem

[11] This means maximizing $\min_{v\in\text{image}(\mathbf{V})} \min_{h\in H} ||v - h||$ over hyperplanes H containing the origin.

Author's personal copy

$$\begin{aligned} \max\ & \lambda_{\min}(V) \\ \text{s.t. } & V \in \mathcal{V} \\ & V \geqslant 0 \quad (\text{equivalent to } \lambda_{\min}(V) \geqslant 0) \\ & \|V\|_F = 1, \end{aligned}$$

where $\lambda_{\min}(V)$ denotes the smallest eigenvalue of a Hermitian matrix V. Consequently, we can alternatively regard our computational geometric problem as a problem in eigenvalue optimization. (For eigenvalue optimization, see [24].)

In the case of $\dim \mathcal{V} = 2$ we are dealing with the numerical range of a matrix. Admitting many extensions, (3.4) is among them [20, pp. 85–87] being perhaps the most natural one (except that no assumptions on orthonormality are made). Traditionally, its convexity has been an object of interest, leading to the respective notion for matrix subspaces.

Definition 3.4. A Hermitian matrix subspace $\mathcal{V}$ is said to have convex numerical range if the image of the map (3.4) is convex.

This is well-defined by the fact that the image of (3.4) being convex in a basis $V_1, \ldots, V_k$ of $\mathcal{V}$ is necessary and sufficient for being convex in any basis of $\mathcal{V}$. This follows from composing $M\mathbf{V}(x)$ with any invertible matrix $M \in \mathbb{R}^{k\times k}$ and recovering the corresponding map (3.4).

Establishing convexity is, in general, a tough problem for $\dim \mathcal{V} > 2$. In the case of $\dim \mathcal{V} = 3$ there are some results [20, p. 86]. An interesting open problem (not considered here) is to identify cases in which $\mathcal{V}$ having convex numerical range is a generic property among Hermitian subspaces of the same dimension in $\mathbb{C}^{n\times n}$. The following is a consequence of Proposition 3.2.

Proposition 3.5. *Suppose a Hermitian matrix subspace $\mathcal{V}$ has convex numerical range. Then vanishing of (3.4) at a point is a necessary and sufficient condition on $\mathcal{V}$ not to possess positive definite elements.*

For a Hermitian matrix subspace $\mathcal{V}$, set

$$v(\mathcal{V}) = \min_{||x||=1} ||\mathbf{V}(x)||, \tag{3.5}$$

i.e., the distance of the image of $\mathbf{V}$ from the origin requiring the basis be orthonormal. Whether or not $\mathcal{V}$ has convex numerical range, this is certainly a quantity of interest. (The Crawford number[12] for two, not necessarily orthonormal, Hermitian matrices is defined analogously.) In a way, $v(\mathcal{V})$ yields an opposite of the numerical radius which would correspond to taking the maximum instead. Recall that the numerical radius of a matrix $A \in \mathbb{C}^{n\times n}$ is $w(A) = \max_{\lambda\in F(A)} |\lambda|$, where $F(A)$ denotes the numerical range of A.[13] (For its computation, see [29].)

A minimizer yields a good candidate for constructing a positive definite element, yielding an optimal solution in the following case.

Theorem 3.6. *Suppose a Hermitian matrix subspace $\mathcal{V}$ has convex numerical range. If a unit vector $x \in \mathbb{C}^n$ satisfies $v(\mathcal{V}) = ||\mathbf{V}(x)|| > 0$, then*

$$\sum_{j=1}^{k} \frac{x^* V_j x}{v(\mathcal{V})} V_j \tag{3.6}$$

is the most positive definite element of $\mathcal{V}$.

[12] Called the Crawford number of a Hermitian pair.

[13] As a curiosity, a result by T. Ando states $w(A) < 1$ if and only if the $\mathcal{V} \cap \mathcal{S}_{++} \neq \emptyset$ for a certain Hermitian matrix subspace. See [5, Theorem 3.5.1].

Author's personal copy

Proof. Take the unit vector with the components $u_j = \frac{x^*V_jx}{v(\mathcal{V})}$. Then, because of convexity, $\mathbf{V}(x)$ is the nearest point to the hyperplane containing the origin with this normal vector. Consequently, because of convexity, this hyperplane is farthest away from the joint numerical range. □

In the next section an algorithm for computing the distance of the convex hull of the image of $\mathbf{V}$ from the origin is devised. Thus, $v(\mathcal{V})$ is computable in case $\mathcal{V}$ has convex numerical range. Otherwise we obtain a lower bound which still suffices for locating the most positive definite element.

Let us emphasize that computing values of $\mathbf{V}$ at random points of $\mathbb{S}^{2n-1}$ can be expected to approximate the image poorly. (Numerical experiments support this in the case of numerical range of a matrix.) The boundary of the image is more accessible. This is due to the fact that for the convex hull of the image we can find support planes by computing extreme eigenvalues and corresponding eigenvectors of Hermitian matrices; see Algorithm 1. Recall that a support plane of a closed set has at least one common point with the set such that the entire set lies in one of the two half-spaces determined by the plane.

Algorithm 1 Computing a boundary point of the image of $\mathbf{V}$.

1: Choose a unit vector $u = (u_1, \ldots, u_k)$ and set $V = \sum_{j=1}^k u_jV_j$
2: Compute an extreme eigenvalue and respective unit eigenvector x of V
3: Set $p = (x^*V_1x, \ldots, x^*V_kx)$

Observe that the vector p is on that part of the boundary of the image of $\mathbf{V}$ which intersects the boundary of the convex hull of the image of $\mathbf{V}$.

In the algorithm, there are two alternatives for the extreme; either the smallest or the largest eigenvalue of V. We denote by $\lambda(u)$ the smallest. (Clearly, $\lambda(u) > 0$ if and only if V is positive definite.) In both cases,

$$\left\{(t_1, \ldots, t_k) \ : \ \sum_{j=1}^k u_j(t_j - p_j) = 0\right\} \tag{3.7}$$

yields a support plane of the image of $\mathbf{V}$; see Fig. 3.1(a). It is noteworthy that with the Hermitian Lanczos method, numerical computation of the extreme eigenvalues and corresponding eigenvectors is inexpensive for sparse matrix subspaces.[14] For the Hermitian Lanczos method, see [25]. These are readily programmed, e.g. in MATLAB.

3.2. Computational geometry for the convex hull of the joint numerical range

Denote by $F \subset \mathbb{R}^k$ the convex hull of the image of $\mathbf{V}$ and by $\mathbb{S}_{\mathbb{R}}^{k-1} = \{u \in \mathbb{R}^k : \|u\| = 1\}$ the set of unit vectors in $\mathbb{R}^k$. By executing Algorithm 1, for any $u \in \mathbb{S}_{\mathbb{R}}^{k-1}$ we can compute $\lambda(u) \in \mathbb{R}$ and a boundary point $p \in F$ such that

$$\lambda(u) = u^Tp = \min_{t\in F} u^Tt. \tag{3.8}$$

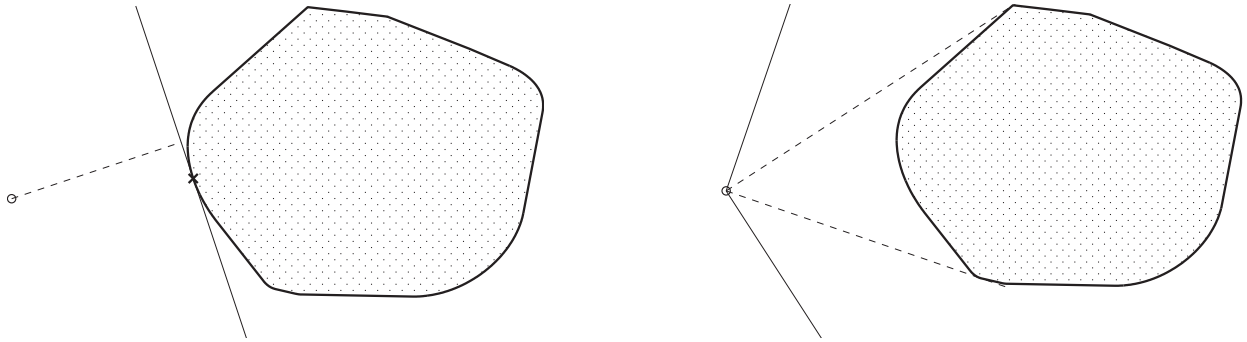
A graphical illustration of this is given in Fig. 3.1(a). For the compact set F, its *dual cone* is defined as

$$F^* = \{t \in \mathbb{R}^k : p^Tt \geqslant 0 \text{ for all } p \in F\}.$$

A graphical illustration of a dual cone in two dimensions is given in Fig. 3.1(b). Notice that the boundary lines of F^* are perpendicular to the opposite boundary lines of the smallest convex cone containing F.

[14] A matrix subspace is sparse if its members are sparse with a common sparsity pattern.

Author's personal copy



(a) The support plane (solid line) and a boundary point p (cross) corresponding to direction u (a unit vector such that $\lambda(u)u = p^T uu = \tilde{u}$ is the dashed line).

(b) The boundary lines of the dual cone F^* (solid lines) and the smallest convex cone containing F (dashed lines).

Fig. 3.1. A $k = 2$ dimensional F (thick outline). The origin is marked with a circle.

For a Hermitian matrix subspace $\mathcal{V}$, we are interested in solving the minimization problem

$$\begin{aligned} &\max\ \lambda_{\min}(V) \\ &\text{s.t. } V \in \mathcal{V} \\ &\quad V \geqslant 0 \quad (\text{equivalent to } \lambda_{\min}(V) \geqslant 0) \\ &\quad \|V\|_F = 1, \end{aligned} \tag{3.9}$$

where $\lambda_{\min}(V)$ denotes the smallest eigenvalue of a Hermitian matrix V. (The norm constraint $\|V\|_F = 1$ guarantees that we have a bounded solution.) If $V_1, \ldots, V_k$ is an orthonormal basis of $\mathcal{V}$, then (3.9) is equivalent to

$$\begin{aligned} &\max\ \lambda(u) \\ &\text{s.t. } u \in F^* \quad (\text{equivalent to } \lambda(u) \geqslant 0) \\ &\quad \|u\| = 1 \end{aligned} \tag{3.10}$$

with $V = \sum_{i=1}^k u_i V_i$. The strict feasibility problem (3.9) means locating the most positive definite matrix in $\mathcal{V}$. The strict feasibility problem (3.10) means locating the element $u \in \mathbb{S}_{\mathbb{R}}^{k-1}$ satisfying $v^T u > 0$ for all $v \in F$, as given in Definition 3.3 in terms of the corresponding hyperplane. Observe that the latter problem (3.10) can also be seen as a "dual" of the convex optimization problem

$$\begin{aligned} &\min\ \|p\| \\ &\text{s.t. } p \in F \end{aligned} \tag{3.11}$$

under the assumption $0 \notin F$.

Theorem 3.7. *If p' solves (3.11), then $u' = p'/\|p'\|$ solves (3.10) with $\lambda(u') = \|p'\|$.*

Proof. If $\lambda(u') = p^T u' < \|p'\|$ for some $p \in F$, then there exists a point on the line segment between p and p' closer to the origin than p', which is a contradiction. Therefore $\lambda(u') = \|p'\|$. In addition, for any $u \in \mathbb{S}_{\mathbb{R}}^{k-1}$,

$$\lambda(u) = p^T u \leqslant (p')^T u \leqslant \|p'\| = \lambda(u'),$$

for some $p \in F$, which proves the claim. □

Author's personal copy

From Theorem 3.7 we can conclude that for any feasible $u \in \mathbb{S}^{k-1}_{\mathbb{R}}$ and $p \in F$ holds

$$\lambda(u) \leqslant \lambda(u') = \|p'\| \leqslant \|p\| \tag{3.12}$$

and therefore any pair of primal and dual feasible points (p, u) can be used to bound the optimal values in (3.11) and (3.10).

In the next section we devise methods to so solve the computational geometric problems (3.10) and (3.11) relying, in essence, only on a "least eigenvalue solver", i.e., using Algorithm 1 we assume that for a given $u \in \mathbb{S}^{k-1}_{\mathbb{R}}$ we can generate $\lambda(u)$ and $p \in F$ such that $\lambda(u) = p^T u$.

4. Algorithms

Next algorithms for solving the positive definiteness problems are devised.

4.1. Perceptron algorithm for feasibility

The strict feasibility problem related with (3.10) consists of finding a unit vector u such that $p^T u > 0$ for all $p \in F$. Provided that the problem is strictly feasible, that is, at least one such u exists (i.e. $0 \notin F$), it can be found using a simple method known as the *perceptron algorithm*, which is a general method for finding a separating hyperplane between two arbitrary subsets of $\mathbb{R}^k$. Algorithm 2 shows one possible way to solve the strict feasibility problem related with (3.10) using the perceptron algorithm. The algorithm progresses by repeatedly updating a single vector, $\tilde{u}$ which corresponds to the normal of a hyperplane $\{t \in \mathbb{R}^k : t^T \tilde{u} = \lambda\}$ that should separate the origin from the set F with some $\lambda > 0$. Step 5 of the algorithm finds a point $p_j \in F$, that is on the wrong side of the current hyperplane and step 6 corrects this by tilting the hyperplane normal $\tilde{u}_j$ towards p_j. Namely,

$$p_j^T \tilde{u}_{j+1} = p_j^T (\tilde{u}_j + p_j) = p_j^T \tilde{u}_j + \|p_j\|^2 > 0,$$

where equality $p_j^T \tilde{u}_{j+1} = 0$ is excluded by the fact that $p^T u = \|p\|$ if and only if u and p are the pair of primal and dual solutions. The use of Algorithm 2 is best justified by the following theorem, establishing the convergence of the algorithm under minimal assumptions. The proof is adapted to our setting from [14].

Algorithm 2 Perceptron algorithm for finding a unit vector in the interior of F^*.

1: Set $j = 0$ and $\tilde{u}_1 = \mathbf{V}(x)$ for an arbitrary $x \in \mathbb{S}^{2n-1}_{\mathbb{C}}$
2: **repeat**
3: Increase j
4: Set $u_j = \frac{\tilde{u}_j}{\|\tilde{u}_j\|}$
5: Compute $p_j \in F$ such that $\lambda(u_j) = p_j^T u_j$
6: Set $\tilde{u}_{j+1} = \tilde{u}_j + p_j$
7: **until** $\lambda(u_j) > 0$
8: **return** u_j

Theorem 4.1 (Perceptron Convergence Theorem). *Algorithm 2 will converge in at most* $\max_{p\in F} \|p\|^2 / \lambda^2(u')$ *steps, where* u' *is the solution to* (3.10) *such that* $\lambda(u') > 0$.

Proof. Let u' be the solution to (3.10) and $\lambda(u_i) = p_i^T u_i \leqslant 0$ for $i = 1, \ldots, j-1$. Then, for any $i = 2, \ldots, j$, holds

$$\|\tilde{u}_i\|^2 = \|\tilde{u}_{i-1}\|^2 + 2\tilde{u}_{i-1}^T p_{i-1} + \|p_{i-1}\|^2 \leqslant \|\tilde{u}_{i-1}\|^2 + \max_{p\in F} \|p\|^2$$

Author's personal copy

and hence $\|\tilde{u}_j\|^2 \leqslant j \max_{p\in F} \|p\|^2$. On the other hand for any $i = 2, \ldots, j$ we have

$$\tilde{u}_i^T u' = \tilde{u}_{i-1}^T u' + p_{i-1}^T u' \geqslant \tilde{u}_{i-1}^T u' + \lambda(u')$$

which yields $\tilde{u}_j^T u' \geqslant j\lambda(u')$. Thereby

$$\sqrt{j} \max_{p\in F} \|p\| \geqslant \|\tilde{u}_j\| \geqslant \tilde{u}_j^T u' \geqslant j\lambda(u'),$$

that is, $j \leqslant \max_{p\in F} \|p\|^2/\lambda^2(u')$. □

Let us mention that the perceptron algorithm is used, in a bit different manner, to locate a positive definite element in Hermitian matrix subspaces in [28]. This subject has also been studied in [30].

4.2. Ellipsoid algorithm

The problem (3.10) is an optimization problem on the $(k-1)$-sphere $\mathbb{S}_{\mathbb{R}}^{k-1}$ and thereby not as such a convex optimization problem in $\mathbb{R}^k$. However, some standard convex optimization techniques may still be applied to the problem. In what follows, an ellipsoid algorithm is devised to solve the task.

To this end, consider an ellipsoid

$$\mathcal{E} = \mathcal{E}(A, t) = \{v \in \mathbb{R}^k : (v-t)^T A^{-1}(v-t) \leqslant 1\},$$

where $A \in \mathbb{R}^{k\times k}$ is positive definite and $t \in \mathbb{R}^k$ is the center of $\mathcal{E}$. Assume that $\mathcal{E}$ contains an optimal point $u' \in \mathbb{R}^k$ of an optimization problem. If $c \in \mathbb{R}^k$ and $\beta \in \mathbb{R}$ are such that $c^T u' \geqslant \beta$, then also

$$\mathcal{E} \cap \overline{H}_{c,\beta} = \mathcal{E} \cap \{v \in \mathbb{R}^k : c^T v \geqslant \beta\}$$

contains the solution. Define

$$\alpha = \frac{\beta - c^T t}{\sqrt{c^T A c}}. \tag{4.1}$$

If $\alpha = 0$, the center of the ellipsoid is on the border of the half-space $\overline{H}_{c,\beta}$. If $\alpha > 1$, the intersection is empty and if $\alpha < -1$, then $\mathcal{E} \subset \overline{H}_{c,\beta}$ [13].

For any $1 > \alpha > -1/k$, it is possible to construct an updated ellipsoid

$$\mathcal{E}' = \mathcal{E}(A', t'),$$

where

$$\begin{aligned} t' &= t - \tfrac{1+k\alpha}{k+1} b, \quad b = \tfrac{-Ac}{\sqrt{c^T A c}}, \\ A' &= \tfrac{k^2(1-\alpha^2)}{k^2-1} \left(A - \tfrac{2(1+k\alpha)}{(k+1)(1+\alpha)} bb^T \right), \end{aligned} \tag{4.2}$$

such that $\mathcal{E}' \supset \mathcal{E} \cap \overline{H}_{c,\beta}$ and the volume of $\mathcal{E}'$ is strictly less than that of $\mathcal{E}$ [13]. If $\alpha \geqslant 0$, then

$$\text{volume}(\mathcal{E}') \leqslant e^{\frac{-1}{2k}} \text{volume}(\mathcal{E})$$

[8]. This leads to the iterative algorithm known as the (deep cut) *ellipsoid algorithm* [13]. It consists of finding an initial ellipsoid $\mathcal{E}_0$, such that $u' \in \mathcal{E}_0$ and then constructing a sequence $(\mathcal{E}_j)$ of ellipsoids, where $\mathcal{E}_{j+1}$ is constructed from $\mathcal{E}_j$ using (4.2). Such an update is possible if, for each $\mathcal{E}_j$, one is able to find a half-space $\overline{H}_{c_j,\beta_j}$, for which the corresponding α_j, as defined in (4.1) satisfies $\alpha_j > -1/k$. If, in addition, $\alpha_j \geqslant 0$ for all j, then the solution u' is contained in a sequence of ellipsoids $\mathcal{E}_j$ whose volume tends geometrically to zero.

Author's personal copy

For the problem (3.10) an ellipsoid algorithm can be devised as follows. Initially, set $t_0 = 0$ and $A_0 = I$. For any $j > 0$, set $u_j = t_j/\|t_j\|$ if $\|t_j\| > 0$. Otherwise, pick an arbitrary $u_j \in \mathbb{S}_{\mathbb{R}}^{k-1}$. As shown later, it can be ensured that $\|t_j\| \leqslant 1$. Let

$$\lambda_j^{\text{best}} = \max(\{0\} \cup \{\lambda(u_i), i = 0, 1, \ldots, j\}) \tag{4.3}$$

and $p_j \in F$ be such that $\lambda(u_j) = p_j^T u_j$. By setting $c_j = p_j$, $\beta_j = \lambda_j^{\text{best}}$ we get

$$\beta_j - c_j^T t_j = \lambda_j^{\text{best}} - \|t_j\| p_j^T u_j = \lambda_j^{\text{best}} - \|t_j\|\lambda(u_j) \geqslant 0,$$

which means that $\alpha_j \geqslant 0$. This corresponds to a valid cut because

$$c_j^T u' = p_j^T u' \geqslant \lambda(u') \geqslant \lambda_j^{\text{best}} = \beta_j.$$

One may then construct an ellipsoid $\mathcal{E}(\tilde{A}_{j+1}, \tilde{t}_{j+1})$ according to (4.2). If $\|\tilde{t}_{j+1}\| \leqslant 1$, set $t_{j+1} = \tilde{t}_{j+1}$ and $A_{j+1} = \tilde{A}_{j+1}$ and continue.

In case $\|\tilde{t}_{j+1}\| > 1$, it one must conduct a *norm cut* by setting

$$\tilde{c}_j = -\tilde{t}_{j+1}/\|\tilde{t}_{j+1}\|, \quad \tilde{\beta}_j = -1 \tag{4.4}$$

and then cutting the ellipsoid $\mathcal{E}(\tilde{A}_{j+1}, \tilde{t}_{j+1})$ with the half-space $\overline{H}_{\tilde{c}_j, \tilde{\beta}_j}$ according to (4.2). The norm cut procedure can be repeated several times, if necessary, in order to have $\|t_{j+1}\| < 1$.

Regarding the relative error of the solution, the error bounds given in [13] cannot be used since the value $\lambda(u_j)$ is not an evaluation of the objective function at the center t_j of the ellipsoid. For the problem (3.10), strict error bounds are given by

$$\lambda_j^{\text{best}} =: \lambda(u_j^{\text{best}}) \leqslant \lambda(u') \leqslant P_j^{\text{best}} := \min_{0 \leqslant i \leqslant j} \|p_i\|. \tag{4.5}$$

Another upper bound is [15]

$$\lambda(u') \leqslant p^T u' = p^T t + p^T(u' - t) \leqslant p^T t + \sqrt{p^T A p}$$

for any $p \in F$ and $\mathcal{E}(A, t) \ni u'$. Therefore set

$$\lambda_j^{\max} = \min(p_j^T t_j + \sqrt{p_j^T A_j p_j}, \|p_j\|, \lambda_{j-1}^{\max}), \tag{4.6}$$

where $\lambda_0^{\max} = \infty$. With these, the resulting ellipsoid algorithm is summarized as Algorithm 3.

Algorithm 3 Deep cut ellipsoid algorithm to solve (3.10).

1: Set $j = 0$, $A_0 = I$, $t_0 = 0$, pick $u_1 \in \mathbb{S}_{\mathbb{R}}^{k-1}$
2: **repeat**
3: Increase j
4: Compute $p_j \in F$ such that $\lambda(u_j) = p_j^T u_j$
5: Compute t_j, A_j from (4.1) and (4.2), applying norm cuts (4.4) if necessary
6: Update λ_j^{best}, $\lambda_j^{\max}$ and u_j^{best} according to (4.3), (4.5) and (4.6)
7: Set $u_{j+1} = t_j/\|t_j\|$
8: **until** $\lambda_j^{\max} - \lambda_j^{\text{best}} < \epsilon$ (or some other stopping criterion is satisfied)
9: **return** u_j^{best}

[15] If $v = \arg\max_{v \in \mathcal{E}(A,0)} p^T v$, then $\nabla_v(p^T v - \mu v^T A^{-1} v) = 0 \Leftrightarrow v = \frac{1}{2\mu} A p$, and from $v A^{-1} v \leqslant 1$ we get $p^T v = \sqrt{p^T A p}$.

Author's personal copy

4.3. Accelerated ellipsoid algorithm

The principles described in Section 4.2 can be used to construct other *cutting-plane*-based methods to solve the problem (3.10). For instance, the ellipsoid method may be sped up by storing multiple points p_j computed so far. Namely, for any $p \in F$ and $j \geqslant 0$ holds

$$p^T u' \geqslant \lambda_j^{\text{best}}. \tag{4.7}$$

If $\frac{\lambda_j^{\text{best}} - p^T t_j}{\sqrt{p^T A_j p}} > -1/k$, then (4.7) defines a cut which can be used to decrease the volume of the ellipsoid when updated according to (4.2). In theory, an ellipsoid $\mathcal{E}(A, t)$ may be cut during the same iteration until it satisfies

$$\frac{\lambda_j^{\text{best}} - p_i^T t}{\sqrt{p_i^T A p_i}} \leqslant -1/k, \quad \forall 0 \leqslant i \leqslant j \quad \text{and} \quad \frac{\|t\|^2 - \|t\|}{\sqrt{t^T A t}} \leqslant -1/k.$$

It may not be feasible to find such an ellipsoid exactly, but an approximation may be computed by iterating and cutting over all p_i, $i = 0, 1, \ldots, j$ multiple times.

Algorithm 5 describes a relatively straightforward multiple cutting-plane scheme that can be used to speed up the ellipsoid algorithm if the execution time is dominated by the eigenvalue computations. It works as Algorithm 3, except that the ellipsoid is also cut with (at most M_2) constraints from the previous eigenvalue computation rounds. As in the initialization phase, the cutting is repeated over multiple (M_1) rounds. On lines 18–19 the stored constraints are pruned so that only the ones that contributed with the deepest cuts (greatest α) remain. The upper bound $\lambda_{\max}$ is updated whenever possible.

A way to further speed up the method is to use additional inequalities (3.3) to construct an initial ellipsoid. For example, a necessary condition for the positive semidefiniteness of a matrix $V \in \mathcal{V}$ is that all its diagonal elements are nonnegative. This yields n initial linear constraints in $\mathbb{R}^k$. Algorithm 4 describes a method to compute an initial ellipsoid $\mathcal{E}(A_0, t_0)$ for Algorithms 3 and 5. This method performs M_0 rounds of cutting the ellipsoid by imposing the diagonal positivity requirement.

Algorithm 4 Initialization scheme for ellipsoid algorithms for the problem (3.9).

```
1:  Set A ← I, t ← 0
2:  for i = 1, ..., n do
3:      Define d_i = [(V_1)_ii, ..., (V_k)_ii]^T
4:  end for
5:  for j = 1, ..., M_0 do
6:      for i = 1, ..., n do
7:          Calculate α for the diagonal equation d_i^T t ⩾ 0 according to (4.1)
8:          Update A and t according to (4.2) if α > −1/k
9:      end for
10:     Apply a norm cut (4.4) to A and t if ‖t‖ > 1
11: end for
12: return A,t
```

5. Numerical experiments

The difficulty of the problem (3.10) is closely related to $\lambda(u')$, the distance between the origin and F, and the size of F. With respect to these parameters, we designed *easy* and *challenging* experiments. This makes the construction of matrix subspaces $\mathcal{V} \subset \mathbb{C}^{n\times n}$ somewhat involved.

Author's personal copy

Algorithm 5 Accelerated ellipsoid algorithm to solve (3.10).

Set $A \leftarrow I$, $t \leftarrow 1$ or calculate them using Algorithm 4
Initialize $P \leftarrow [\ ]$ with an empty matrix
repeat
  Set $u \leftarrow t/\|t\|$ if $t \neq 0$ or an arbitrary $u \in \mathbb{S}_{\mathbb{R}}^{k-1}$ otherwise
  Compute $p \in F$ such that $\lambda(u) = p^T u$, store $P \leftarrow [p\ P]$
  Update λ^{best}, u^{best} and $\lambda^{\max}$ according to (4.3), (4.5) and (4.6)
  Let m be the number of columns in P
  Initialize $\alpha_i = -1$ for $i = 1, \ldots, m$
  for $j = 1, \ldots, M_1$ **do**
    for $i = 1, \ldots, m$ **do**
      Let p be the ith column of P
      Calculate α for the cutting-plane $p^T t \geqslant \lambda^{\text{best}}$
      If $\alpha > -1/k$, update A,t and $\lambda^{\max}$
      Set $\alpha_i \leftarrow \max(\alpha_i, \alpha)$
    end for
    Apply norm cut (4.4) to A and t if $\|t\| > 1$
  end for
  Sort the columns of P to descending order of α_i's
  Drop all i columns with $\alpha_i < -1/k$, keeping at most M_2 columns
until $\lambda^{\max} - \lambda^{\text{best}} < \epsilon$ (or some other stopping criterion is satisfied)
return u^{best}

Start with a Hermitian matrix subspace $\tilde{\mathcal{V}}$ spanned by the matrices $\hat{V}_j = \frac{1}{2}(A_j + A_j^*)$, where each $A_j \in \mathbb{C}^{n\times n}$, $j = 1, \ldots, k$ is a random band matrix with normally distributed complex elements having bandwidth $2j + 1$. Band matrices are used because full random Hermitian matrices were observed to produce $\mathbf{V}$ whose range seemed to resemble the k-ball. (This we regard as an unfounded bias.) Obviously, matrix subspaces constructed in this way are sparse if $k \ll n$. These matrix subspaces typically cannot be expected to contain positive definite elements (based on numerical experiments). Therefore we translate the basis matrices to have feasible problems.

For any given $b \geqslant 0$, we construct a Hermitian matrix subspace $\mathcal{V}$ such that

$$\min_{p\in F} \|p\| = b. \tag{5.1}$$

First take any Hermitian matrix subspace $\tilde{\mathcal{V}}$ with an orthonormal basis $\tilde{V}_1, \ldots, \tilde{V}_k$. Choose an arbitrary unit vector $u \in \mathbb{R}^k$ and calculate a boundary point p corresponding to $\lambda(u) = p^T u$ for the convex hull $\tilde{F}$ of the image of $\tilde{\mathbf{V}}$. Then form

$$V_j' = \tilde{V}_j + (bu_j - p_j)I$$

and orthonormalize to have $\mathcal{V} = \operatorname{span}\{V_1', \ldots, V_k'\} = \operatorname{span}\{V_1, \ldots, V_k\}$, where the matrices $V_1, \ldots, V_k$ are orthonormal and (5.1) holds.

Based on this construction, for various n and k, two types of random problems are generated. The size of the image of $\mathbf{V}$ is approximated by $d = \|p' - p\|$, where p' is a boundary point corresponding to $\lambda(-u)$. Using this number d, *easy* problems with $b = \frac{d}{10}$ and *challenging* problems with $b = \frac{d}{1000}$ are generated.

The performance of Algorithm 2, Algorithm 3 and our accelerated Algorithm 5, initialized with Algorithm 4, are compared. The performances are measured in terms of the number of iterations (or equivalently, the number of eigenvalue computations) required to solve the problems. Each cell is an average over at least ten runs with different random matrix subspaces $\mathcal{V}$. As the parameters of Algorithms 4 and 5 we used (quite arbitrarily) $M_0 = 10, M_1 = 3, M_2 = 50$. Algorithm 3 was initialized with $u_1 = \mathbf{V}(x)/\|\mathbf{V}(x)\|$ for a random $x \in \mathbb{S}_{\mathbb{C}}^{2n-1}$.

Author's personal copy

Table 5.1
Average number of iterations needed to locate a positive definite matrix.

b	k	n	Perceptron	Ellipsoid	Acc. Ellipsoid
$10^{-1}d$	5	100	2.0	2.2	1.00
$10^{-1}d$	5	1000	1.1	1.1	1.00
$10^{-1}d$	5	5000	1.1	1.1	1.00
$10^{-1}d$	15	100	5.3	6.1	1.11
$10^{-1}d$	15	1000	1.1	1.1	1.00
$10^{-1}d$	15	5000	1.00	1.00	1.00
$10^{-1}d$	100	100	6.4	8.3	2.1
$10^{-1}d$	100	1000	5	6	1.5
$10^{-3}d$	5	100	79	23	8.7
$10^{-3}d$	5	1000	17	22	8.1
$10^{-3}d$	5	5000	15	22	8.1
$10^{-3}d$	15	100	294	92	25
$10^{-3}d$	15	1000	97	88	23.9
$10^{-3}d$	15	5000	24	89	24
$10^{-3}d$	100	100	1041	438	43
$10^{-3}d$	100	1000	652	370	35

Table 5.2
Average number of iterations needed to solve (3.9) to the relative precision 10^{-6}.

b	k	n	Ellipsoid	Acc. Ellipsoid
$10^{-1}d$	5	100	193	30.3
$10^{-1}d$	5	1000	217	31.2
$10^{-1}d$	5	5000	223	32
$10^{-1}d$	15	100	1500	140
$10^{-1}d$	15	1000	1734	148
$10^{-1}d$	15	5000	–	149
$10^{-1}d$	100	100	–	6125
$10^{-1}d$	100	1000	–	5764
$10^{-3}d$	5	100	265	41.5
$10^{-3}d$	5	1000	262	40.7
$10^{-3}d$	5	5000	258	41
$10^{-3}d$	15	100	2909	168
$10^{-3}d$	15	1000	2903	166
$10^{-3}d$	15	5000	–	166

Table 5.1 compares the performance of all three algorithms on the strict feasibility problem, i.e., the problem of locating a positive definite element in a Hermitian subspace. In Algorithms 3 and 5 this is achieved by setting the stopping criterion to $\lambda^{\text{best}} > 0$. In case $b = 10^{-1}d$, the initial guess (calculated with Algorithm 4 or as $\mathbf{V}(x)/\|\mathbf{V}(x)\|$) was often enough to have a valid solution and the corresponding average iteration counts are therefore close to 1. With Algorithm 2, the iteration counts varied considerably within a class of problems with same parameters (e.g. from 9 to occasionally hundreds with $k = 15$, $n = 1000$, $b = 10^{-3}d$). The iteration counts for Algorithms 3 and 5 were more stable in all problems, primarily depending on b and secondarily on k.

Table 5.2 compares the performance of Algorithms 3 and 5 on the least eigenvalue maximization problem (3.9). A relative stopping criterion $\frac{\lambda^{\max}-\lambda^{\text{best}}}{\lambda^{\max}} < \epsilon = 10^{-6}$ was used. The problems marked with a dash took too long to solve. The results indicate that the problem (3.9) can be solved reasonably efficiently in matrix subspaces with low dimension k, if solving the extremal eigenvalue problems is feasible. Otherwise the iteration count does not seem to have much dependence on n. The difficulty of solving the maximization problem seems to primarily depend on the dimension k (and not so much on b). The average CPU time needed to solve the problem using Algorithm 5 with $b = 10^{-3}d$, $k = 15$, $n = 1000$ on the test workstation [16] was 24 s such that 85% of it was spent inside the least eigenvalue solver routine (MATLAB's `eigs` function).

[16] 2.66 GHz Intel Core 2 Quad Q8400, 3.2 GB RAM, MATLAB R2010b on Debian GNU/Linux.

Author's personal copy

Appendix A. The case dim $\mathcal{V} = 2$

The two dimensional case is instructive, classical and can be solved satisfactorily. For the two dimensional case, discussed in terms of matrix pairs, see [23]. See also [16].

Denote by $F(M)$ the numerical range of a matrix $M \in \mathbb{C}^{n\times n}$.

Theorem A.1 [12,1]. *Assume the matrices V_1 and V_2 span a Hermitian matrix subspace $\mathcal{V}$. Then $\mathcal{V}$ contains possess positive definite if and only if $0 \notin F(V_1 + iV_2)$.* [17]

The location of the numerical range determines the positive definite linear combinations completely as follows. (Recall that the numerical range is convex.)

Corollary A.2. *Let $\theta_1 \leqslant \theta_2$ be the angles of the smallest cone centered at the origin containing $F(V_1+iV_2)$ with $\theta_2 - \theta_1 < \pi$. Then, with $\gamma = \frac{\theta_2+\theta_1}{2}$, exactly*

$$\cos(\gamma - \theta)V_1 + \sin(\gamma - \theta)V_2 \tag{A.1}$$

for $\theta \in (-\frac{1}{2}(\pi + \theta_1 - \theta_2), \frac{1}{2}(\pi + \theta_1 - \theta_2))$ are positive definite.

For a graphical illustration, see Fig. 3.1(b), where the dashed lines yield the cone of the corollary.

It is the dual cone in Fig. 3.1(b) which is of importance in quantifying the size of $\mathcal{V} \cap \mathcal{S}_{++}$. For an isometric isomorphism between the parameter space $\mathbb{R}^2$ and $\mathcal{V}$, assume that V_1 and V_2 are orthonormal with respect to the inner product (2.1). The matrices

$$\cos(\theta_2 - \pi/2)V_1 + \sin(\theta_2 - \pi/2)V_2 \text{ and } \cos(\theta_1 + \pi/2)V_1 + \sin(\theta_1 + \pi/2)V_2$$

determine the boundaries of the cone $\mathcal{V} \cap \mathcal{S}_{++}$. This angle is independent of the matrices V_1 and V_2 spanning $\mathcal{V}$, as long as they are orthonormal. Interpreted in terms the dual cone, see Fig. 3.1(b). If $\theta_2 - \theta_1 \approx \pi$, then the dual cone is needle-like. Then and only then positive definite elements occupy a tiny portion in $\mathcal{V}$.

Hence, analogously, in the general case of dim $\mathcal{V} = k$, the size of the dual cone needs to be compared against the solid angle of the sphere in $\mathbb{R}^k$ which is

$$\frac{k\pi^{k/2}}{\Gamma(k/2+1)},$$

where Γ denotes Gamma function.

Appendix B. Related problems

Locating a positive definite element in a Hermitian matrix subspace resembles a class of convex optimization problems known as *semidefinite programs*, which can be formulated as

$$\begin{aligned} &\min\ c^T u \\ &\text{s.t. } V_0 + \textstyle\sum_{i=1}^k u_i V_i \geqslant 0, \end{aligned} \tag{B.1}$$

where $V_0, \ldots, V_k$ are Hermitian matrices. The feasibility problem of finding a positive semidefinite matrix from an affine Hermitian subspace is also known as a *linear matrix inequality*. Semidefinite programs and linear matrix inequalities for linear subspaces (that is $V_0 = 0$) are not much of interest since they are either trivially solved by the zero matrix or unbounded. The problem (3.9) is not a special case of a semidefinite program, but an optimization problem on $\mathbb{S}_{\mathbb{R}}^{k-1}$.

[17] In [4, p.76] a related result is called Finsler's theorem.

Author's personal copy

Acknowledgments

We would like to thank Gerard Meurant for bringing to our attention relevant references concerning positive definite linear combinations. We are also very grateful to an anonymous referee for an exceptionally detailed referee report and the suggestions for rewriting.

References

[1] Y.-H. Au-Yeung, Y.T. Poon, A remark on the convexity and positive definiteness concerning Hermitian matrices, Southeast Asian Bull. Math. 3 (1979) 85–92.
[2] A. Barvinok, A course in convexity, in: Graduate Studies in Math., vol. 54, AMS, Providence, RI, 2002.
[3] S. Basu, R. Pollack, M.-F. Roy, Algorithms in real algebraic geometry, in: Algorithms and Computation in Mathematics, vol. 10, Springer-Verlag, Berlin, 2006.
[4] R. Bellman, Introduction to Matrix Analysis, McGraw-Hill, New York, 1960.
[5] R. Bhatia, Positive Definite Matrices, Princeton Series in Applied Mathematics, Princeton University Press, Princeton, NJ, 2007.
[6] P. Binding, Hermitian forms and the fibration of spheres, Proc. Amer. Math. Soc. 94 (1985) 581–584.
[7] S. Boyd, L. Vandenberghe, Convex Optimization, Cambridge University Press, Cambridge, 2004.
[8] S. Boyd, L. El Ghaoui, E. Feron, V. Balakrishnan, Linear matrix inequalities in system and control theory, in: Studies in Applied Mathematics, vol. 15, SIAM, 1994.
[9] M. Byckling, M. Huhtanen, Approximate factoring of the inverse, Numer. Math. 117 (2011) 507–528.
[10] S. Cheng, N. Higham, The nearest definite pair for the Hermitian generalized eigenvalue problem, Linear Algebra Appl. 302/303 (1999) 63–76.
[11] E.G. Effros, Z.-J. Ruan, Operator Spaces, Oxford University Press, Oxford, 2000.
[12] P. Finsler, Über das Vorkommen definiter und semidefiniter Formen in Scharen quadratischer Formen, Comment. Math. Helv. 9 (1937) 188–192.
[13] J.B.G. Frenk, J. Gromicho, S. Zhang, A deep cut ellipsoid algorithm for convex programming: theory and applications, Math. Program. 63 (1994) 83–108.
[14] Y. Freund, R.E. Schapire, Large margin classification using the perceptron algorithm, Mach. Learn. 37 (1999) 277–296.
[15] G.H. Golub, C.F. van Loan, Matrix Computations, third ed., The Johns Hopkins University Press, 1996.
[16] E. Jonckheere, F. Ahmad, E. Gutkin, Differential topology of numerical range, Linear Algebra Appl. 279 (1–3) (1998) 227–254.
[17] E. Gutkin, E. Jonckheere, M. Karow, Joint numerical range: topological and differential geometric viewpoints, Linear Algebra Appl. 376 (2004) 143–171.
[18] N. Higham, F. Tisseur, P. Van Dooren, Detecting a definite Hermitian pair and a hyperbolic or elliptic quadratic eigenvalue problem, and associated nearness problems, Linear Algebra Appl. 351–352 (2002) 455–474.
[19] R.A. Horn, C.R. Johnson, Matrix Analysis, Cambridge University Press, Cambridge, 1987.
[20] R.A. Horn, C.R. Johnson, Topics in Matrix Analysis, Cambridge University Press, Cambridge, 1991.
[21] M. Huhtanen, Factoring matrices into the product of two matrices, BIT 47 (2007) 793–808.
[22] M. Huhtanen, Differential geometry of matrix inversion, Math. Scand. 107 (2010) 267–284.
[23] P. Lancaster, L. Rodman, Canonical forms for Hermitian matrix pairs under strict equivalence and congruence, SIAM Rev. 47 (2005) 407–443.
[24] A.S. Lewis, The mathematics of eigenvalue optimization, Math. Program. 97 (2003) 155–176.
[25] B. Parlett, The symmetric eigenvalue problem, in: Classics in Applied Mathematics, vol. 20, SIAM, Philadelphia, 1997.
[26] V. Paulsen, Completely Bounded Maps and Operator Algebras, Cambridge University Press, Cambridge, 2002.
[27] M. Todd, Semidefinite programming, Acta Numer. 10 (2001) 515–560.
[28] L. Tong, Y. Inouye, R.W. Liu, A finite-step global convergence algorithm for the parameter estimation of multichannel MA processes, IEEE Trans. Signal Process. 40 (1992) 2547–2558.
[29] G.A. Watson, Computing the numerical radius, Linear Algebra Appl. 234 (1996) 163–172.
[30] A. Zaidi, Positive definite combination of symmetric matrices, IEEE Trans. Signal Process. 53 (2005) 4412–4416.