Deciding polyhedrality of spectrahedra

Jules BESSON

April 20, 2023

Abstract

This document describes and add explanations/descriptions to the article *Deciding polyhedrality of spectrahedra*[1]. The goal here is not to paraphrase but to give further enlightment on rather implicit elements of the article.

1   Introduction

Since spectrahedra are a generalisation of polyhedra, it is natural to ask oneself if the spectrahedron we are working on is polyhedral. This allows us to solve our optimisation problem with linear programming methods, such as the simplex or ellipsoïd method. The article provides an algorithm stating if a spectrahedron is polyhedral.

The first step is a simplification allowing us to go from an affine point of view to a linear one. To do so, we go from a spectrahedron

$$\mathcal{S} = \{x \in \mathbb{R}^d : x_1 A_1 + \cdots + x_d A_d + B \succcurlyeq 0\}.$$

and associate to it the spectrahedral cone

$$S = \{x \in \mathbb{R}^{d+1} : x_1 A_1 + \cdots + x_d A_d + x_{d+1} B \succcurlyeq 0, \ x_{d+1} \geqslant 0\}.$$

Adding this optional parameter allows us to get back to a linear map $A'(x)$ instead of the affine map $A(x)$.
The proposition 1.1 states that $\mathcal{S}$ is polyhedral *iff* S is, consequently we only need to provide an algorithm for cones. The proof is very clear, thus no detail will be added.

The first statement we can make is an easy criterion for polyhedrality: the codiagonalisation of $A(x)$. If we have it, the positivity condition transforms in a positivity condition for all diagonal coefficients, which gives us a polyhedron.

Why the codiagonalisation criterion is sufficient but not necessary? The first reason is a problem with commutativity. Indeed it is necessary and sufficient for the codiagonalisation of all $A(x)$, but it is not preserved by congruence which preserves positivity...
The second (more problematic) issue is when a spectrahedron is redundantly defined. Let us take S a spectrahedral cone and C a polyhedral cone such that $C \subset S$. The formal intersection will be defined by the block matrix map

$$\begin{pmatrix} A(x) & 0 \\ 0 & D(x) \end{pmatrix}$$

where A is the matrix map associated to S and D the diagonal matrix map associated to C. Of course in this case, if S is not polyhedral, A cannot be diagonalised, thus the diagonalisation criterion fails. If we were to observe this example it would be clear that the condition $A(x) \succcurlyeq 0$ was optionnal. We must find a way to get around this.

2   Criterion for unital matrice maps

In this section S is a spectrahedral cone defined by the unital matrix map $A(x)$. This means that there exists $x \in S$, $A(x) = I_n$. The items 2.1 to 2.4 of the article show us that we can operate a similar reduction as the one used for quadratic forms, we keep in mind that the spectrahedral inequality is preserved by congruence (Sylvester's inertia). Proposition 2.7 states that we can find an orthogonal matrix U such that $A(x)$ is in normal form:

$$U A(x) U^T = \begin{pmatrix} Q(x) & 0 \\ 0 & D(x) \end{pmatrix}$$

with $D(x)$ diagonal and $Q(x)$ proper, i.e. it admits no common eigenvector for all $Q(x)$, $x \in S$. This property ensures us some canonicity by making it impossible to diagonalise $A(x)$ further.

If S was to be polyhedral, it would then only be defined by $D(x)$. Consequently, we introduce its *polyhedral hull* $\hat{S}$ defined by

$$\hat{S} := \{x \in \mathbb{R}^d : D(x) \succcurlyeq 0\}.$$

If S was polyhedral, we would have $S = \hat{S}$, or equivalently $\forall x \in S$, $Q(x) \succcurlyeq 0$. Now we just have to find a way to get to this normal form and to verify in finite time the assertion $\forall x \in S$, $Q(x) \succcurlyeq 0$.

For the first step the technique used is to find the *joint invariant subspace* $\mathcal{N}$ of all $A(x)$, $x \in S$. It is the linear hull of all common eigenvectors of $A(x)$s. It is clear that $D(x)$ lives on this subspace, indeed the block matrix gives us two invariant subspaces, the second one is $\mathcal{N}$ since the codiagonalisation is possible with the existence of common eigenvectors.

We will now characterise this subspace in a computable way.

2.1 THEOREM:
Two square matrices A and B over $\mathbb{C}$ have common eigenvectors *iff*

$$\mathcal{N} = \bigcap_{i,j=1}^{n-1} \ker[A^i, B^j] \neq 0$$

proof: The first direction is obvious, for the second we can see that $\mathcal{N}$ is invariant by the action of A and B, hence we can restrict them to $\mathcal{N}$. On this space, A and B commute, since A has an eigenvector, it has an eigenspace V, on which B is stable, thus B has an eigenvector on V which is also an eigenvector for A. □

2.2 COROLLARY:
If A and B are two diagonalisable matrices, the smallest linear space containing all common eigenvectors, called the *joint invariant subspace* is the set

$$\mathcal{N} = \bigcap_{i,j=1}^{n-1} \ker[A^i, B^j].$$

proof: $\mathcal{N}$ is still invariant for A and B, which are diagonalisable, hence codiagonalisable on $\mathcal{N}$, thus $\mathcal{N}$ contains all eigenvectors and is minimal. □

With this corollary, we now can determine the joint invariant subspace, but only for two points $A(p)$ and $A(q)$. The next theorem explains how chossing two "good" points called *generic* suffice to find $\mathcal{N}$.

2.3 THEOREM:
Let $A(x)$ be a unital matrix map and let p, $q \in R^d$ be two distinct generic points. Let $\mathcal{N} \subset \mathbb{R}^n$ the smallest subspace containing all eigenvectors common to $A(p)$ and $A(q)$. Then $\mathcal{N}$ is invariant under any matrix in the image of $A(x)$ and $\mathcal{N}^\perp$ is the largest invariant subspace on which $A(x)$ restricts to a proper matrix map.

proof: Suppose that A is already in normal form. We just need to verify that there exists two points (that we can compute) p and q for which $Q(p)$ and $Q(q)$ have no common eigenvectors. If we have that, this will prove the property since $\mathcal{N} = \bigcap_{i,j=1}^{n-1} \ker[A(p)^i, A(q)^j]$.

For this, we consider the polynomial $f(x,y) = \sum_{i,j=1}^{n-1} \det[Q(x)^i, Q(y)^j]^2$. The pair of points (x,y) for which $Q(x)$ and $Q(y)$ share a common eigenvector is defined by $V(f)$, thus it is a proper closed set under Zariski's topology and it is nowhere dense, S beeing full-dimensional, it is possible to find two such *generic* points p and q. □

Two such points can be found in polynomial time.

Now that we have the joint invariant subspace, we only need to make an orthogonal congruence to get the expression of $A(x)$ in a suitable base for the decomposition $\mathbb{R}^d = \mathcal{N}^\perp \oplus \mathcal{N}$ and diagonalise the symetric matrix living in the subspace $\mathcal{N}$ to get the normal form.

This means we have the data of Q, the only problem is to find a way to evaluate the condition $Q(x) \succcurlyeq 0$ in finite time.

To do so, we need to introduce *extreme rays* which are a generalisation of generative family for polyhedral cones.

2.4 PROPOSITION:
Let C be a polyhedral cone, there exists a finite set $R = R(C) \subset C$ of vectors such that each point of C is a positive linear combination of elements of R, called *extreme rays*.
Equivalently,

$$C = \left\{ \sum_{r \in R} \lambda_r r : \forall r \in R, \; \lambda_r \geqslant 0 \right\}$$

proof: If C is a vector space, it is direct.
Else, C is convex, hence it is the convex hull of all its faces of codimension one of which there is a finite number. Each face is either a vector space or a smaller dimension polyhedral cone, thus by induction, it remains to show that 1th dimensional polyhedral cones verify this property. It is obvious since any non-zero vector fits as an extreme ray. □

Then if S was polyhedral, it would be equal to $\hat{S}$, thus we only need to verify $Q(x) \succcurlyeq 0$ on $\hat{S}$'s extreme rays $R(\hat{S})$. This can be done in finite time.

Now that we have done all this work it is easy to see that operating steps 3 to 7 of the algorithm 1 tells us, wether or not S is polyhedral.

3 Cases where the matrice map is not unital

The above case is probably very frequent but not general, there are two issues possible:

1. S is not full dimensional.

2. S is full dimensional but not unital.

For the first problem, we can see that the spectrahedron is contained in a slice of $\mathbb{R}^d$, the following items give us a way to diminish the dimension of the space in which S is living, to become fully dimensional.

3.1 PROPOSITION:
If $p \in \text{rel int}(S)$ then for each $x \in S$, $\ker A(p) \subset \ker A(x)$.

proof: Let $v \in \ker A(p)$ and $x \in S$. The application $f : t \mapsto v^T A((1-t)p + tx)v$ is linear. Since p is in the relative interior of S, there exists $\epsilon > 0$ such that $|t| \leqslant \epsilon \Rightarrow (1-t)p + tx \in S$, thus $f(\pm\epsilon) \geqslant 0$. By linearity we have $f(\pm\epsilon) = 0$, hence $f \equiv 0$ and $v^T A(x)v = f(1) = 0$, so $v \in \ker A(x)$. □

3.2 COROLLARY:
Let $S = \{x : A(x) \succcurlyeq 0\}$ be a spectrahedral cone and let $p \in \text{rel int} S$ a point in the relative interior. Then the linear hull of S is given by
$$\text{lin}(S) = \{x \in R^d : \ker A(p) \subset \ker A(x)\}.$$
If $\bar{A}(x)$ is the restriction of $A(x)$ to $(\ker A(p))^\perp$, then

$$S = \{x \in \text{lin}(S) : \bar{A}(x) \succcurlyeq 0\}$$

and $\bar{A}(p) \succ 0$.

For the second case, which can be after the previous step, we can get back to a unital matrix up to congruence. For example we chose a point $p \in S$ for which $A(p)$ is invertible, and L a Cholesky inverse of $A(p)$ (i.e. $LA(p)L^T = I_n$) or the radical of $A(p)^{-1}$, which will also be a Cholesky inverse, and study $LA(x)L^T$ instead of $A(x)$.

All this work beeing done, the algorithm 1 is now fully explained and tells us if our spectrahedral cone is polyhedral.

References

[1] Philipp ROSTALSKI Avinash BHARDWAJ and Raman SANYAL. Deciding polyhedrality on spectrahedra. 2011.