The Positive Semidefinite Rank of Matrices and Polytopes

Richard Z. Robinson

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2014

Reading Committee:

Rekha R. Thomas, Chair

Isabella Novik

Paul Beame

Program Authorized to Offer Degree:
UW Mathematics

University of Washington

Abstract

The Positive Semidefinite Rank of Matrices and Polytopes

Richard Z. Robinson

Chair of the Supervisory Committee:
Professor Rekha R. Thomas
Mathematics

The positive semidefinite (psd) rank of a nonnegative $p \times q$ matrix M is defined to be the smallest integer k such that there exist $k \times k$ psd matrices $A_1, \ldots, A_p$ and $B_1, \ldots, B_q$ with $M_{ij} = \langle A_i, B_j \rangle$. The psd rank of a polytope is defined to be the psd rank of its slack matrix, a special matrix encoding the vertex-facet structure of the polytope, which is intimately related to representations of the polytope as a projection of a slice of a closed convex cone. We apply tools from algebraic geometry, optimization, linear algebra, and other fields to gain a more complete understanding of psd rank. We present lower and upper bounds on the psd rank of polytopes and investigate the class of polytopes achieving the lower bound. We classify the set of all slack matrices and show computational complexity results relating to psd rank. We explore the space of possible psd factorizations as a topological space and exhibit a special case where it is connected.

TABLE OF CONTENTS

ACKNOWLEDGMENTS

I would especially like to thank my adviser, Rekha Thomas, who always had time to help and ideas to share. She taught me to value not only a beautiful proof, but also a well-written paper and a clear presentation. I cannot imagine a more helpful mentor.

From grade school through graduate school, I have been taught by many skilled and encouraging mathematicians. In particular, I would like to thank Peggy Tholen, Estela Gavosto, Rodolfo Torres, and João Gouveia for their time and tutelage.

I would like to thank the Department of Mathematics at the University of Washington and the Graduate Research Fellowship Program of the National Science Foundation for their financial support during this project.

Lastly, I would like to thank my parents and my wife. I owe you everything.

Chapter 1

INTRODUCTION

The rank of a matrix is a fundamental quantity in mathematics whose importance requires no justification. As is the case with many mathematical objects, we can tweak the definition to obtain related values that are also worthy of study. We define the rank of a matrix M as the smallest integer k such that we can find vectors in $\mathbb{R}^k$ with $M_{ij} = \langle a_i, b_j \rangle$. Now if we alter the definition so that the a_i's and b_j's are required to be nonnegative vectors, then we have defined *nonnegative rank*, a quantity that has received much attention in the past few decades [CR93, LS99]. Alternatively, if we modify the rank definition so that the a_i's and b_j's are positive semidefinite (psd) matrices, then we have defined *positive semidefinite rank* (which we often shorten to psd rank). This quantity was defined recently in [GPT13] as an example of a *cone rank* and relatively little is known about it.

If psd rank was merely a mathematical tweak, then perhaps there would not be much value in its study. Fortunately for this thesis, psd rank finds application in several interesting problems. For example, it is known that no polynomially-sized linear programming formulation exists for the cut polytope [FMP$^+$12], but it is unknown whether there exists a polynomially-sized semidefinite programming formulation. Answering this question turns out to be equivalent to determining the psd rank of a certain matrix. In the area of information theory, the psd rank of the matrix defining a bipartite probability distribution characterizes the amount of quantum information that must be shared for two separate parties to generate the distribution [JSWZ13]. Given how recently psd rank was defined, we hope that further applications will emerge as more people are exposed to the definition. One promising area is in the field of machine learning, where nonnegative matrix factorizations have recently found wide application [LS99, AGH$^+$12].

This thesis is dedicated to studying the properties of psd rank and is organized as follows. The material for each of the chapters beyond the introduction is taken from the following articles coauthored on psd rank.

- J. Gouveia, R. Grappe, V. Kaibel, K. Pashkovich, R.Z. Robinson, and R.R. Thomas. Which nonnegative matrices are slack matrices? *Linear Algebra and its Applications*, 439(10):2921 – 2933, 2013.

- J. Gouveia, R. Z. Robinson, and R. R. Thomas. Polytopes of minimum positive semidefinite rank. *Discrete & Computational Geometry*, 50(3):679–699, 2013.

- J. Gouveia, R. Z. Robinson, and R. R. Thomas. Worst-case results for positive semidefinite rank. *arXiv:1305.4600*, 2013.

- H. Fawzi, J. Gouveia, and R. Z. Robinson. Rational and real positive semidefinite rank can be different. *arXiv:1404.4864*, 2014.

- H. Fawzi, J. Gouveia, P. Parrilo, R.Z. Robinson, and R.R. Thomas. Positive semidefinite rank. *arXiv:1407.4095*, 2014.

The specific articles that a chapter draws its material from are referenced in the chapter title. We believe that this body of work contains many insights into the properties and geometric structure of psd rank and lays the foundation for future discoveries in this area.

The rest of this chapter is organized as follows. In Section 1.1, we define psd rank and explore some basic facts that follow from the definition. Section 1.2 describes the geometric problem that led to the definition of psd rank. Section 1.3 contains a summary of all the results presented in later chapters.

1.1 Definition and first properties of psd rank

In this section we define psd rank and derive some of its basic properties. For a thorough introduction to the known properties of psd rank, please see the survey paper [FGP+14].

Let $\mathcal{S}^k$ denote the vector space of all $k \times k$ real symmetric matrices and let $\mathcal{S}^k_+$ denote the closed convex cone of psd matrices that lies within $\mathcal{S}^k$. We equip $\mathcal{S}^k$ with the standard inner product defined by:

$$\langle A, B\rangle = \operatorname{trace}(AB) = \sum_{1\le i,j\le k} A_{ij}B_{ij}.$$

Additionally, let $\mathbb{R}_+$ denote the set of nonnegative real numbers. We are now ready to formally define psd rank.

Definition 1.1.1 ([GPT13])**.** Given a nonnegative matrix $M \in \mathbb{R}_+^{p\times q}$, a *psd factorization* of M of size k is a collection of psd matrices $A_1, \dots, A_p \in \mathcal{S}_+^k$ and $B_1, \dots, B_q \in \mathcal{S}_+^k$ such that $M_{ij} = \langle A_i, B_j\rangle$ for all $i = 1, \dots, p$ and $j = 1, \dots, q$. The *psd rank* of M, denoted $\operatorname{rank}_{\text{psd}}(M)$, is the smallest integer k for which M admits a psd factorization of size k.

Remark 1.1.2. If we replace $\mathcal{S}_+^k$ in the previous definition with $\mathbb{R}_+^k$, then we have defined the *nonnegative rank* of a nonnegative matrix M, which we denote by $\operatorname{rank}_+(M)$. By mapping vectors in $\mathbb{R}_+^k$ to diagonal matrices in $\mathcal{S}_+^k$, we see that nonnegative rank is an upper bound to psd rank. This quantity has received much attention recently [CR93, LS99, AGKM12] and we will often compare our results on psd rank to known results on nonnegative rank.

Since the inner product of any two psd matrices is nonnegative, it is apparent that the psd rank can only be defined for nonnegative matrices M. Furthermore, since any psd factorization of size k induces a matrix factorization $M = UV$ with inner dimension $\binom{k+1}{2}$ (obtained by writing the psd matrices as vectors), we see that psd rank cannot be much smaller than the usual rank. The next proposition formally states this bound.

Proposition 1.1.3 ([FGP+14])**.** *If $M \in \mathbb{R}_+^{p\times q}$ is a nonnegative matrix, then*

$$\frac{1}{2}\sqrt{1 + 8\operatorname{rank}(M)} - \frac{1}{2} \le \operatorname{rank}_{\text{psd}}(M) \le \min(p, q). \tag{1.1}$$

Proof. Let M_i denote the ith row of M and let e_j be the jth standard unit vector. Then

$$\langle \operatorname{diag}(M_i), \operatorname{diag}(e_j)\rangle = \langle M_i, e_j\rangle = M_{ij}.$$

Hence, there exists a psd factorization of M of size q. Similarly, there exists a psd factorization of M of size p and we have shown the second inequality.

Suppose that $A_1, \dots A_p, B_1, \dots, B_q$ give an $\mathcal{S}_+^k$-factorization of M. Consider the vectors

$$a_i = \operatorname{vec}(A_i) \quad \text{and} \quad b_j = \operatorname{vec}(B_j)$$

where for $X \in \mathcal{S}^k$ we define $\text{vec}(X) \in \mathbb{R}^{\binom{k+1}{2}}$ by:

$$\text{vec}(X) = (X_{11}, \ldots, X_{kk}, \sqrt{2}X_{12}, \ldots, \sqrt{2}X_{1k}, \sqrt{2}X_{23}, \ldots, \sqrt{2}X_{(k-1)k}).$$

Then $\langle a_i, b_j \rangle = \langle A_i, B_j \rangle = M_{ij}$ so M has rank at most $\binom{k+1}{2}$. By solving for k we get the desired inequality. □

Example 1.1.4. As a first example of a psd factorization, consider the following matrix M known as the 3×3 derangement matrix:

$$M = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}.$$

This matrix M has usual rank three. One can show that M admits a psd factorization of size two. Indeed, define:

$$A_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad A_2 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \quad A_3 = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$$
$$B_1 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \quad B_2 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad B_3 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

One can easily check that the matrices A_i and B_j are positive semidefinite, and that $M_{ij} = \langle A_i, B_j \rangle$ for all $i = 1, \ldots, 3$ and $j = 1, \ldots, 3$. This factorization shows that $\text{rank}_{\text{psd}}(M) \leq 2$. In fact, we must have equality since the first inequality in (1.1) gives $\text{rank}_{\text{psd}}(M) \geq 2$.

It is straightforward to see that psd rank is subadditive and invariant under positive row and column scalings as we show in the following proposition.

Proposition 1.1.5 ([FGP+14]). *Given a nonnegative matrix $M \in \mathbb{R}_+^{p \times q}$, we have:*

(i) $\text{rank}_{\text{psd}}(M) = \text{rank}_{\text{psd}}(M^T)$.

(ii) If $D_1 \in \mathbb{R}_+^{p \times p}, D_2 \in \mathbb{R}_+^{q \times q}$ are diagonal matrices with strictly positive elements on the diagonal, then $\text{rank}_{\text{psd}}(D_1 M D_2) = \text{rank}_{\text{psd}}(M)$.

(iii) If $N \in \mathbb{R}_+^{p \times q}$, then $\text{rank}_{\text{psd}}(M + N) \leq \text{rank}_{\text{psd}}(M) + \text{rank}_{\text{psd}}(N)$.

Proof.

(i) Property (i) is immediate from the definition.

(ii) If $M_{ij} = \langle A_i, B_j \rangle$ is a psd factorization of M, then

$$(D_1 M D_2)_{ij} = \langle (D_1)_{ii} A_i, (D_2)_{jj} B_j \rangle$$

is a psd factorization of $D_1 M D_2$ of the same size. Thus since the diagonal elements of D_1 and D_2 are strictly positive we easily get that $\operatorname{rank}_{\text{psd}}(M) = \operatorname{rank}_{\text{psd}}(D_1 M D_2)$.

(iii) Let $M_{ij} = \langle A_i, B_j \rangle$ and $N_{ij} = \langle A'_i, B'_j \rangle$ be psd factorizations of M and N of size respectively $\operatorname{rank}_{\text{psd}}(M)$ and $\operatorname{rank}_{\text{psd}}(N)$. Define

$$C_i = \begin{pmatrix} A_i & 0 \\ 0 & A'_i \end{pmatrix} \quad \text{and} \quad D_j = \begin{pmatrix} B_j & 0 \\ 0 & B'_j \end{pmatrix}.$$

Note that C_i and D_j are psd matrices of size $\operatorname{rank}_{\text{psd}}(M) + \operatorname{rank}_{\text{psd}}(N)$. Furthermore we clearly have $M_{ij} + N_{ij} = \langle C_i, D_j \rangle$. Thus $\operatorname{rank}_{\text{psd}}(M + N) \leq \operatorname{rank}_{\text{psd}}(M) + \operatorname{rank}_{\text{psd}}(N)$.

□

1.2 Geometric motivation for psd rank

Now that we have seen the definition of psd rank and derived a few of its properties, we will use this section to explore the geometric problem that led to the definition of psd rank.

In the past few decades, the field of extended formulations has received a lot of attention [LS91, Yan91]. This field is based off of a simple observation: a linear projection of a polyhedron may contain many more facets than the original polyhedron. For a simple example of this phenomenon consider the following polytope:

$$P_n = \left\{ (x, y) \in \mathbb{R}^{2n} \ : \ \sum_{i=1}^{n} y_i = 1, -y_i \leq x_i \leq y_i, \forall i = 1, \ldots, n \right\}$$

which is a polytope in $\mathbb{R}^{2n}$ with $2n$ facets. If we project P_n onto the x-coordinates, however, we obtain the n-dimensional cross polytope

$$C_n = \{x \in \mathbb{R}^n \; : \; \pm x_1 \pm x_2 \ldots \pm x_n \leq 1\}$$

which has 2^n facets. We refer to the higher-dimensional polytope P_n as a *lift* of the lower-dimensional C_n. Once we have the lift, it is straightforward to see that any linear optimization problem over the lower-dimensional polytope may be reformulated as a linear optimization problem over the lift. Since the running time of linear programming algorithms is determined by the number of facets of the feasible set, the reformulated problem may be far more efficient. This lift-and-project approach has led to more efficient algorithms for some LPs and even proposed P=NP proofs (which of course were flawed, see [Yan91]).

In constructing a traditional extended formulation, we attempt to lift the given polytope to some higher dimensional polyhedron. We note two things about polyhedra - they can be written as slices of the nonnegative orthant by an affine space, and they admit an efficient optimization scheme. Now *spectrahedra* are a class of convex sets that are defined as slices of the psd cone by an affine space. Furthermore, linear optimization over spectrahedra can also be accomplished efficiently through semidefinite programming. In this way, spectrahedra are a natural generalization of the polyhedra used in extended formulations and we can extend the lift-and-project framework to this setting.

Let $P \subset \mathbb{R}^n$ be a polytope and suppose that P admits a representation of the form

$$P = \pi(\mathcal{S}^k_+ \cap \mathcal{L}) \tag{1.2}$$

where π is a linear map and $\mathcal{L} \subset \mathcal{S}^k$ is an affine subspace. We call such a representation a *psd lift* of size k. As with linear lifts, we are interested in the case where the size of the psd lift is smaller than the number of facets of P.

Example 1.2.1. For a first example of a psd lift, let's consider the square $P = [-1, 1]^2 \subset \mathbb{R}^2$. Let E be the affine slice of $\mathcal{S}^3_+$ defined as follows:

$$E = \left\{ (x, y, z) \; : \; \begin{pmatrix} 1 & x & y \\ x & 1 & z \\ y & z & 1 \end{pmatrix} \succeq 0 \right\}.$$

Then the projection of E onto the (x, y)-coordinates yields the square P. Hence, E is a psd lift of P of size 3. The set E (known as the elliptope) and this projection are shown in Figure 1.1. We will return to this example later in the section.

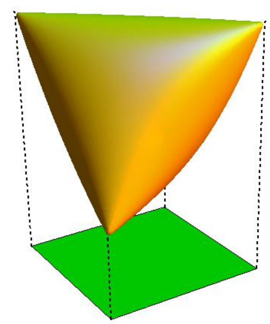


Figure 1.1. Here we see the elliptope $\{X \in \mathcal{S}_+^3 \ : \ \text{diag}(X) = \mathbb{1}\}$ and its linear projection onto the square $[-1, 1]^2$.

Example 1.2.2. For higher dimensional examples of psd lifts, we recall the Lovász theta body introduced in [Lov79]. This construction initiated the application of semidefinite programming to combinatorial optimization and is thoroughly explored in [GLS88, Chapter 9].

Let G be a perfect graph on n nodes and consider the *stable set polytope* $\text{STAB}(G)$ (a stable, or independent, set is a subset of the nodes such that no two nodes are adjacent. The convex hull of all incidence vectors of stable sets forms $\text{STAB}(G)$). Then $\text{STAB}(G)$ is a full-dimensional polytope in $\mathbb{R}^n$. Define the theta body of G to be

$$\text{TH}(G) = \left\{ x \in \mathbb{R}^n \ : \ \begin{pmatrix} 1 & x^T \\ x & U \end{pmatrix} \succeq 0 \right\}$$

where U is allowed to be any matrix in $\mathcal{S}^n$ such that $\text{diag}(U) = x$ and $U_{ij} = 0$ for any edge $\{i, j\}$ in G (since G has n nodes, each off-diagonal entry of U corresponds to a pair of nodes of G). Then $\text{TH}(G)$ is a projection of an affine slice of $\mathcal{S}_+^{n+1}$. Furthermore, it can be shown that $\text{TH}(G) = \text{STAB}(G)$ whenever G is perfect. Hence, $\text{TH}(G)$ forms a psd lift of $\text{STAB}(G)$ of size $n+1$. Since the number of facets of $\text{STAB}(G)$ may be exponential in n, this example shows that the psd lift may have dramatically smaller size than the original polytope.

Now that we have seen examples of psd lifts with size smaller than the size of the original polytope, we naturally begin to wonder how far this technique can be pushed. Specifically, given a polytope P, what is the smallest k such that P admits a psd lift of size k? The answer to this question turns out to be closely related to psd rank. Before we can answer the question, however, we need to introduce a special matrix known as the slack matrix.

Definition 1.2.3 ([Yan91])**.** Let $P \subset \mathbb{R}^n$ be a polytope and $Q \subset \mathbb{R}^n$ be a polyhedron with $P \subset Q$. Let $x_1, \ldots, x_v$ be such that $P = \text{conv}(x_1, \ldots, x_v)$ and let $a_j \in \mathbb{R}^n, b_j \in \mathbb{R}$, $(j = 1, \ldots, f)$ be such that $Q = \{x \in \mathbb{R}^n : a_j^T x \leq b_j \ \forall j\}$. Then the *slack matrix* of the pair P, Q, denoted $S_{P,Q}$ is the nonnegative $v \times f$ matrix whose (i, j)-th entry is $b_j - a_j^T x_i$. When $P = Q$ we write $S_P := S_{P,P}$ and we call it the slack matrix of P.

Remark 1.2.4. It is important to note that the slack matrix for a system $P \subset Q$ is not uniquely defined. It depends on the particular descriptions of P and Q that are used. This actually does not pose much of a problem, however, since all of the slack matrices for a system will have the same rank and the same psd rank. Hence, for ease of exposition, we will often sweep this detail under the rug and just refer to "the" slack matrix of $P \subset Q$.

The next theorem gives an answer to our question about psd lifts above: the smallest k such that P admits a psd lift of size k is equal to the psd rank of the slack matrix of P. We refer to this quantity as the psd rank of P. This result is a generalization of Yannakakis' well-known result on extended formulations [Yan91]. The $P = Q$ case was proven in [GPT13] (see also [FMP$^+$12]) and this more general version first appeared in [GRT13b].

Theorem 1.2.5 ([GRT13b, FGP$^+$14])**.** *Let $P \subset \mathbb{R}^n$ be a polytope and $Q \subset \mathbb{R}^n$ be a polyhedron such that $P \subset Q$, and let $S_{P,Q}$ be the slack matrix of the pair P, Q. Then* $\text{rank}_{\text{psd}}\, S_{P,Q}$ *is the smallest integer k for which there exists an affine subspace $\mathcal{L}$ of $\mathcal{S}^k$ and a linear map π such that $P \subset \pi(\mathcal{S}^k_+ \cap \mathcal{L}) \subset Q$.*

Sketch of proof. Let $k = \text{rank}_{\text{psd}}\, S_{P,Q}$. We first show how to construct a spectrahedron $\mathcal{S}^k_+ \cap \mathcal{L}$ of size k such that $P \subset \pi(\mathcal{S}^k_+ \cap \mathcal{L}) \subset Q$ for some linear map π. Let $x_1, \ldots, x_v$ be the vertices of P and let $Q = \{x \in \mathbb{R}^n : a_j^T x \leq b_j \ \forall j = 1, \ldots, f\}$ be a facet description of Q.

Let $A_1, \ldots, A_v, B_1, \ldots, B_f \in \mathcal{S}_+^k$ be a psd factorization of $S_{P,Q}$ of size k:

$$b_j - a_j^T x_i = \langle A_i, B_j \rangle \quad \forall i = 1, \ldots, v,\ j = 1, \ldots, f.$$

Consider the convex set C:

$$C = \{x \in \mathbb{R}^n : \exists A \in \mathcal{S}_+^k \text{ such that } b_j - a_j^T x = \langle A, B_j \rangle \ \forall j = 1, \ldots, f\}. \tag{1.3}$$

It is easy to verify that C is contained between P and Q: indeed $C \subset Q$ because any $x \in C$ satisfies $b_j - a_j^T x \geq 0$ for all $j = 1, \ldots, f$; also $P \subset C$ because the vertices x_i of P satisfy (1.3) with $A = A_i$. Also it is not too difficult to show that C can be expressed in the desired form $C = \pi(\mathcal{S}_+^k \cap \mathcal{L})$ where $\mathcal{L}$ is an affine subspace of $\mathcal{S}^k$ and π is a linear projection map (we refer to Proposition 5.2.6 for the details). This proves the first direction.

Assume now that we can write $P \subset \pi(\mathcal{S}_+^k \cap \mathcal{L}) \subset Q$ where $\mathcal{L}$ is an affine subspace of $\mathcal{S}^k$ and π is a linear map. We show how to construct a psd factorization of $S_{P,Q}$ of size k. Let $C = \pi(\mathcal{S}_+^k \cap \mathcal{L})$. Using a suitable choice of basis for $\mathcal{L}$, we can assume that C has the form:

$$C = \{x \in \mathbb{R}^n \ : \ \exists y \in \mathbb{R}^m \text{ such that } T(x, y) \in \mathcal{S}_+^k\}$$

where $T : \mathbb{R}^n \times \mathbb{R}^m \to \mathcal{S}^k$ is an affine linear map (i.e., T has the form $T(x, y) = U_0 + x_1 U_1 + \cdots + x_n U_n + y_1 V_1 + \cdots + y_m V_m$ for some $U_0, U_1, \ldots, U_n, V_1, \ldots, V_m \in \mathcal{S}^k$). Observe that since $C \subset Q$ we have for any $(x, y) \in \mathbb{R}^n \times \mathbb{R}^m$:

$$T(x, y) \in \mathcal{S}_+^k \ \Rightarrow \ b_j - a_j^T x \geq 0 \ \forall j = 1, \ldots, f.$$

By Farkas' lemma this means that, for any $j = 1, \ldots, f$, there exists $B_j \in \mathcal{S}_+^k$ such that:

$$b_j - a_j^T x = \langle T(x, y), B_j \rangle \quad \forall (x, y) \in \mathbb{R}^n \times \mathbb{R}^m.$$

Furthermore, since $P \subset C$ we know that for any x_i vertex of P there exists y_i such that $T(x_i, y_i) \in \mathcal{S}_+^k$. Thus if we let $A_i = T(x_i, y_i)$ then we get the following psd factorization of size k of the slack matrix $S_{P,Q}$:

$$b_j - a_j^T x_i = \langle A_i, B_j \rangle \quad \forall i = 1, \ldots, v, j = 1, \ldots, f.$$

□

Example 1.2.6. Now we return to the setting where P is the unit square $[-1,1]^2$. The square has four vertices and facets and a quick calculation shows that the slack matrix is:

$$S_P = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \end{pmatrix}.$$

Example 1.2.1 showed us that P has a psd lift of size 3, so by Theorem 1.2.5 we know that there must exist a psd factorization of S_P of size 3. Here is one such factorization:

$$A_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad A_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad A_3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad A_4 = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

$$B_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad B_2 = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad B_3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{pmatrix} \quad B_4 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

This shows that $\text{rank}_{\text{psd}}(S_P) \leq 3$. Using results from later sections (Theorem 1.3.3), it is possible to see that the psd rank of P is exactly 3.

This framework can actually be extended to general convex sets, but we will not require that generality in our results. In many instances, however, it will be preferable to work with polyhedral cones instead of polytopes. The slack matrix definition and the theorem characterizing psd lifts both extend easily to this setting with analogous proofs.

Definition 1.2.7. Let $P \subset \mathbb{R}^n$ and $Q \subset \mathbb{R}^n$ be polyhedral cones with $P \subset Q$. Let $x_1, \ldots, x_v$ be such that $P = \text{cone}(x_1, \ldots, x_v)$ and let $a_j \in \mathbb{R}^n$, $(j = 1, \ldots, f)$ be such that $Q = \{x \in \mathbb{R}^n : a_j^T x \geq 0 \ \forall j\}$. Then the *slack matrix* of the pair P, Q, denoted $S_{P,Q}$ is the nonnegative $v \times f$ matrix whose (i,j)-th entry is $a_j^T x_i$. When $P = Q$ we write $S_P := S_{P,P}$ and we call it the slack matrix of P.

Theorem 1.2.8. *Let $P \subset \mathbb{R}^n$ and $Q \subset \mathbb{R}^n$ be polyhedral cones such that $P \subset Q$, and let $S_{P,Q}$ be the slack matrix of the pair P, Q. Then $\text{rank}_{\text{psd}} S_{P,Q}$ is the smallest integer k for which there exists a linear subspace $\mathcal{L}$ of $\mathcal{S}^k$ and a linear map π such that $P \subset \pi(\mathcal{S}^k_+ \cap \mathcal{L}) \subset Q$.*

1.3 Results and contributions

This section provides an overview of the contributions of this thesis to the study of psd rank, with each subsection covering a chapter of the thesis. The focus here will be on the motivation and results. Proofs and details are available in the corresponding chapters.

1.3.1 Which nonnegative matrices are slack matrices?

Since slack matrices play such an integral role in the geometric interpretation of psd rank, it is natural to ask what properties differentiate slack matrices from general matrices. We investigate these properties in Chapter 2. Clearly, slack matrices must be nonnegative, but are there other properties that they must satisfy? In one sense of the question, the answer is a simple "no" since any nonnegative matrix is a slack matrix of a pair of polyhedral cones $P \subset Q$ that can be constructed from a *rank factorization* of the matrix. If we no longer consider slack matrices of pairs $P \subset Q$ and only consider slack matrices of polytopes (that is, restricting the slack matrix definition to the case $P = Q$), then the answer becomes more subtle. It is clear that not every nonnegative matrix is a slack matrix of a polytope, since every slack matrix will have zero entries where the vertex defining the row is tight to the facet defining the column.

For every nonnegative matrix, the cone spanned by the rows will be contained in the linear span of the rows. Whether a matrix is a slack matrix or not is related to the "tightness" of this containment. Specifically, we show the following in Theorem 2.2.1 and Corollary 2.2.8.

Theorem 1.3.1. *Let $M \in \mathbb{R}_+^{p\times q}$ be a matrix with* $\operatorname{rank}(M) \geq 2$.

(i) M is a slack matrix of a polyhedral cone if and only if

$$\operatorname{cone}(\operatorname{rows}(M)) = \operatorname{lin}(\operatorname{rows}(M)) \cap \mathbb{R}_+^q$$

(ii) M is a slack matrix of a polytope if and only if

$$\operatorname{conv}(\operatorname{rows}(M)) = \operatorname{aff}(\operatorname{rows}(M)) \cap \mathbb{R}_+^q$$

where cone, conv, lin, *and* aff *respectively denote the conic, convex, linear, and affine hull of a set of vectors.*

This theorem gives a clean geometric statement of what it means to be a slack matrix. It is not a perfect answer, however, since deciding whether a matrix satisfies the theorem is computationally equivalent to deciding if a vertex-description of a polytope and a facet-description of a polytope yield the same polytope. This is known as the *polytope verification problem* and its complexity is a long-standing open problem in computational geometry.

We know going forward that we are not going to find a simple, easy-to-compute characterization of slack matrices, since doing so would answer a long-standing open problem. We do, however, offer one more characterization that might yield some intuition on slack matrices, but we need a couple of observations first. First, we observe that the zero pattern of a slack matrix corresponds to the vertex-facet incidence structure of the polytope. Second, it is straightforward to show that the slack matrix of an n-dimensional polytope will have rank $n+1$. Our final result says that if a matrix has the correct zero pattern for a slack matrix and the correct rank for a slack matrix, then it must be a slack matrix.

Theorem 1.3.2. *A nonnegative matrix M with $\operatorname{rank}(M) \geq 2$ is a slack matrix of some polytope if and only if its zero pattern is an incidence matrix of some $(\operatorname{rank}(M)-1)$-dimensional polytope and $\mathbb{1}$, the all ones vector, is contained in the column span of M.*

1.3.2 Polytopes of minimum positive semidefinite rank

From Proposition 1.1.3, we know that the slack matrix of an n-dimensional polytope (which has rank $n+1$) must have psd rank bounded below by $\sqrt{n}$. In Chapter 3, we investigated how tight this bound could be and what types of polytopes achieve this bound.

Using the zero pattern of slack matrices and an induction argument, we were able to improve upon the standard lower bound.

Theorem 1.3.3. *If $P \subset \mathbb{R}^n$ is a full-dimensional polytope with slack matrix S_P, then the psd rank of S_P is at least $n+1$. Furthermore, if $\operatorname{rank}_{\text{psd}}(S_P) = n+1$, then every $\mathcal{S}_+^{n+1}$-factorization of S_P uses only rank one matrices as factors.*

The psd lifts we saw in Examples 1.2.1 and 1.2.2 immediately show that this new lower bound can be tight. We call polytopes that achieve the lower bound *psd minimal* polytopes. Using the characterization of Theorem 1.3.3, we were able to classify all psd minimal polytopes in two and three dimensions. In $\mathbb{R}^2$, a straightforward application of the characterization let us show that the psd minimal polytopes are precisely the triangles and quadrilaterals (Theorem 3.4.7). In $\mathbb{R}^3$, the situation was more subtle. We used Euler's formula and the fact that any face of a psd minimal polytope must also be psd minimal to rule out most combinatorial types of polytope in $\mathbb{R}^3$. From here, we constructed a combinatorial structure that could not appear in a psd minimal polytope and used this structure to rule out the remaining non-psd minimal combinatorial types. Now there were four remaining combinatorial classes that were always psd minimal and two combinatorial classes (octahedra and cuboids) that contained some psd minimal and some non-psd minimal polytopes. The final step was characterizing the psd minimal octahedra and cuboids through a planarity condition which we called "biplanarity." Finally, we were able to show that the psd minimal polytopes in $\mathbb{R}^3$ are precisely the simplices, quadrilateral pyramids, bisimplicies, triangular prisms, "biplanar" octahedra, and "biplanar" cuboids (Theorem 3.4.11).

By the Lovász theta body construction (see Example 1.2.2), it is clear that the stable set polytope STAB(G) is psd minimal if G is a perfect graph. It is not immediately clear whether there exist non-perfect graphs such that STAB(G) is psd minimal. Using the strong perfect graph theorem, we were able to prove that STAB(G) is psd minimal if and only if G is a perfect graph (Theorem 3.4.12).

1.3.3 Rational and real positive semidefinite rank can be different

In their seminal paper on nonnegative rank [CR93], Cohen and Rothblum asked whether nonnegative rank could increase when the nonnegative factorizations are restricted to have rational entries, as opposed to the usual real entries. This question has remained open and inspired us to investigate the analogous question for psd rank, which we answer affirmatively by exhibiting an example in Chapter 4.

We define rational psd rank in the same manner as the usual psd rank, except that we

$$M = \begin{pmatrix} 0 & 0 & 2 & 1 & 0 & 1 \\ 1 & 0 & 0 & 2 & 0 & 1 \\ 0 & 1 & 2 & 0 & 0 & 1 \\ 1 & 2 & 0 & 0 & 0 & 1 \\ 0 & 0 & 2 & 1 & 1 & 0 \\ 1 & 0 & 0 & 2 & 1 & 0 \\ 0 & 1 & 2 & 0 & 1 & 0 \\ 1 & 2 & 0 & 0 & 1 & 0 \end{pmatrix}$$

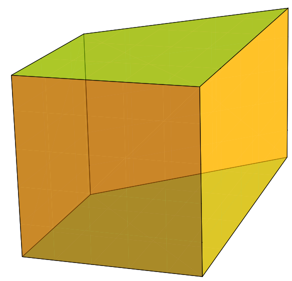


Figure 1.2. The example matrix showing rational and real psd rank can be different, along with the polytope from which it arises.

require that the matrix factors have rational entries. Then clearly rational psd rank is an upper bound to the usual psd rank. We show that there exists an octahedron for which rational psd rank is five and usual psd rank is four (Proposition 4.0.5). The octahedron is biplanar, and hence, psd minimal, and the proof relies on the fact that any size four psd factorization must use only rank one matrices as factors.

1.3.4 Worst-case results for positive semidefinite rank of polygons

Chapter 5 contains two results on the psd rank of polygons. The first is a lower bound on the psd rank of "bad" polygons. The second is an upper bound that applies to any polygon, both the "good" and "bad."

We define a *generic* polytope to be a polytope such that the coordinates of its vertices form an algebraically independent set over the rationals, i.e. the vertex coordinates do not satisfy any non-trivial polynomial equation with rational cooefficients. Such polytopes are clearly the "worst-case" scenario and we expect them to have high psd rank. Using some facts about semidefinite programming and applying the Tarski-Seidenberg Theorem of real algebraic geometry, we were able to show the following.

Theorem 1.3.4. *If $P \subset \mathbb{R}^n$ is a generic polytope with v vertices, then its psd rank is at least $(nv)^{\frac{1}{4}}$.*

One nice application of this theorem is to see that psd rank can vary dramatically over a fixed combinatorial type of polytope. For example, consider 2^n-gons in the plane. Ben-Tal and Nemirovski showed that the nonnegative rank, and hence the psd rank, of the regular 2^n-gon is at most $2n$ [BTN01]. On the other hand, this result shows that a generic 2^n-gon has psd rank at least $2^{n/4}$. Hence, there is an exponentially large gap between the smallest and largest psd ranks achieved by 2^n-gons.

The next result relies on the fact that every hexagon has psd rank four. We proved this by showing that for every hexagon, there exists a biplanar octahedron that projects onto the hexagon. With this fact in hand, we could show that every v-gon has psd rank at most $4\left\lceil\frac{v}{6}\right\rceil$ by writing a v-gon as the convex hull of many hexagons. Finally, this result can be extended to general nonnegative rank three matrices by viewing a general rank three matrix as a slack matrix of $P \subset Q$ for two polygons P and Q.

Theorem 1.3.5. *Let M be a nonnegative $p \times q$ matrix with usual rank three. Then $\operatorname{rank}_{\text{psd}}(M) \leq 4\left\lceil\frac{\min\{p,q\}}{6}\right\rceil$. In particular, the psd rank of a v-gon is at most $4\left\lceil\frac{v}{6}\right\rceil$.*

1.3.5 Complexity of computing positive semidefinite rank

Computing the psd rank of a matrix is a difficult problem in practice, but little is known about the theoretical computational complexity of computing psd rank. In Chapter 6, we investigated two natural questions. First, is it NP-hard to compute psd rank? Second, if we fix the size of the psd factorization we are looking for, then does there exist a polynomial time algorithm for deciding if a psd factorization exists?

Both of these questions have been answered in the affirmative in the case of nonnegative rank. Vavasis [Vav09] used a reduction to 3-SAT to show that computing nonnegative rank is NP-hard. Arora et al. [AGKM12] showed that there exists an algorithm to decide if a $p \times q$ matrix has nonnegative rank r that runs in time $pq^{O(r^2 2^r)}$.

We have been able to make partial progress on both of these questions. We were able to show that for certain special values of the rank of the input matrix and the size of the

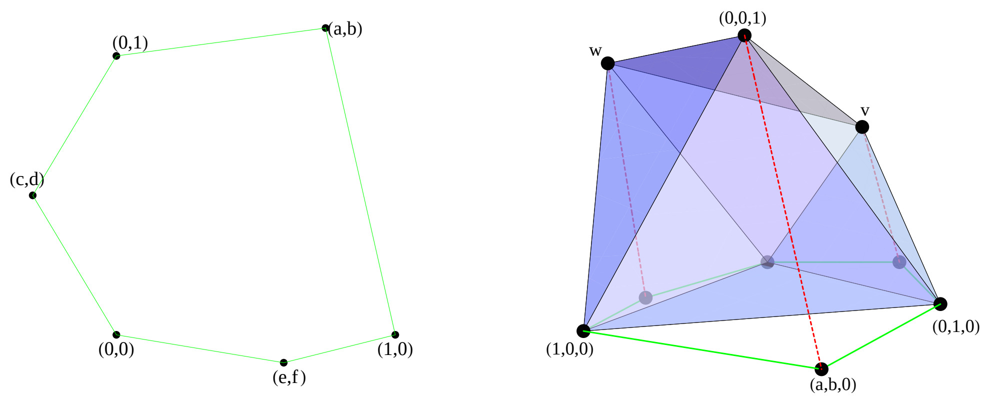


Figure 1.3. This figure depicts the lift of a hexagon to an octahedron showing that all hexagons have psd rank four. The left picture shows a hexagon in a normalized form. The right picture shows the biplanar octahedron and its linear projection onto the hexagon.

psd factorization, there does indeed exist a polynomial time algorithm for deciding if a psd factorization exists. In particular, we were able to show the following.

Theorem 1.3.6. *Suppose M is a nonnegative matrix of size $p \times q$ with rank equal to $\binom{k+1}{2}$. Then using the Blum-Shub-Smale model of complexity, it can be decided if there exists a psd factorization of M of size k in time $(pq)^{O(k^5)}$. In particular, if k is fixed, then this problem can be solved in polynomial time.*

The proof relies on reducing the existence of a psd factorization of M to the existence of an invertible linear map that maps the psd cone so that it is sandwiched between two particular polyhedral cones. With this reduction, we can apply quantifier elimination to get the result. For other values of k, unfortunately, the linear map in question is no longer invertible and the argument falls apart.

We have not been able to show that psd rank is NP-hard to compute, but we were able to show that a closely related quantity, the square root rank, is NP-hard to compute. The square root rank of a nonnegative matrix M is defined as the minimum rank of any possible entry-wise square root of M. It is straightforward to see that square root rank

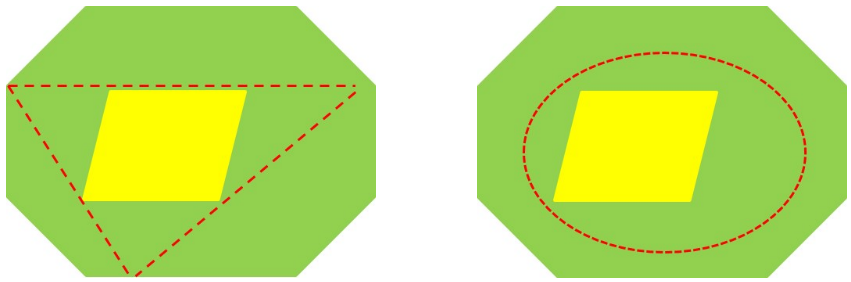


Figure 1.4. Vavasis reduced the problem of finding a nonnegative factorization to a problem known as "Intermediate Simplex" which we depict in the left figure. This problem asks if there exists a simplex nested between two given polytopes. The analogous problem for psd factorizations can be reduced to a geometric question asking if there exists a free spectrahedron nested between two given polytopes. This situation is depicted in the right figure. We used this reduction in the proof of Theorem 1.3.6.

is equivalent to a restricted version of psd rank where the psd factors are required to be rank one matrices (for more details on square root rank, see Section 3.2). By reducing the NP-complete problem PARTITION to a question of square root rank, we showed in Theorem 6.2.1 that computing square root rank is NP-hard.

1.3.6 Spaces of positive semidefinite factorizations

In Chapter 7, we fix a nonnegative matrix M and consider the set of all valid psd factorizations of M as a topological space. Specifically, if $M \in \mathbb{R}_+^{p\times q}$ has psd rank k, then we define the space of factorizations $\mathcal{SF}(M)$ to be the set of all valid psd factorizations of size k. Hence, $\mathcal{SF}(M)$ lives in the space $\left(\mathcal{S}^k\right)^{p+q}$. Now we observe that for any invertible $k \times k$ matrix L and psd factorization $(A_1, \ldots, B_q)$, the matrices $\left(L^T A_1 L, \ldots, L^{-1} B_q L^{-T}\right)$ also form a valid psd factorization. This group action makes it natural to consider the space $\mathcal{SF}(M)/GL(k)$, the space of psd factorization orbits.

We were able to prove several results about the space $\mathcal{SF}(M)/GL(k)$ in the special case that $\operatorname{rank}(M)$ is $\binom{k+1}{2}$. The main result we were able to show is the following.

Theorem 1.3.7. *Let $M \in \mathbb{R}_+^{p\times q}$ have psd rank k and usual rank $\binom{k+1}{2}$. Fix a rank fac-*

torization of $M = UV$ where u_i is the ith row of U and v_j is the jth column of V. Let $P = \text{cone}(u_1, \ldots, u_p)$ and $Q = \left\{x \,:\, v_j^T x \geq 0 \text{ for all } j\right\}$ be the polyhedral cones generated by this rank factorization so that $M = S_{P,Q}$. Then $\mathcal{SF}(M)/GL(k)$ is homeomorphic to the space of all linear images C of $\mathcal{S}_+^k$ such that $P \subset C \subset Q$.

In other words, if we view M as a slack matrix of two polyhedral cones, then the space of factorization orbits is equivalent to the space of images of the psd cone that nest between the two polyhedral cones.

This geometric perspective gives a new way to look at psd factorizations and can assist in the construction of interesting factorizations. For example, during our first efforts studying psd rank, all of the explicit psd factorizations that we constructed had the property that either the row or column factors could all be chosen to be rank one matrices. When questioned about this by a reviewer, we were unsure if this indicated a larger phenomenon or whether our examples were lacking diversity. With this geometrical machinery, it is quite easy to construct a matrix whose psd factorizations must always use at least one rank two matrix on a row and at least one rank two matrix on a column (see Example 7.0.7).

For the special case where M has rank three and we are considering 2×2 psd factorizations, we can show that the space of factorization orbits is connected (Proposition 7.0.8). The proof amounts to showing that for any two polytopes P, Q and any two ellipses E, F such that the ellipses are nested between the polytopes (i.e. $P \subset \{E, F\} \subset Q$), we can find a path of ellipses that connect E to F such that every ellipse in the path is also nested between P and Q. This result contrasts with the case of nonnegative rank, where in the first nontrivial case (rank three matrices of nonnegative rank three), it is known that the space of factorizations is disconnected [MSvS03].

1.3.7 Notation

We conclude with a quick warning about notation. Since the remaining chapters draw on different source material, there may be some notational inconsistencies between chapters. For example, in Chapter 2 a rank factorization of M is written as $M = AB$ while in later chapters the letters A and B are reserved for psd factors and a rank factorization is written

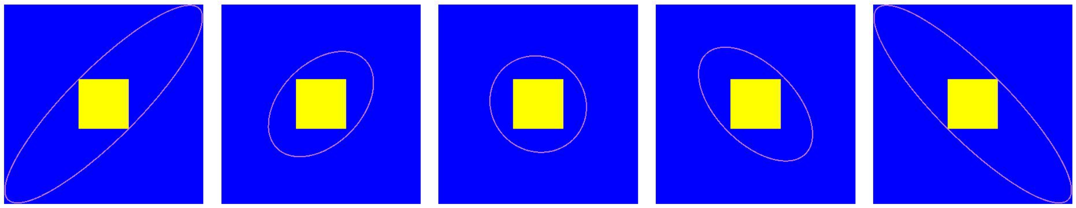


Figure 1.5. This sequence of pictures shows how we can transform one ellipse into another via a sequence of ellipses that remain sandwiched between the two bounding polytopes. This technique was used in the proof of Proposition 7.0.8.

as $M = UV$. The meaning of any notation should be made clear by the surrounding context.

Chapter 2

WHICH NONNEGATIVE MATRICES ARE SLACK MATRICES? [GGK+13]

2.1 Introduction

This chapter is concerned with a class of nonnegative matrices with real entries, called *slack matrices*, that arise naturally from polyhedral cones and polytopes. As we saw in Section 1.2, given a polytope $P \subset \mathbb{R}^n$ with vertices $v_1, \ldots, v_p$ and facet inequalities $a_j^T x \leq \beta_j$ for $j = 1, \ldots, q$, a *slack matrix* of P is the $p \times q$ nonnegative matrix whose (i, j)-entry is $\beta_j - a_j^T v_i$, the *slack* (distance from equality), of the ith vertex v_i in the jth facet inequality $a_j^T x \leq \beta_j$ of P. A similar definition holds for polyhedral cones.

Slack matrices form an interesting class of nonnegative matrices with many special properties. Most obviously, if M is a slack matrix of a polytope P, then the zeros in M record the face lattice of P and hence the combinatorial structure of P. In its entirety, M specifies an embedding of P up to affine transformation. However, slack matrices carry much more (and surprising) information about P. In [Yan91], Yannakakis proved that the *nonnegative rank* of a slack matrix of P is the minimum k such that P is the linear image of an affine slice of the positive orthant $\mathbb{R}^k_+$. We use $\mathbb{R}_+$ to denote the set of nonnegative real numbers. The nonnegative rank of a matrix $M \in \mathbb{R}^{p \times q}_+$ is the smallest k such that there exist vectors $a_1, \ldots, a_p \in \mathbb{R}^k_+$ and $b_1, \ldots, b_q \in \mathbb{R}^k_+$ such that $M_{ij} = a_i^T b_j$. Affine slices of positive orthants that project onto P are called *polyhedral lifts* or *polyhedral extended formulations* of P and the smallest k such that $\mathbb{R}^k_+$ admits a lift of P is called the *(polyhedral) extension complexity* or nonnegative rank of P. If the extension complexity of P is small (polynomial in the dimension of P), then usually it is possible to optimize a linear function over P in polynomial time by optimizing an appropriate function on the lift. This is a powerful technique in optimization that yields polynomial time algorithms for linear optimization over complicated polytopes. There are many instances of n-dimensional polytopes with exponentially many (in n) facets that allow small polyhedral lifts.

Yannakakis' result was generalized in [GPT13] to lifts of convex sets by affine slices of convex cones via *cone factorizations* of *slack operators*. Even in the larger context of cone lifts of convex sets, the case of polytopes is the simplest and the theory relies on slack matrices of polytopes and their factorizations through cones. Thus, understanding the structure of these matrices is fundamental for this theory. There are several phenomena that occur in the class of nonnegative matrices that have not yet been observed for slack matrices. For instance, an important open question is whether there exists a family of slack matrices of polytopes that exhibit an exponential gap between nonnegative rank and *positive semidefinite rank*. (If $\mathcal{S}_+^k$ denotes the cone of $k \times k$ real symmetric positive semidefinite matrices, then the positive semidefinite rank of a matrix $M \in \mathbb{R}_+^{p\times q}$ is the smallest k such that there exists matrices $A_i \in \mathcal{S}_+^k$, $i = 1, \ldots, p$ and $B_j \in \mathcal{S}_+^k$, $j = 1, \ldots, q$ such that $M_{ij} = \langle A_i, B_j \rangle$.) While there are simple families of matrices that exhibit even arbitrarily large gaps between nonnegative and positive semidefinite ranks [GPT13, Example 5], no family of slack matrices with this property is known. Such a family would be a clear witness for the power of *semidefinite programming* over linear programming in lifts of polytopes.

This chapter was motivated by the many open questions about slack matrices that rely on understanding the structure of these matrices. We establish two main characterizations of slack matrices of polyhedral cones and polytopes. In Section 2.2 we establish linear algebraic characterizations: Theorem 2.2.1 for cones and Theorem 2.2.6 for polytopes. In Section 2.4 we give combinatorial characterizations: Theorem 2.4.1 for polytopes and Theorem 2.4.3 for polyhedral cones. In Section 2.3 we use our characterization from Section 2.2 to give an algorithm for recognizing slack matrices. The computational complexity of this problem is unknown and is equivalent to the *polyhedral verification problem*. There are several further geometric and complexity results about slack matrices throughout the chapter.

Notation: For a set of vectors $\mathcal{A} = \{a_1, \ldots, a_p\}$, $\text{cone}(\mathcal{A}) := \{\sum \lambda_i a_i \,:\, \lambda_i \geq 0\}$ is the cone spanned by $\mathcal{A}$; $\text{conv}(\mathcal{A}) := \{\sum \lambda_i a_i \,:\, \lambda_i \geq 0, \sum \lambda_i = 1\}$ is the convex hull of $\mathcal{A}$; $\text{lin}(\mathcal{A}) := \{\sum \lambda_i a_i \,:\, \lambda_i \in \mathbb{R}\}$ is the linear span of $\mathcal{A}$, and $\text{aff}(\mathcal{A}) := \{\sum \lambda_i a_i \,:\, \sum \lambda_i = 1\}$ is the affine span of $\mathcal{A}$. The above sets can also be defined for an infinite $\mathcal{A}$ by taking unions over all finite subsets of $\mathcal{A}$. For a $n \times q$ matrix M, we let $\text{rows}(M)$ and $\text{cols}(M)$ denote the sets of all rows and columns of M. We let $\mathcal{A} \cdot M$ be the set of vectors $\{x^T M \,:\, x \in \mathcal{A}\}$. For a set $K \subset \mathbb{R}^n$, $\text{lineal}(K)$ is the largest subspace contained in K, known as the *lineality space* of K. The dimension of a polytope P, $\dim(P)$ is the dimension of $\text{aff}(P)$, the affine hull of P, and the dimension of a cone K is the dimension of $\text{lin}(K)$. For any matrix $M \in \mathbb{R}^{p\times q}$ of rank k, we will call a factorization of the form $M = AB$ with $A \in \mathbb{R}^{p\times k}$, $\text{rank}\,(A) = k$ and $B \in \mathbb{R}^{k\times q}$, $\text{rank}\,(B) = k$ a *rank factorization* of M.

2.2 Geometric characterizations of slack matrices

2.2.1 Slack matrices of polyhedral cones

Let $\mathbf{0}$ denote the all-zero vector and consider the polyhedral cone

$$K = \{x \in \mathbb{R}^n \,:\, x^T B \geq \mathbf{0}\} = \mathbb{R}^p_+ \cdot A$$

in $\mathbb{R}^n$ constrained by the columns of the matrix $B \in \mathbb{R}^{n\times q}$ and generated by the rows of the matrix $A \in \mathbb{R}^{p\times n}$. We call (the set of rows of) A a $\mathcal{V}$*-representation* and (the set of columns of) B an $\mathcal{H}$*-representation* of K. As referenced in the introduction, the *slack matrix* of K with respect to the representation (A, B) is $S = AB \in \mathbb{R}^{p\times q}_+$. Its (i, j)-entry records the "slack" of the ith generator of K with respect to the jth inequality of K in the given description of K.

Let $\mathcal{S}_K$ denote the set of all slack matrices of K. For $S \in \mathcal{S}_K$, any matrix obtained by scaling the rows and columns of S by positive reals is again in $\mathcal{S}_K$ since scaling the vectors in a $\mathcal{V}$ and/or $\mathcal{H}$-representation of K does not change K. Also, $\mathcal{S}_K$ can have matrices of different sizes as adding redundant inequalities and/or generators to the representations of K does not change K. From

$$\begin{aligned}(\mathbb{R}^n \cdot B) \cap \mathbb{R}^q_+ = K \cdot B = (\mathbb{R}^p_+ \cdot A) \cdot B = \mathbb{R}^p_+ \cdot S \qquad&\\ \subseteq (\mathbb{R}^p \cdot S) \cap \mathbb{R}^q_+ = (\mathbb{R}^p \cdot AB) \cap \mathbb{R}^q_+ \subseteq (\mathbb{R}^n \cdot B) \cap \mathbb{R}^q_+&\end{aligned}$$

we find that $\mathbb{R}^p_+ \cdot S = \mathbb{R}^p \cdot S \cap \mathbb{R}^q_+$ which says that the *cone generated by the rows of* S coincides with the *nonnegative part of the row span of* S. In fact, this relation characterizes slack matrices of cones as we now show.

Theorem 2.2.1. *A nonnegative matrix $M \in \mathbb{R}_+^{p\times q}$ is a slack matrix of a polyhedral cone if and only if*

$$\mathbb{R}_+^p \cdot M = \mathbb{R}^p \cdot M \cap \mathbb{R}_+^q, \tag{2.1}$$

or in other words, the cone spanned by the rows of M coincides with the nonnegative part of the row span of M.

Proof. It remains to show that every matrix $M \in \mathbb{R}_+^{p\times q}$ with $\mathbb{R}_+^p \cdot M = \mathbb{R}^p \cdot M \cap \mathbb{R}_+^q$ is a slack matrix of some cone. Let $k = \operatorname{rank}(M)$ and consider a rank factorization $M = AB$ with $A \in \mathbb{R}^{p\times k}$ and $B \in \mathbb{R}^{k\times q}$. Let $K = \operatorname{cone}(\{a_1, \ldots, a_p\})$ and $\widetilde{K} = \{x \in \mathbb{R}^k : x^T b^j \geq 0,\ j = 1, \ldots, q\}$ where a_i is the ith row of A and b^j is the jth column of B. We need to show that $K = \widetilde{K}$.

Since M is nonnegative, we get that $K \subseteq \widetilde{K}$. In order to show the inclusion $\widetilde{K} \subseteq K$, consider a vector x from $\widetilde{K}$. Since the matrix $A \in \mathbb{R}^{p\times k}$ has full column rank, the vector x^T lies in $\mathbb{R}^p \cdot A$, and thus $x^T B$ lies in $\mathbb{R}^p \cdot M \cap \mathbb{R}_+^q$. Thus, due to equation (2.1) the vector $x^T B$ lies in $\mathbb{R}_+^p \cdot M$; i.e., $x^T B$ lies in $\mathbb{R}_+^p \cdot AB$. Since the matrix B has full row rank the vector x^T hence lies in $\mathbb{R}_+^p \cdot A$; i.e., x lies in K. ☐

Recall that the dual cone of K is the cone

$$K^\star = \{y \in \mathbb{R}^n : x^T y \geq 0 \text{ for all } x \in K\} = \{y \in \mathbb{R}^n : Ay \geq 0\} = B \cdot \mathbb{R}_+^q.$$

Thus, if S is the slack matrix of K, then S^T is a slack matrix of $K^\star$ and we get the following.

Proposition 2.2.2. *A nonnegative real matrix is a slack matrix of a polyhedral cone if and only if its transpose is also the slack matrix of a polyhedral cone.*

In particular, we obtain the following consequence of Theorem 2.2.1.

Corollary 2.2.3. *A nonnegative matrix $M \in \mathbb{R}_+^{p\times q}$ is a slack matrix of a polyhedral cone if and only if*

$$M \cdot \mathbb{R}_+^q = M \cdot \mathbb{R}^q \cap \mathbb{R}_+^p, \tag{2.2}$$

or in other words, the cone spanned by the columns of M coincides with the nonnegative part of the column span of M.

We say that a matrix M satisfies the *row cone generating condition* ($RCGC$) if (2.1) holds and the *column cone generating condition* ($CCGC$) if (2.2) holds.

Corollary 2.2.4. *For a nonnegative matrix $M \in \mathbb{R}_+^{p\times q}$ the following statements are pairwise equivalent:*

- *M is a slack matrix of a polyhedral cone.*
- *M satisfies the RCGC.*
- *M satisfies the CCGC.*

The equivalence of RCGC and CCGC for a general nonnegative matrix is not obvious. However, its proof becomes transparent via the theory of slack matrices of polyhedral cones and cone duality.

For a nonnegative M with RCGC/CCGC, the proof of Theorem 2.2.1 showed how to produce a cone K such that $M \in \mathcal{S}_K$, which is captured by the next lemma.

Lemma 2.2.5. *Let $M \in \mathbb{R}_+^{p\times q}$ be the slack matrix of a polyhedral cone and let $M = AB$ be a rank factorization of M. Then if K is the cone generated by the rows of A, the columns of B form an $\mathcal{H}$-representation of K. In particular, $M \in \mathcal{S}_K$.*

Proof. For the slack matrix M equation (2.1) is valid, and thus the statement follows from the proof of Theorem 2.2.1. ☐

2.2.2 Slack matrices of polytopes

We now investigate the slack matrices of polytopes. Let $V \in \mathbb{R}^{p\times n}$ and $P = \operatorname{conv}(\operatorname{rows}(V))$ be the polytope in $\mathbb{R}^n$ that is the convex hull of the rows of V. Suppose also that $P = \{x \in \mathbb{R}^n : Wx \leq w\}$ with $W \in \mathbb{R}^{q\times n}$ and $w \in \mathbb{R}^q$. To avoid unnecessary inconveniences, we assume that $\dim(P) \geq 1$. We call (the set of rows of) V a $\mathcal{V}$-representation and (the set of columns of) $[w, -W]^T$ an $\mathcal{H}$-representation of P. The *slack matrix* of P with respect to the representation (V, W, w) is then

$$S = [\mathbb{1}, V] \cdot [w, -W]^T \in \mathbb{R}_+^{p\times q}\,. \tag{2.3}$$

We denote the set of all slack matrices of P by $\mathcal{S}_P$. Clearly, scaling the columns of a slack matrix of P by positive scalars yields another slack matrix of P, because scaling the vectors in an $\mathcal{H}$-representation of P yields another $\mathcal{H}$-representation of P. However, we cannot scale the rows of a matrix $S \in \mathcal{S}_P$ and still stay in $\mathcal{S}_P$.

The matrix S is also the slack matrix of the *homogenization cone* of P:

$$P^h = \mathbb{R}^p_+ \cdot [\mathbb{1}, V] = \{(x_0, x) \in \mathbb{R} \times \mathbb{R}^n \,:\, Wx \leq x_0 w\} \tag{2.4}$$

with respect to the representation $([\mathbb{1}, V], \begin{bmatrix} w^T \\ -W^T \end{bmatrix})$. Since $\dim(P) \geq 1$, there is some $c \in \mathbb{R}^n$ with

$$\max\{c^T x \,:\, x \in P\} - \min\{c^T x \,:\, x \in P\} = 1\,,$$

and hence, due to LP-duality, we get

$$(1, \mathbf{0}^T) \in \mathbb{R}^q \cdot (w, W) \text{ and so also, } (1, \mathbf{0}^T) \in \mathbb{R}^q \cdot (w, -W). \tag{2.5}$$

From (2.3) and (2.5) we get that $\mathbb{1} \in S \cdot \mathbb{R}^q$, the column span of S. These properties characterize the slack matrices of polytopes of dimension at least one:

Theorem 2.2.6. *A matrix $M \in \mathbb{R}^{p \times q}_+$ with* $\operatorname{rank}(M) \geq 2$ *is a slack matrix of a polytope if and only if M is a slack matrix of a polyhedral cone and $\mathbb{1} \in M \cdot \mathbb{R}^q$.*

Proof. It suffices to show that a matrix $M \in \mathbb{R}^{p \times q}_+$ with $\mathbb{1} \in M \cdot \mathbb{R}^q$ that is the slack matrix of some cone $K \subseteq \mathbb{R}^n$ with respect to a representation (A, B) is also the slack matrix of some polytope. To construct such a polytope, choose any $\mu \in \mathbb{R}^q$ such that $\mathbb{1} = M\mu$ and define $c = B\mu$. Since $M = AB$, we have $Ac = \mathbb{1}$ and each row of A must be contained in the hyperplane defined by $x^T c = 1$. Thus, M will be a slack matrix of the polytope $P = \operatorname{conv}(\operatorname{rows}(A))$. □

Corollary 2.2.7. *A matrix $M \in \mathbb{R}^{p \times q}_+$ with* $\operatorname{rank}(M) \geq 2$ *is a slack matrix of some polytope if and only if it satisfies the RCGC (or, equivalently, the CCGC) and $\mathbb{1} \in M \cdot \mathbb{R}^q$ holds.*

Theorem 2.2.1 geometrically characterizes the slack matrices of cones as those matrices $M \in \mathbb{R}^{p \times q}_+$ that satisfy

$$\operatorname{cone}(\operatorname{rows}(M)) = \operatorname{lin}(\operatorname{rows}(M)) \cap \mathbb{R}^q_+\,. \tag{2.6}$$

There is an analogous geometric characterization of slack matrices of polytopes.

Corollary 2.2.8. *A matrix* $M \in \mathbb{R}_+^{p\times q}$ *with* $\operatorname{rank}(M) \geq 2$ *is a slack matrix of some polytope if and only if*

$$\operatorname{conv}(\operatorname{rows}(M)) = \operatorname{aff}(\operatorname{rows}(M)) \cap \mathbb{R}_+^q \,. \tag{2.7}$$

Proof. First, suppose that M is a slack matrix of some polytope. Then by Corollary 2.2.7, we have that M satisfies (2.6) and $\mathbb{1} \in M \cdot \mathbb{R}^q$. Hence, there exists some $c \in \mathbb{R}^q$ such that $Mc = \mathbb{1}$ and the affine hyperplane $L = \{x \in \mathbb{R}^q \,:\, x^T c = 1\}$ contains the rows of M. Intersecting L with both sides of (2.6), we obtain (2.7).

For the reverse implication, let $M \in \mathbb{R}_+^{p\times q}$ be a nonnegative matrix satisfying (2.7). Using any isometry φ between the d-dimensional affine subspace $\operatorname{aff}(\operatorname{rows}(M))$ and $\mathbb{R}^d$, we find that M is a slack matrix of the φ-image of the polytope defined in (2.7). □

We have seen above that every slack matrix of a polytope P has the all-ones vector in its column span and is also a slack matrix of the homogenization cone P^h of P. The next example shows that not all slack matrices of P^h are slack matrices of P. This does not even hold for the slack matrices of P^h that have the all-ones vector in their column span.

Example 2.2.9. Let P be the square $[-1, 1]^2$. The matrix

$$M = \begin{pmatrix} \frac{4}{3} & 0 & \frac{4}{3} & 0 \\ 2 & 0 & 0 & 2 \\ 0 & 2 & 2 & 0 \\ 0 & 4 & 0 & 4 \end{pmatrix} = \begin{pmatrix} \frac{2}{3} & \frac{2}{3} & \frac{2}{3} \\ 1 & 1 & -1 \\ 1 & -1 & 1 \\ 2 & -2 & -2 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix}$$

is in $\mathcal{S}_{P^h}$ and $\mathbb{1}$ is in the column span of M. It is clear, however, that M is not in $\mathcal{S}_P$ since each facet of $[-1, 1]^2$ is equidistant from the two vertices not on the facet. On the other hand, since M has the RCGC/CCGC and $\mathbb{1}$ is in its column span, it is the slack matrix of some other polytope Q. To obtain it, write a new rank factorization of M (note that $\operatorname{rank}(M) = 3$) so that the first factor contains the all ones vector as its first column as follows:

$$M = \begin{pmatrix} \frac{2}{3} & \frac{2}{3} & \frac{2}{3} \\ 1 & 1 & -1 \\ 1 & -1 & 1 \\ 2 & -2 & -2 \end{pmatrix} UU^{-1} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix}, \quad U = \begin{pmatrix} 1 & 0 & 0 \\ 1/4 & 1 & 0 \\ 1/4 & 0 & 1 \end{pmatrix}$$

to get

$$M = \begin{pmatrix} \frac{4}{3} & 0 & \frac{4}{3} & 0 \\ 2 & 0 & 0 & 2 \\ 0 & 2 & 2 & 0 \\ 0 & 4 & 0 & 4 \end{pmatrix} = \begin{pmatrix} 1 & \frac{2}{3} & \frac{2}{3} \\ 1 & 1 & -1 \\ 1 & -1 & 1 \\ 1 & -2 & -2 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 3/4 & -5/4 & -1/4 & -1/4 \\ -1/4 & -1/4 & 3/4 & -5/4 \end{pmatrix}.$$

By Lemma 2.2.5, M is the slack matrix of the cone with $\mathcal{V}$-representation the rows of the first factor and $\mathcal{H}$-representation the columns of the second factor. Assuming the coordinates of this three-dimensional cone are x_0, x_1, x_2, and slicing the cone with the hyperplane $\{(x_0, x_1, x_2) : x_0 = 1\}$ gives a polytope Q with vertices $(2/3, 2/3), (1, -1), (-1, 1), (-2, -2)$ and $\mathcal{H}$-representation given by the columns of the second factor. Then $M \in \mathcal{S}_Q$.

2.2.3 Further results on slack matrices of cones and polytopes

In this section we derive some more insight into the geometric relations between cones, polytopes, and their slack matrices that will be useful in later parts of this chapter. We return to the setup used earlier: K is assumed to be a cone and S the slack matrix of K with respect to its representation (A, B) where $A \in \mathbb{R}^{p \times n}$ and $B \in \mathbb{R}^{n \times q}$.

First, we will show that every slack matrix of a cone is the slack matrix of some pointed cone. Recall that we use $\text{lin}(K)$ to denote the linear hull of K and $\text{lineal}(K)$ to denote the lineality space of K. Then we have $\text{lin}(K) = \mathbb{R}^p \cdot A$ and $\text{lineal}(K) = \text{leftkernel}(B)$. A cone K is *pointed* if $\text{lineal}(K) = \{\mathbf{0}\}$. Define

$$L := \text{lin}(K) \cap \text{lineal}(K)^{\perp} = (\mathbb{R}^p \cdot A) \cap (B \cdot \mathbb{R}^q).$$

Then we have

$$\text{lin}(K) = L + \text{lineal}(K)$$

(where the summands are orthogonal to each other) and

$$K = (K \cap L) + \operatorname{lineal}(K)\,,$$

where $K \cap L \subseteq L$ is a pointed (i.e., having trivial lineality space) cone with $\dim(K \cap L) = \dim(L)$. Denoting by $A' \in \mathbb{R}^{p \times n}$ the matrix obtained from A by orthogonal projections of all rows to L, we have

$$K \cap L = \mathbb{R}^p_+ \cdot A' \qquad \text{and} \qquad S = A'B\,.$$

By mapping L isometrically to $\mathbb{R}^{\dim(L)}$, we thus find that S is a slack matrix of the pointed cone that is the image of $K \cap L$ under that map and we get the following:

Lemma 2.2.10. *A matrix is a slack matrix of a polyhedral cone if and only if it is a slack matrix of some pointed polyhedral cone.*

If the cone K is pointed, then for every zero-row of $S = AB$ the corresponding row of A is a zero-row as well. Hence, removing any zero-row from S results in another slack matrix of K. A similar statement clearly holds for adding zero-rows.

Lemma 2.2.11. *If a matrix S is a slack matrix of a pointed polyhedral cone K then every matrix obtained from S by adding or removing zero-rows is a slack matrix of K as well.*

Lemmas 2.2.10 and 2.2.11 together also imply this statement:

Lemma 2.2.12. *If a matrix is a slack matrix of some polyhedral cone then every matrix obtained from it by adding or removing zero-rows is a slack matrix of some polyhedral cone as well.*

Let us further investigate the linear map $x \mapsto x^T B$. It induces the isomorphism

$$L \xrightarrow[\text{isomorphism}]{\cdot B} \mathbb{R}^p \cdot S \tag{2.8}$$

between the linear space L and the row span of S because of the relations:

$$L \subseteq \operatorname{lineal}(K)^\perp = \operatorname{leftkernel}(B)^\perp$$

and

$$L \cdot B = (L + \mathrm{lineal}(K)) \cdot B = \mathrm{lin}(K) \cdot B = (\mathbb{R}^p \cdot A) \cdot B = \mathbb{R}^p \cdot S\,.$$

It also induces the isomorphism

$$K \cap L \xrightarrow[\text{isomorphism}]{\cdot B} \mathbb{R}^p_+ \cdot S$$

between the cone $K \cap L$ and the cone spanned by the rows of S since $(K \cap L) \cdot B = ((K \cap L) + \mathrm{lineal}(K)) \cdot B = K \cdot B = (\mathbb{R}^p_+ \cdot A) \cdot B = \mathbb{R}^p_+ \cdot S$. In particular, we have shown the following result:

Lemma 2.2.13. *A polyhedral cone K is pointed if and only if $\dim(K) = \mathrm{rank}\,(S)$ for any slack matrix S of K.*

Recall that if P is a polytope with representation (V, W, w) and slack matrix $S = [\mathbb{1}, V] \cdot B$ where

$$B = \begin{bmatrix} w^T \\ -W^T \end{bmatrix},$$

then the homogenization P^h of P is a pointed cone that also has S as a slack matrix. Since P^h is pointed, L contains the entire cone and we can restrict the isomorphism in (2.8) to the set $\{1\} \times P = \mathrm{conv}(\mathrm{rows}([\mathbb{1}, V]))$. Thus we have that $\{1\} \times P$ is isomorphic to $\mathrm{conv}(\mathrm{rows}([\mathbb{1}, V])) \cdot B = \mathrm{conv}(\mathrm{rows}(S))$. This establishes the first part of the following:

Theorem 2.2.14. *If S is a slack matrix of the polytope P, then P is isomorphic to $\mathrm{conv}(\mathrm{rows}(S))$. In addition, we have $\dim(P) = \mathrm{rank}\,(S) - 1$.*

Proof. To prove the second statement, note that $\dim(P^h) = \dim(P) + 1$. By Lemma 2.2.13, we have that $\dim(P^h) = \mathrm{rank}\,(S)$. □

In the conic case, we had that $M \in \mathcal{S}_K$ if and only if $M^T \in \mathcal{S}_{K^*}$. This correspondence breaks down for polytopes as we see in the example below. The reason behind this is that we cannot scale $\mathcal{V}$-representations of polytopes by positive scalars.

Example 2.2.15. Let P be the triangular prism with vertices $(0,0,0)$, $(1,0,0)$, $(0,1,0)$, $(0,0,1)$, $(1,0,1)$, and $(0,1,1)$. Then

$$M = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix}$$

is a slack matrix for P. Thus, by Corollary 2.2.7, M satisfies both the RCGC and the CCGC, and the all ones-vector is in the column span of M. However, the all-ones vector is not in the row span of M, so M^T is not the slack matrix of any polytope.

Despite this complication, we can still derive some results for transposes of slack matrices of polytopes. Recall that the *polar* of a polytope $P \subset \mathbb{R}^n$ is

$$P^\circ = \{y \in \mathbb{R}^n \,:\, x^T y \leq 1 \text{ for all } x \in P\}\,.$$

Then P° is a polytope whenever $\mathbf{0} \in \text{int}(P)$, the interior of P. Since translating P does not change its slack matrices, we may assume that $\mathbf{0} \in \text{int}(P)$. Therefore, P has an $\mathcal{H}$-representation of the form $P = \{x \in \mathbb{R}^n \,:\, Wx \leq \mathbb{1}\}$ and $P^\circ = \text{conv}(\text{rows}(W))$. Similarly, if $P = \text{conv}(\text{rows}(V))$, then $P^\circ = \{x \in \mathbb{R}^n \,:\, Vx \leq \mathbb{1}\}$. This implies that the slack matrix of P with respect to the representation $(V, W, \mathbb{1})$ is the transpose of the slack matrix of P° with respect to the representation $(W, V, \mathbb{1})$ and we get the following result that is analogous to Proposition 2.2.2 for cones.

Proposition 2.2.16. *For any polytope P, there exists a slack matrix $M \in \mathcal{S}_P$ such that M^T is also a slack matrix of a polytope.*

In the light of Theorem 2.2.6, this says that slack matrices of polytopes (which already have $\mathbb{1}$ in their column span) allow positive scalings of their columns that puts $\mathbb{1}$ into their row span as well. This is false for general nonnegative matrices.

Example 2.2.17. Continuing Example 2.2.15, we see that the following matrix M' obtained by scaling the columns of M is also a slack matrix of the same prism and does have $\mathbb{1}$ in its row span:

$$M' = \begin{pmatrix} 2 & 2 & 0 & 0 & 0 \\ 2 & 0 & 4 & 0 & 0 \\ 2 & 0 & 0 & 4 & 0 \\ 0 & 2 & 0 & 0 & 2 \\ 0 & 0 & 4 & 0 & 2 \\ 0 & 0 & 0 & 4 & 2 \end{pmatrix}.$$

The prism has vertices:

$$(0,1,-1),(2,-1,-1),(-2,-1,-1),(0,1,1),(2,-1,1),(-2,-1,1)$$

and M' comes from the facet description:

$$z \leq 1, -y \leq 1, -x+y \leq 1, x+y \leq 1, -z \leq 1.$$

Therefore, P° has vertices $(0,0,1),(0,-1,0),(-1,1,0),(1,1,0),(0,0,-1)$ and is a bisimplex (i.e., a bipyramid over a simplex) with slack matrix M'^T.

We can also show a converse to Proposition 2.2.16.

Proposition 2.2.18. *Suppose $M \in \mathbb{R}_+^{p\times q}$ such that M and M^T are both slack matrices of polytopes. Then there exists a polytope P, with $0 \in \operatorname{int}(P)$, such that $M \in \mathcal{S}_P$ and $M^T \in \mathcal{S}_{P^\circ}$.*

Proof. Since M^T is a slack matrix of a polytope, we have that $\mathbb{1} \in \mathbb{R}_+^p \cdot M$. Without loss of generality, we can scale M by a positive scalar so that $\mathbb{1} \in \operatorname{conv}(\operatorname{rows}(M))$.

Let M be a slack matrix of a polytope R with $\dim(R) = d$. By Theorem 2.2.14, $\operatorname{rank}(M) = d+1$. Since the convex hull of the rows of M is isomorphic to R, we have that the affine hull of the rows of M has dimension d. Let J denote the all-ones matrix of dimension $p \times q$. Since $\mathbb{1}$ is contained in the affine hull of the rows of M, we have that the affine hull of the rows of $M - J$ passes through the origin and has dimension d. Hence, rank $(M - J) = d$. This implies that we can write $M - J = AB$ with $A \in \mathbb{R}^{p \times d}$ and $B \in \mathbb{R}^{d \times q}$.

Let $A' = (\mathbb{1}, A)$ and let $B' = \left(\mathbb{1}, B^T\right)^T$. Then $M = A'B'$ is a rank factorization of M. Let $P := \text{conv}(\text{rows}(A))$ and $Q := \{x \in \mathbb{R}^d : \mathbb{1} + x^T B \geq \mathbf{0}\}$. Then the rows of A' form a $\mathcal{V}$-representation of P^h and the columns of B' form a $\mathcal{H}$-representation for $Q^h = \{(x_0, x) \in \mathbb{R}^{d+1} : \mathbb{1}x_0 + x^T B \geq \mathbf{0}\}$. By Lemma 2.2.5, $P^h = Q^h$ which implies that $P = Q$. Therefore, M is a slack matrix of P and M^T is a slack matrix of P°. □

2.3 An algorithm to recognize slack matrices

In this section, we discuss the algorithmic problem of deciding whether a given nonnegative matrix has the RCGC (or, equivalently, the CCGC). According to Corollaries 2.2.4 and 2.2.7 this is the crucial step to be performed in order to decide whether a given matrix is a slack matrix of a cone or a polytope.

We start with a promising result:

Theorem 2.3.1. *The problem to decide whether a nonnegative matrix satisfies the RCGC (or the CCGC) is in coNP. In particular, the same holds for checking the property of being a slack matrix (of a cone or of a polytope).*

Proof. If the given matrix $M \in \mathbb{R}_+^{p \times q}$ does not satisfy the RCGC, then this can be certified by exhibiting an extreme ray of $x \in \mathbb{R}^p \cdot M \cap \mathbb{R}_+^q$ and a facet F of $\mathbb{R}_+^p \cdot M$ such that F separates x from $\mathbb{R}_+^p \cdot M$. This certificate can be chosen to have polynomial size in the encoding length of M since the complexity of the $\mathcal{H}$-representation and the $\mathcal{V}$-representation of a polyhedral cone are polynomially related [Sch86, Chapter 10]. □

Next, we are going to describe an algorithm to check the CCGC (equivalently, the RCGC) for a nonnegative matrix. By Corollary 2.2.4, this algorithm will then provide a method to check if a given nonnegative matrix is a slack matrix of a cone. To check if the matrix is the slack matrix of a polytope (see Corollary 2.2.7), we can add the additional step of checking if the all-ones vector is in the column span of the matrix which is doable in polynomial time. A SAGE worksheet implementing this code can be found at `http://www.math.washington.edu/~rzr`.

Algorithm to check if a nonnegative matrix has the CCGC

Input: A matrix $M \in \mathbb{R}_+^{p\times q}$.

Output: `True` if M has the CCGC and `False` otherwise.

1. Compute a basis L for the left kernel of M. For each vector ℓ in L, generate the equation $\ell^T x = 0$.

2. Generate an $\mathcal{H}$-representation of the cone K with the equations from the previous step and the inequalities $x_1 \geq 0, \ldots, x_p \geq 0$.

3. Compute a minimal $\mathcal{V}$-representation of K.

4. Normalize the vectors in the $\mathcal{V}$-representation and the columns of M.

5. Check that each normalized vector in the $\mathcal{V}$-representation is a normalized column of M. If so, return `True`. If not, return `False`.

Proof. We have $K = M \cdot \mathbb{R}^q \cap \mathbb{R}_+^p$ and $M \cdot \mathbb{R}_+^q \subseteq K$ due to the nonnegativity of M. Thus, M satisfies the CCGC if and only if $K \subseteq M \cdot \mathbb{R}_+^q$ holds, which is what the algorithm checks in the last three steps (note that all cones involved are pointed because they are contained in $\mathbb{R}_+^p$). □

The only computationally challenging part of the algorithm is converting from the $\mathcal{H}$-representation of K to a $\mathcal{V}$-representation. There are several algorithms to do this, and we refer to [Jos04], [MR80], and [Sei97] for information on the different techniques. No polynomial time algorithm for this conversion exists, since the $\mathcal{V}$-representation may have size exponential in that of the $\mathcal{H}$-representation. If the dimension of the cone is fixed, however, then there do exist polynomial time algorithms for the conversion [Dye83]. Thus, we obtain the following complexity results.

Theorem 2.3.2. *For fixed r, checking whether a rank r matrix satisfies the RCGC (CCGC) can be done in polynomial time. In particular, checking whether matrices of fixed rank are slack matrices of cones or polytopes can be done in polynomial time.*

Given an $\mathcal{H}$-polyhedron P and a $\mathcal{V}$-polytope Q contained in P, the problem of deciding whether $P = Q$ is known as the *polyhedral verification problem.* The complexity of this problem is unknown [Sey94]. However, a polynomial time algorithm for the polyhedral verification problem would yield an *output sensitive* algorithm for the problem of computing the facets of a polytope given in $\mathcal{V}$-representation, and thus solve a decades old open problem in computational geometry (see [JZ04]).

Clearly, given a $\mathcal{V}$-polytope it is easy to check whether it is contained in an $\mathcal{H}$-polyhedron. The reverse problem of checking whether an $\mathcal{H}$-polyhedron is contained in a $\mathcal{V}$-polytope is known to be coNP-complete [FO85]. Note that the polyhedral verification problem is the restriction of the latter problem to those instances in which the $\mathcal{V}$-polytope is contained in the $\mathcal{H}$-polyhedron (see also [KP03]).

Theorem 2.3.3. *The following problems can be reduced in polynomial time to each other:*

1. *The polyhedral verification problem*
2. *Is a given matrix a slack matrix of a polytope?*
3. *Is a given matrix a slack matrix of a cone?*
4. *Does a given matrix satisfy the RCGC/CCGC?*

Proof. Corollary 2.2.7 shows that (2) can be reduced (in polynomial time) to (4) (since checking whether $\mathbb{1}$ is contained in the column space can be done in polynomial time) and Corollary 2.2.4 shows that (4) can be reduced to (3).

We can also reduce (3) to (2): Suppose we need to check whether a given matrix M is a slack matrix of a cone. By Lemma 2.2.11, we can assume that M has no zero rows. We can also scale the rows of M by positive scalars without effect on M being a slack matrix of a cone. Using these two facts, we can assume that $\mathbb{1}$ is in the column span of M. Then, being a slack matrix of a cone is equivalent to being a slack matrix of a polytope due to Theorem 2.2.6.

Since Corollary 2.2.8 shows how to reduce (2) to (1), it thus remains to establish a reduction of (1) to (2). Let $Q = \text{conv}(\text{rows}(V))$ with $V \in \mathbb{R}^{p \times n}$ and $P = \{x \in \mathbb{R}^n :$

$Wx \le w\}$ with $W \in \mathbb{R}^{q\times n}$ and $w \in \mathbb{R}^q$ with $Q \subseteq P$. Suppose we need to decide whether $P = Q$. First, we check whether P is pointed (i.e., W has a trivial right kernel) and $\dim(P) = \dim(Q)$ (both checks can be done in polynomial time, the second one using linear programming). If either check fails, then $P \ne Q$.

So let us assume $\dim(P) = \dim(Q)$ and that P is pointed. The latter fact implies that the affine map $\varphi : \mathbb{R}^n \to \mathbb{R}^q$ defined via $\varphi(x) = w - Wx$ is injective. Let M be the matrix arising from V by applying φ to each row. Then, due to $Q \subseteq P$, we have that M is nonnegative. According to Corollary 2.2.8, the matrix M is a slack matrix of a polytope if and only if

$$\operatorname{conv}(\operatorname{rows}(M)) = \operatorname{aff}(\operatorname{rows}(M)) \cap \mathbb{R}^q_+. \tag{2.9}$$

Since we have

$$\operatorname{conv}(\operatorname{rows}(M)) = \varphi(\operatorname{conv}(\operatorname{rows}(V))) = \varphi(Q)$$

and

$$\begin{aligned}\operatorname{aff}(\operatorname{rows}(M)) \cap \mathbb{R}^q_+ &= \varphi(\operatorname{aff}(\operatorname{rows}(V))) \cap \mathbb{R}^q_+ = \varphi(\operatorname{aff}(Q)) \cap \mathbb{R}^q_+ \\ &= \varphi(\{x \in \operatorname{aff}(Q) \,:\, \varphi(x) \ge \mathbf{0}\}) = \varphi(P \cap \operatorname{aff}(Q)) = \varphi(P)\end{aligned}$$

(here we used that $\dim(P) = \dim(Q)$), condition (2.9) is equivalent to $\varphi(P) = \varphi(Q)$. In turn, this is equivalent to $P = Q$ since φ is injective. Thus, $P = Q$ is equivalent to M being the slack matrix of a polytope. □

2.4 A combinatorial characterization of slack matrices

Our second characterization of slack matrices of cones and polytopes relies on incidence structures. For a (nonnegative) matrix M, we denote by M_{inc} the 0/1-matrix with $(M_{\text{inc}})_{ij} = 1$ if and only if $M_{ij} = 0$. Hence, M_{inc} will have a complementary zero pattern to M (we chose this convention so that M_{inc} would correspond to the usual definition of an incidence matrix). The matrices M_{inc} arising from slack matrices M of a polyhedral cone K or of a polytope P are called the *incidence matrices* of K or P, respectively.

We start by characterizing the slack matrices of polytopes, since the corresponding statement for cones can easily be deduced from the one for polytopes. The characterization is restricted to nonnegative matrices of rank at least two. It is easy to see that no matrix of rank one is a slack matrix of a nontrivial polytope. One may (or may not) want to consider a rank-zero matrix as a slack matrix of the polytope consisting of the zero-vector in $\mathbb{R}^0$.

Theorem 2.4.1. *A nonnegative matrix M with $\operatorname{rank}(M) \geq 2$ is a slack matrix of some polytope if and only if M_{inc} is an incidence matrix of some $(\operatorname{rank}(M) - 1)$-dimensional polytope and $\mathbb{1}$ is contained in the column span of M.*

Proof. If M is a slack matrix of a polytope P, then $\mathbb{1}$ is contained in the column span of M (Theorem 2.2.6), and by Theorem 2.2.14, $\dim(P) = \operatorname{rank}(M) - 1$.

In order to establish the non-trivial implication of the claim, let $M \in \mathbb{R}_+^{p\times q}$ be a nonnegative matrix with $\operatorname{rank}(M) = d + 1 \geq 2$, $\mathbb{1} \in M \cdot \mathbb{R}^q$ and M_{inc} an incidence matrix of some d-dimensional polytope R. Denote by $V \subseteq \mathbb{R}_+^q$ the set of rows of M and define the polytope $P := \operatorname{conv}(V)$ and the polyhedron $Q := \operatorname{aff}(V) \cap \mathbb{R}_+^q$. Clearly, $P \subseteq Q$, and since $\mathbb{1} \in M \cdot \mathbb{R}^q$, $\dim(Q) = \dim(P) = d$. By Corollary 2.2.8, in order to show that M is a slack matrix of a polytope, it suffices to prove $P = Q$.

In order to establish $Q \subseteq P$, let us define

$$V_i = \{v \in V : v_i = 0\} \quad \text{and} \quad F_i = \operatorname{conv}(V_i) \quad \text{for } 1 \leq i \leq q\,.$$

The set

$$F = \bigcup_{i=1}^{q} F_i$$

is contained in the relative boundary ∂Q of Q. Note that as an incidence matrix of some polytope of dimension at least one, M_{inc} does not have an all-ones column. Since $Q = \operatorname{conv}(\partial Q)$ (note that Q is a *pointed* polyhedron of dimension $d \geq 2$, which is important here in case of Q being unbounded), if we show that $F = \partial Q$, then we will have that $Q = \operatorname{conv}(F) \subseteq P$.

Thus, our goal is to establish $F = \partial Q$. As mentioned above, we have $F \subseteq \partial Q$. It suffices to show that F is homotopy-equivalent to a $(d-1)$-dimensional sphere[1], because then F cannot be *properly* contained in the $(d-1)$-dimensional connected (recall $\dim(Q) \geq 2$)

[1] Our proof of this is inspired by [JZ04].

manifold ∂Q. This follows from the fact that a closed proper subset of a $(d-1)$-dimensional connected manifold must have a trivial $(d-1)$-st cohomology group and the fact that the $(d-1)$-st cohomology group of a $(d-1)$-dimensional sphere is non-trivial [Bre93].

To show that F is homotopy-equivalent to a $(d-1)$-dimensional sphere, observe that for every subset $I \subseteq \{1, \dots, q\}$, we have $\cap_{i \in I} F_i \neq \varnothing$ if and only if the submatrix of M_{inc} formed by the columns indexed by I has an all-ones row. Now let R be a polytope of which M_{inc} is an incidence matrix. Let $G_1, \dots, G_q$ be the faces of R that correspond to the columns of M_{inc}. Then $\cap_{i \in I} G_i \neq \varnothing$ holds if and only if the submatrix of M_{inc} formed by the columns indexed by I has an all-ones row.

Therefore, the abstract simplicial complexes

$$\{I \subseteq \{1, \dots, q\} : \bigcap_{i \in I} F_i \neq \varnothing\}, \quad \text{and} \quad \{I \subseteq \{1, \dots, q\} : \bigcap_{i \in I} G_i \neq \varnothing\}$$

(known as the *nerves* of the polyhedral complexes induced by $F_1, \dots, F_q$ and by $G_1, \dots, G_q$, respectively) are identical. Since all intersections $\bigcap_{i \in I} F_i$ and $\bigcap_{i \in I} G_i$ are contractible (in fact, they are even convex), this simplicial complex is homotopy equivalent to both F and to the $(d-1)$-dimensional (polyhedral) sphere ∂R (this result is known as the *Nerve Theorem*, see [Bjö95, Thm. 10.6]). □

Since polygons have a very simple combinatorial structure, Theorem 2.4.1 readily yields a simple characterization of their slack-matrices. Here, a *vertex-facet slack matrix* of a polytope P is a slack matrix of P whose rows and columns are in one-to-one correspondence with the vertices and facets of P, respectively.

Corollary 2.4.2. *A matrix $M \in \mathbb{R}_{+}^{n \times n}$ ($n \geq 3$) is a vertex-facet slack matrix of an n-gon if and only if its rows span an affine space of dimension exactly two and its rows and columns can be permuted such that the non-zero entries appear exactly at the positions (i, i) (for $1 \leq i \leq n$), and $(i, i-1)$ (for $2 \leq i \leq n$), and $(1, n)$.*

Steinitz' theorem [SR34] says that a graph G is the 1-skeleton of a three-dimensional polytope if and only if G is planar and three-connected. Using this, one can check in polynomial time whether a given 0/1-matrix is an incidence matrix of a three-dimensional polytope. For every fixed $d \geq 4$, however, it is NP-hard to decide whether a given 0/1-matrix is an incidence matrix of a d-dimensional polytope [RG96].

In the following combinatorial characterization of slack matrices of cones we restrict our attention to matrices of rank at least two as for polytopes. Clearly, every nonnegative matrix of rank one is a slack matrix of the ray $\mathbb{R}^1_+$, and, we may consider a matrix of rank zero as a slack matrix of the trivial cone $\{0\}$ in $\mathbb{R}^0$.

Theorem 2.4.3. *A nonnegative matrix M with* $\operatorname{rank}(M) \geq 2$ *is a slack matrix of a polyhedral cone if and only if M_{inc} is an incidence matrix of some* $\operatorname{rank}(M)$*-dimensional pointed polyhedral cone.*

Proof. If M is a slack matrix of some polyhedral cone then, by Lemma 2.2.10, M is a slack matrix (and hence M_{inc} is an incidence matrix) of a pointed polyhedral cone K. By Lemma 2.2.13 this cone has dimension $\operatorname{rank}(M)$.

In order to prove the reverse implication, we can assume by the results in Section 2.3 that M does not have any zero-row. Since M is also nonnegative, there exists a positive diagonal matrix D such that DM contains $\mathbb{1}$ in its column span.

Given a pointed cone K, we can slice K by an affine hyperplane L such that the slice is a polytope of dimension $\dim(K) - 1$ and the incidence structures of K and $K \cap L$ are identical. Thus, $(DM)_{\text{inc}}$ is an incidence matrix of some $(\operatorname{rank}(M) - 1)$-dimensional polytope. By Theorem 2.4.1, we have that DM is a slack matrix of a polytope. Hence, M is a slack matrix of the homogenization cone of this polytope. $\square$

Note that dropping *pointed* from the formulation of Theorem 2.4.3 makes the statement false. Indeed,

$$M = \begin{bmatrix} 1 & 2 \\ 2 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \quad \text{with} \quad M_{\text{inc}} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 1 \\ 1 & 1 \end{bmatrix}$$

and $\operatorname{rank}(M) = 2$ is not a slack matrix (since M does not satisfy the RCGC), but M_{inc} is the incidence matrix of the non-pointed cone $\{(x_1, x_2) : x_2 \geq 0\}$ with $\mathcal{V}$-representation

$(0,1)$, $(0,1)$, $(1,0)$, $(-1,0)$ and $\mathcal{H}$-representation $(0,1)$, $(0,1)$.

Chapter 3

POLYTOPES OF MINIMUM POSITIVE SEMIDEFINITE RANK [GRT13a]

3.1 Introduction

Efficient representations of polytopes are of fundamental importance in contexts such as linear optimization where the complexity of many algorithms depends on the size of the representation. A standard idea to find a compact description of a complicated polytope $P \subset \mathbb{R}^n$ is to look for a simpler convex set of higher dimension that has P as a linear image of it. Affine slices of closed convex cones offer a rich source of convex sets and the following definition was introduced in [GPT13].

Definition 3.1.1. Let $P \subset \mathbb{R}^n$ be a polytope. If $K \subset \mathbb{R}^m$ is a closed convex cone, L an affine space in $\mathbb{R}^m$, and $\pi : \mathbb{R}^m \to \mathbb{R}^n$ a linear map such that $P = \pi(K \cap L)$, then we say that $K \cap L$ is a K-*lift* of P.

If linear optimization over affine slices of K admits efficient algorithms, then often, linear optimization over P can be done rapidly as well. Well studied cones in this context are nonnegative orthants and the cones of real symmetric positive semidefinite (psd) matrices. We will denote the m-dimensional nonnegative orthant by $\mathbb{R}^m_+$ and the cone of $m \times m$ psd matrices by $\mathcal{S}^m_+$. Affine slices of $\mathbb{R}^m_+$ are polyhedra over which linear optimization can be done efficiently via *linear programming*. Affine slices of $\mathcal{S}^m_+$ are called *spectrahedra*, and linear optimization over them can be done efficiently via *semidefinite programming*. Recall that $\mathbb{R}^m_+$ embeds into $\mathcal{S}^m_+$ via diagonal matrices and hence, polyhedra are special cases of spectrahedra, and semidefinite programming generalizes linear programming.

There are many families of polytopes in $\mathbb{R}^n$ with exponentially many facets (in n) that admit small (polynomial in n) polyhedral or spectrahedral lifts. Examples are the *parity* and *spanning tree polytopes* [Yan91], the *permutahedron* [Goe14] and the *stable set polytope* of a *perfect graph* [LS91]. When the lifts come from families of cones such as $\{\mathbb{R}^m_+\}$ or $\{\mathcal{S}^m_+\}$, it is useful to determine the smallest cone in the family that admits a lift of the polytope. This allows the notion of *cone rank* of a polytope with respect to a family of cones [GPT13]. We recall the definitions needed in this chapter.

Definition 3.1.2. [GPT13]

1. The *nonnegative rank* of a polytope $P \subset \mathbb{R}^n$, denoted as $\operatorname{rank}_+ P$, is the smallest k such that P has an $\mathbb{R}^k_+$-lift.

2. The *positive semidefinite rank* of a polytope $P \subset \mathbb{R}^n$, denoted as $\operatorname{rank}_{\text{psd}} P$, is the smallest k such that P has an $\mathcal{S}^k_+$-lift.

To describe our results, we recall the slack matrix definition from Section 1.2.

Definition 3.1.3. Let P be a full-dimensional polytope in $\mathbb{R}^n$ with vertex set $\{p_1, \dots, p_v\}$ and an irredundant (facet) inequality representation

$$P = \{x \in \mathbb{R}^n : \beta_1 - \langle a_1, x\rangle \geq 0, \dots, \beta_f - \langle a_f, x\rangle \geq 0\}$$

where $\beta_j \in \mathbb{R}$ and $a_j \in \mathbb{R}^n$. Then the nonnegative matrix in $\mathbb{R}^{v\times f}$ whose (i,j)-entry is $\beta_j - \langle a_j, p_i\rangle$ is called a *slack matrix of* P.

Recall that the *polar dual* of a cone $K \subset \mathbb{R}^m$ is the cone

$$K^* := \{y \in \mathbb{R}^m \,:\, \langle x, y\rangle \geq 0 \;\; \forall \;\; x \in K\}.$$

In the vector space of $m \times m$ symmetric matrices we use the trace inner product $\langle A, B\rangle = \operatorname{Tr}(AB)$. Both $\mathcal{S}^k_+$ and $\mathbb{R}^k_+$ are *self dual* cones, meaning that $K^* = K$, and we will identify them with their polar duals in what follows. The notion of *cone factorizations* of slack matrices plays a central role in the theory of cone lifts of polytopes.

Definition 3.1.4. [GPT13] Let $M = (M_{ij}) \in \mathbb{R}^{p\times q}_+$ be a nonnegative matrix and K a closed convex cone whose polar dual is K^*.

- A *K-factorization* of M is a pair of ordered sets $a^1, \dots, a^p \in K$ and $b^1, \dots, b^q \in K^*$ (called *factors*) such that $\langle a^i, b^j\rangle = M_{ij}$.

- When $K = \mathbb{R}^m_+$ (respectively, $\mathcal{S}^m_+$), a K-factorization of M is called a *nonnegative* (respectively, *psd*) *factorization* of M.

- The smallest k for which M has an $\mathbb{R}^k_+$-factorization (respectively, $\mathcal{S}^k_+$-factorization) is called the *nonnegative rank* (respectively, *psd rank*) of M. We denote these invariants of M as $\operatorname{rank}_+ M$ and $\operatorname{rank}_{\text{psd}} M$.

As we discussed in Chapter 2, any positive scaling of a facet inequality of a polytope P can be used in Definition 3.1.3 and so the slack matrix of P is only defined up to positive scalings of its columns. We denote any such slack matrix of P by S_P. Since scaling rows or columns of a matrix M by arbitrary positive real numbers does not affect the existence of a K-factorization of M, all slack matrices of P will have the same behavior with respect to K-factorizations and, in particular, have the same nonnegative (respectively, psd) rank.

In what follows, $P \subset \mathbb{R}^n$ is always an n-dimensional polytope. Yannakakis showed in [Yan91] that $\operatorname{rank}_+ P = \operatorname{rank}_+ S_P$ by proving that P has an $\mathbb{R}^k_+$-lift if and only if S_P has an $\mathbb{R}^k_+$-factorization. The nonnegative rank of a polytope has been the subject of many recent papers [FKPT13, FMP+12, FRT12, GG12, KP11]. The psd rank of a *convex set* $C \subset \mathbb{R}^n$ was introduced in [GPT13] where Yannakakis' theorem was generalized (Theorem 2.4 [GPT13]). Specializing to polytopes, this theorem says that P has a K-lift (in particular, $\mathcal{S}^k_+$-lift) if and only if S_P has a K-factorization ($\mathcal{S}^k_+$-factorization), and so, $\operatorname{rank}_{\text{psd}} P = \operatorname{rank}_{\text{psd}} S_P$. (The extension of Yannakakis' theorem in the case of polytopes also appeared in [FMP+12].) Since $\mathbb{R}^k_+$ embeds into $\mathcal{S}^k_+$ for each k, we always have $\operatorname{rank}_{\text{psd}} P \leq \operatorname{rank}_+ P$. It is easy to see that $\operatorname{rank}_+ P \geq \operatorname{rank} S_P = n+1$. In Proposition 3.3.2 we show that $\operatorname{rank}_{\text{psd}} P$ is also at least $n+1$. This is not immediate since for a general nonnegative matrix M, $\operatorname{rank} M$ is not a lower bound for $\operatorname{rank}_{\text{psd}} M$, and the correct relationship is that $\frac{1}{2}(\sqrt{1 + 8\operatorname{rank} M} - 1) \leq \operatorname{rank}_{\text{psd}} M$ [GPT13]. Theorem 3.3.5 characterizes those n-polytopes whose psd rank equals $n+1$, and we give several families of n-dimensional polytopes whose psd rank equals this lower bound.

We now recall a few useful facts about nonnegative and psd ranks of polytopes that will be needed in this chapter. It follows from [GPT13, Prop. 2] that $\operatorname{rank}_+ P$ and $\operatorname{rank}_{\text{psd}} P$ are invariant under projective (and hence also, affine) transformations of P. Further, transposing a matrix M does not effect the existence of a K-factorization of M if K is self-dual.

Therefore, if P contains the origin in its interior, its *polar* polytope is $P^\circ := \{y \in \mathbb{R}^n : \langle x, y \rangle \leq 1 \; \forall \; x \in P\}$, and $\text{rank}_+ P = \text{rank}_+ P^\circ$ and $\text{rank}_{\text{psd}} P = \text{rank}_{\text{psd}} P^\circ$ since we can obtain a slack matrix of P° by transposing a slack matrix of P and rescaling rows. It is common to define the slack matrix of a polytope using any inequality description of the polytope, including redundant inequalities. This will not affect the nonnegative or psd rank of the polytope. However, since some of our results will become more cumbersome to state using this more general definition of a slack matrix, we restrict ourselves to Definition 3.1.3.

The psd rank of a polytope P quantifies the power of semidefinite programming to provide efficient algorithms for linear optimization over P. For example, the stable set polytope of a perfect graph on n vertices is known to have psd rank $n + 1$ which provides the only known polynomial time algorithm (via semidefinite programming) for finding the highest weight stable set in a perfect graph. The connection between psd rank and semidefinite lifts allows psd rank to become a possible tool for settling questions concerning semidefinite programming in combinatorial optimization. A question that is currently active is whether the psd rank of the *perfect matching polytope* of a complete graph K_n is polynomial in n. Another active question concerns the possible gap between $\text{rank}_+ P$ and $\text{rank}_{\text{psd}} P$ which is a measure of the relative strength of linear vs. semidefinite programming for linear optimization over P. No example where this gap is large is known so far. While nonnegative rank has been studied in several papers, the notion of psd rank is new. The results and techniques presented here further our understanding of psd rank of a polytope.

This chapter is organized as follows. In Section 3.2 we introduce tools to study the psd rank of a general nonnegative matrix M using Hadamard square roots of M. In Section 3.3, we specialize to slack matrices of polytopes and derive the lower bound of $n + 1$ for the psd rank of a n-dimensional polytope (Proposition 3.3.2). Theorem 3.3.5 characterizes n-dimensional polytopes with psd rank $n + 1$ in terms of the lowest rank of a Hadamard square root of a slack matrix of P. In Section 3.4 we give several families of polytopes whose psd rank equals this lower bound. In the plane, the full-dimensional polytopes with psd rank three are exactly triangles and quadrilaterals (Theorem 3.4.7). Every polytope in $\mathbb{R}^n$ with at most $n + 2$ vertices has psd rank $n + 1$ (Theorem 3.4.3). In $\mathbb{R}^3$, the situation gets more tricky and we exhibit polytopes of a fixed combinatorial type (octahedra) whose psd rank depends on the embedding of the polytope. Nonetheless, we show that the three dimensional polytopes with psd rank four are exactly tetrahedra, quadrilateral pyramids, bisimplicies, combinatorial triangular prisms, "biplanar" octahedra, and "biplanar" cuboids (Theorem 3.4.11). It follows from [GPT10] that if S_P is a $0/1$ matrix then $\text{rank}_{\text{psd}}\, P = n+1$. Such polytopes are called 2-level polytopes and include the stable set polytopes of perfect graphs. We exhibit polytopes that are not combinatorially equivalent to 2-level polytopes whose psd rank achieves the lower bound. We also show polytopes that are combinatorially equivalent to 2-level polytopes whose psd rank is not the minimum possible. Finally, we prove in Theorem 3.4.12 that for stable set polytopes, the results of Lovász prevail even in our general setting in the sense that the stable set polytope of a graph on n vertices has psd rank $n+1$ if and only if the graph is perfect.

3.2 Hadamard square roots and psd ranks of matrices

Definition 3.2.1. A *Hadamard square root* of a nonnegative real matrix M, denoted as $\sqrt{M}$, is any matrix whose (i,j)-entry is a square root (positive or negative) of the (i,j)-entry of M. Additionally, we let $\sqrt[+]{M}$ denote the all-nonnegative Hadamard square root of M.

Let $\text{rank}_{\sqrt{\ }}\, M := \min\{\text{rank}\, \sqrt{M}\}$ be the minimum rank of a Hadamard square root of a nonnegative matrix M. We recall the connection between the psd rank of a nonnegative matrix M and $\text{rank}_{\sqrt{\ }}\, M$ shown in [GPT13, Proposition 4.8] , and also in [FMP+12].

Proposition 3.2.2. *If M is a nonnegative matrix, then* $\text{rank}_{\text{psd}}\, M \leq \text{rank}_{\sqrt{\ }}\, M$*. In particular, the psd rank of a $0/1$ matrix is at most the rank of the matrix.*

Proof. Let $\sqrt{M}$ be a Hadamard square root of $M \in \mathbb{R}_+^{p\times q}$ of rank r. Then there exist vectors $a_1, \ldots, a_p, b_1, \ldots, b_q \in \mathbb{R}^r$ such that $(\sqrt{M})_{ij} = \langle a_i, b_j\rangle$. Therefore, $M_{ij} = \langle a_i, b_j\rangle^2 = \langle a_i a_i^T, b_j b_j^T\rangle$ where the second inner product is the trace inner product for symmetric matrices defined earlier. Hence, $\text{rank}_{\text{psd}}\, M \leq r$. □

The upper bound in Proposition 3.2.2 can be strict even for simple examples.

Example 3.2.3. For the matrix

$$M := \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix},$$

$\operatorname{rank} M = \operatorname{rank}_{\sqrt{}} M = 3$ while $\operatorname{rank}_{\text{psd}} M = 2$. Assigning the first three psd matrices below to the rows of M, and the next three to the columns, we obtain a $\mathcal{S}^2_+$-factorization of M:

$$\begin{bmatrix} 0.5 & -0.5 \\ -0.5 & 1 \end{bmatrix}, \begin{bmatrix} 0.5 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \text{ and } \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}.$$

Even though $\operatorname{rank}_{\sqrt{}} M$ is only an upper bound on $\operatorname{rank}_{\text{psd}} M$, we cannot find $\mathcal{S}^k_+$-factorizations of M with only rank one factors if $k < \operatorname{rank}_{\sqrt{}} M$ as shown in Lemma 3.2.4 below. Note that the psd factors corresponding to the first row and the third column of the matrix M in Example 3.2.3 both have rank two.

Lemma 3.2.4. *The smallest k for which a nonnegative real matrix M admits a $\mathcal{S}^k_+$-factorization in which all factors are matrices of rank one is $k = \operatorname{rank}_{\sqrt{}} M$.*

Proof. If $k = \operatorname{rank}_{\sqrt{}} M$, then there is a Hadamard square root of $M \in \mathbb{R}^{p\times q}_+$ of rank k and the proof of Proposition 3.2.2 gives a $\mathcal{S}^k_+$-factorization of M in which all factors have rank one. On the other hand, if there exist $a_1 a_1^T, \ldots, a_p a_p^T, b_1 b_1^T, \ldots, b_q b_q^T \in \mathcal{S}^k_+$ such that $M_{ij} = \langle a_i a_i^T, b_j b_j^T \rangle = \langle a_i, b_j \rangle^2$, then the matrix with (i,j)-entry $\langle a_i, b_j \rangle$ is a Hadamard square root of M of rank at most k. □

Example 3.2.5. For a 0/1 matrix M, $\operatorname{rank}_{\text{psd}} M \leq \operatorname{rank}_{\sqrt{}} M \leq \operatorname{rank} M$. In Example 3.2.3 we saw that the first inequality may be strict. We now show that the second inequality may also be strict. The following *derangement* matrix

$$\begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

has rank three and psd rank two. An $\mathcal{S}_+^2$-factorization in which all factors have rank one is gotten by assigning

$$\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$$

to the three rows and the three columns, respectively. A Hadamard square root of M of rank two is

$$\begin{bmatrix} 0 & -1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}.$$

We now show a method to increase the psd rank of any matrix by one. This technique will be used later to study the psd rank of a polytope.

Proposition 3.2.6. *Suppose $M \in \mathbb{R}_+^{p\times q}$ and $\operatorname{rank}_{\text{psd}} M = k$. If M is extended to $M' = \begin{pmatrix} M & \mathbf{0} \\ w & \alpha \end{pmatrix}$ where $w \in \mathbb{R}_+^q$, $\alpha > 0$ and $\mathbf{0}$ is a column of zeros, then $\operatorname{rank}_{\text{psd}} M' = k+1$. Further, the factor associated to the last column of M' in any $\mathcal{S}_+^{k+1}$-factorization of M' has rank one.*

Proof. Suppose M' has a $\mathcal{S}_+^k$-factorization with factors $A_1, \ldots, A_p, A \in \mathcal{S}_+^k$ associated to its rows and $B_1, \ldots, B_q, B \in \mathcal{S}_+^k$ associated to its columns. Then $A, B \neq 0$ since $\langle A, B\rangle = \alpha \neq 0$. Let $r = \operatorname{rank}(B) > 0$. Then there exists an orthogonal matrix U such that $U^{-1}BU = \operatorname{diag}(\lambda_1, \ldots, \lambda_r, 0, \ldots, 0) =: D$ where $\lambda_1, \ldots, \lambda_r$ are the nonzero (positive) eigenvalues of B. Let $A_i' := U^{-1}A_iU$ for $i = 1, \ldots, p$. Then

$$\langle D, A_i'\rangle = \operatorname{Tr}(U^{-1}BA_iU) = \operatorname{Tr}(BA_i) = \langle B, A_i\rangle = 0 \ \forall \ i = 1, \ldots, p.$$

Since the diagonal entries of A_i' are nonnegative, $\langle D, A_i'\rangle = 0$ implies that the first r diagonal entries of A_i' are all zero. Therefore, the first r rows and the first r columns of A_i' are all zero since A_i' is psd. Now let $B_j' := U^{-1}B_jU$ for all $j = 1, \ldots, q$. Then for all $i = 1, \ldots, p$ and $j = 1, \ldots, q$,

$$\langle A_i', B_j'\rangle = \operatorname{Tr}(U^{-1}A_iB_jU) = \langle A_i, B_j\rangle = M_{ij}.$$

However, since A'_i has nonzero entries only in its bottom right $(k-r)\times(k-r)$ block, it also follows that $M_{ij} = \langle \tilde{A}_i, \tilde{B}_j\rangle$ where $\tilde{A}_i$ is the bottom right $(k-r)\times(k-r)$-submatrix of A'_i and $\tilde{B}_j$ is the bottom right $(k-r)\times(k-r)$ submatrix of B'_j. Thus, there exists a $\mathcal{S}_+^{k-r}$-factorization of M which is a contradiction to the fact that the psd rank of M is k. Therefore, $\text{rank}_{\text{psd}}\, M' \geq k+1$.

If $A_1,\dots,A_p$, $B_1,\dots,B_q \in \mathcal{S}_+^k$ form a size k psd factorization of M, then we can obtain a psd factorization of M' of size $k+1$ by setting

$$\tilde{A}_i := \begin{bmatrix} A_i & \mathbf{0} \\ \mathbf{0} & 0 \end{bmatrix}, \tilde{B}_j := \begin{bmatrix} B_j & \mathbf{0} \\ \mathbf{0} & w_j \end{bmatrix}, \tilde{A} := \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & 1 \end{bmatrix}, \tilde{B} := \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \alpha \end{bmatrix}.$$

Now consider an $\mathcal{S}_+^{k+1}$-factorization of M' and let B be the matrix associated to the last column of M' in this factorization. If $\text{rank}\,(B) = r$, then by the same argument as above, there exists an $\mathcal{S}_+^{k+1-r}$-factorization of M. Since $\text{rank}_{\text{psd}}\, M = k$, $k+1-r \geq k$ or equivalently, $r \leq 1$. Since $B \neq 0$, it follows that $\text{rank}\,(B) = 1$. ☐

Example 3.2.7. The psd rank of a $n \times n$ diagonal matrix with positive diagonal entries is n. The statement holds for $n = 1$ and the general case follows by induction on n and the first part of Proposition 3.2.6. Each factor in an $\mathcal{S}_+^n$-factorization of such a diagonal matrix must have rank one. This follows by applying the second part of Proposition 3.2.6 to both the diagonal matrix and its transpose.

3.3 Hadamard square roots and psd ranks of polytopes

In this section we derive a lower bound to the psd rank of any polytope. We begin with the following easy fact.

Lemma 3.3.1. *Let $P \subset \mathbb{R}^n$ be an n-dimensional polytope. Then a slack matrix S_P has rank $n+1$.*

Proof. Let the vertices of P be $p_1,\dots,p_v$ and the facet inequalities of P be $\langle a_j, x\rangle \leq \beta_j$ for $j = 1,\dots,f$. Then the corresponding $v \times f$ slack matrix S_P has (i,j)-entry equal to

$\beta_j - \langle a_j, p_i \rangle$, and we may factorize S_P as

$$\begin{pmatrix} 1 & p_1 \\ \vdots & \vdots \\ 1 & p_v \end{pmatrix} \begin{pmatrix} \beta_1 & \cdots & \beta_f \\ -a_1 & \cdots & -a_f \end{pmatrix}.$$

Since P is full-dimensional and bounded, both of the factors have rank $n+1$. ☐

We now obtain a lower bound on the psd rank of a polytope.

Proposition 3.3.2. *If $P \subset \mathbb{R}^n$ is a full-dimensional polytope, then the psd rank of P is at least $n+1$. Furthermore, if $\operatorname{rank}_{\mathrm{psd}} P = n+1$, then every $\mathcal{S}_+^{n+1}$-factorization of the slack matrix of P only uses rank one matrices as factors.*

Proof. The proof is by induction on n. If $n = 1$, then P is a line segment and we may assume that its vertices are p_1, p_2 and facets are f_1, f_2 with $p_1 = f_2$ and $p_2 = f_1$. Hence its slack matrix is a 2×2 diagonal matrix with positive diagonal entries. By the arguments in Example 3.2.7, $\operatorname{rank}_{\mathrm{psd}} S_P = 2$ and any $\mathcal{S}_+^2$-factorization of it uses only rank one matrices.

Assume the first statement in the theorem holds up to dimension $n-1$ and consider a polytope $P \subset \mathbb{R}^n$ of dimension n. Let F be a facet of P with vertices $p_1, \ldots, p_s$, facets $f_1, \ldots, f_t$ and slack matrix S_F. Suppose f_i corresponds to facet F_i of P for $i = 1, \ldots, t$. By induction hypothesis, $\operatorname{rank}_{\mathrm{psd}} F = \operatorname{rank}_{\mathrm{psd}} S_F \geq n$. Let p be a vertex of P not in F and assume that the top left $(s+1) \times (t+1)$ submatrix of S_P is indexed by $p_1, \ldots, p_s, p$ in the rows and $F_1, \ldots, F_t, F$ in the columns. Then this submatrix of S_P, which we will call S'_F, has the form

$$S'_F = \begin{pmatrix} S_F & \mathbf{0} \\ * & \alpha \end{pmatrix}$$

with $\alpha > 0$. By Proposition 3.2.6, the psd rank of S'_F is at least $n+1$ since the psd rank of S_F is at least n. Hence, $\operatorname{rank}_{\mathrm{psd}} P = \operatorname{rank}_{\mathrm{psd}} S_P \geq n+1$.

Suppose there is now a $\mathcal{S}_+^{n+1}$-factorization of S_P and therefore of S'_F. By Proposition 3.2.6 the factor corresponding to the facet F has rank one. Repeating the procedure for all facets F and all submatrices S'_F we get that all factors corresponding to the facets of P in this $\mathcal{S}_+^{n+1}$-factorization of S_P must have rank one. To prove that all factors indexed by the vertices of P also have rank one, recall that the transpose of a slack matrix of P is (up to row scaling) a slack matrix of the polar polytope P°, concluding the proof. $\square$

Remark 3.3.3. The zero pattern in S_P has been used to provide lower bounds for $\operatorname{rank}_+ P$ (see for instance, [Yan91, FKPT13]). We note that the zero pattern of a slack matrix by itself is not enough to improve the lower bound on psd rank given in Proposition 3.3.2. For example, consider the slack matrix S_k of a k-gon in $\mathbb{R}^2$. Then $\operatorname{rank}_{\text{psd}} S_k$ grows to infinity as k goes to infinity as shown in [GPT13]. The Hadamard square S_k^2, however, has the same zero pattern as S_k and $\operatorname{rank}_{\text{psd}} S_k^2 \leq \operatorname{rank} S_k = 3$ by Lemma 3.3.1.

Example 3.3.4. The *Birkhoff polytope* $B(n)$ is the convex hull of all $n \times n$ permutation matrices. It was shown in [FKPT13] that $\operatorname{rank}_+ B(n) = n^2$ when $n \geq 5$. By Proposition 3.3.2 and the fact that $B(n)$ has dimension $n^2 - 2n + 1$ (since its affine hull is determined by the $2n$ equations requiring the rows and columns to sum to one, one of which is redundant), we have that $\operatorname{rank}_{\text{psd}} B(n) \geq n^2 - 2n + 2$.

The *permutahedron* $\Pi(n)$ is the convex hull of the vectors $(\pi(1), \ldots, \pi(n))$ where π is a permutation on n letters. It was shown in [Goe14] that $\operatorname{rank}_+ \Pi(n) = O(n \log n)$. By Proposition 3.3.2, $\operatorname{rank}_{\text{psd}} \Pi(n) \geq n$.

Theorem 3.3.5. *If $P \subset \mathbb{R}^n$ is a full-dimensional polytope, then $\operatorname{rank}_{\text{psd}} P = n+1$ if and only if $\operatorname{rank}_{\sqrt{}} S_P = n+1$.*

Proof. By Proposition 3.2.2, $\operatorname{rank}_{\text{psd}} P \leq \operatorname{rank}_{\sqrt{}} S_P$. Therefore, if $\operatorname{rank}_{\sqrt{}} S_P = n+1$, then by Proposition 3.3.2, the psd rank of P is exactly $n+1$.

Conversely, suppose $\operatorname{rank}_{\text{psd}} P = n+1$. Then there exists a $\mathcal{S}_+^{n+1}$-factorization of S_P which, by Proposition 3.3.2, has all factors of rank one. Thus, by Lemma 3.2.4, we have $\operatorname{rank}_{\sqrt{}} S_P \leq n+1$. Since $\operatorname{rank}_{\sqrt{}}$ is bounded below by $\operatorname{rank}_{\text{psd}}$, we must have $\operatorname{rank}_{\sqrt{}} S_P = n+1$. $\square$

Theorem 3.3.5 says that if a full-dimensional polytope $P \subset \mathbb{R}^n$ has the minimum possible psd rank $n+1$, then there must be a Hadamard square root of S_P of rank $n+1$ that serves as a witness. In the next section we exhibit several classes of n-polytopes whose psd rank is $n+1$. We now give examples in the plane that show that many of the properties we have derived so far for n-polytopes of psd rank $n+1$ fail when psd rank is larger than $n+1$.

Example 3.3.6. Consider the pentagon P in $\mathbb{R}^2$ with vertices

$$(0,0),(1,0),(2,1),(1,2),(0,1),$$

and a regular hexagon H in $\mathbb{R}^2$. Then we have slack matrices:

$$S_P = \begin{bmatrix} 0 & 4 & 12 & 4 & 0 \\ 0 & 0 & 8 & 8 & 2 \\ 2 & 0 & 0 & 8 & 4 \\ 4 & 8 & 0 & 0 & 2 \\ 2 & 8 & 8 & 0 & 0 \end{bmatrix}, \; S_H = \begin{bmatrix} 0 & 2 & 4 & 4 & 2 & 0 \\ 0 & 0 & 2 & 4 & 4 & 2 \\ 2 & 0 & 0 & 2 & 4 & 4 \\ 4 & 2 & 0 & 0 & 2 & 4 \\ 4 & 4 & 2 & 0 & 0 & 2 \\ 2 & 4 & 4 & 2 & 0 & 0 \end{bmatrix}.$$

Theorem 3.4.7 will show that these polytopes have psd rank at least four which is not the minimum possible in the plane. We make the following observations:

(i) $\operatorname{rank}_{\sqrt{\ }} S_P > \operatorname{rank}_{\text{psd}} P$

This pentagon has psd rank at most four due to the $\mathcal{S}^4_+$-factorization given by the following matrices (the first five matrices correspond to the rows and the second five to the columns):

$$\begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 1 & 1 & -1 \\ 0 & 1 & 1 & -1 \\ 0 & -1 & -1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & -1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix},$$

$$\begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & -1 & -1 & 1 \\ -1 & 1 & 1 & -1 \\ -1 & 1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix},$$

$$\begin{bmatrix} 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \\ 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

One can check that $\operatorname{rank}_{\sqrt{}}\, S_P = 5$ in this case via the following algebraic calculation. Create a symbolic matrix with the same zeros as a S_P, say

$$S := \begin{bmatrix} 0 & a & b & c & 0 \\ 0 & 0 & d & e & f \\ g & 0 & 0 & h & i \\ j & k & 0 & 0 & l \\ m & n & o & 0 & 0 \end{bmatrix}.$$

Then there is a Hadamard square root of S_P of rank at most four if and only if there is a solution to the system of polynomial equations

$$\{\det(S) = 0,\ a^2 = 4,\ b^2 = 12,\ c^2 = 4, \ldots, o^2 = 8\}.$$

Using a computer algebra package such as Macaulay2 [GS], we can see that this system of equations has no solutions. Therefore, when the psd rank of a n-polytope is greater than $n+1$, there need not be any Hadamard square root of the slack matrix whose rank equals the psd rank of the polytope.

(ii) $\operatorname{rank}_{\sqrt{}}\, S_H < \operatorname{rank}\, \sqrt[+]{S_H}$

The all-nonnegative Hadamard square root $\sqrt[+]{S_H}$ has rank 5. The following Hadamard square root has rank 4:

$$\begin{bmatrix} 0 & \sqrt{2} & 2 & 2 & \sqrt{2} & 0 \\ 0 & 0 & \sqrt{2} & 2 & 2 & \sqrt{2} \\ \sqrt{2} & 0 & 0 & \sqrt{2} & 2 & 2 \\ -2 & -\sqrt{2} & 0 & 0 & \sqrt{2} & 2 \\ 2 & -2 & -\sqrt{2} & 0 & 0 & \sqrt{2} \\ \sqrt{2} & 2 & -2 & -\sqrt{2} & 0 & 0 \end{bmatrix}.$$

Thus, it is not enough to check the positive Hadamard square root of S_P to get $\operatorname{rank}_{\sqrt{}}\, S_P$.

(iii) Recall that if Q is an n-dimensional polytope and $\operatorname{rank}_{\text{psd}} Q = n+1$, then $\operatorname{rank}_{\text{psd}} Q = \operatorname{rank}_{\sqrt{}}\, Q$ and all $\mathcal{S}_+^{n+1}$-factorizations of S_Q have factors of rank one. However, even if

$\operatorname{rank}_{\text{psd}} Q = \operatorname{rank}_{\sqrt{}} Q$, but $\operatorname{rank}_{\text{psd}} Q > n+1$, then there can be factorizations of S_Q by psd matrices of size $\operatorname{rank}_{\text{psd}} Q$ in which the factors do not all have rank one as in the case of the hexagon H.

From above, $\operatorname{rank}_{\sqrt{}} S_H = 4$. A $\mathcal{S}^4_+$-factorization of S_H is gotten by assigning the following six psd matrices of rank two to the columns:

$$\begin{bmatrix} 1 & -1 & 0 & 1 \\ -1 & 1 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 1 & -1 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & -1 \\ 0 & 1 & 1 & -1 \\ 0 & -1 & -1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

$$\begin{bmatrix} 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 1 \\ 0 & -1 & 1 & -1 \\ 0 & 1 & -1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & -1 & 1 & 0 \\ -1 & 1 & -1 & 0 \\ 1 & -1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

and the following six psd matrices of rank one to the rows:

$$\begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

$$\begin{bmatrix} 1 & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

There is no systematic algorithm to find exact psd factorizations of the type shown above. The factorizations in the above example were obtained via trial and error with a pen and paper. We always tried to choose row factors of rank one and all factors as sparse as possible.

We now give two applications of Propositions 3.2.6 and 3.3.2. The first yields a method to produce polytopes of psd rank k from polytopes of psd rank $k-1$. Given a polytope Q, we define a *pyramid over* Q to be any polytope P such that Q forms a facet of P and there is exactly one vertex of P not on this facet.

Proposition 3.3.7. *If* $P \subset \mathbb{R}^n$ *is an* n*-dimensional pyramid over a* $(n-1)$*-polytope* Q *and* $\operatorname{rank}_{\mathrm{psd}} Q = k$*, then* $\operatorname{rank}_{\mathrm{psd}} P = k+1$*.*

Proof. Let S_Q be the slack matrix of Q. By assumption, $\operatorname{rank}_{\mathrm{psd}} S_Q = k$. We may assume without loss of generality that Q lies in the hyperplane $x_n = 0$ and that the apex v of P has $v_n > 0$. The facets of P that contain v are in bijection with the facets of Q. The only other facet inequality of P is $x_n \geq 0$. A slack matrix of P is

$$\begin{bmatrix} S_Q & \mathbf{0} \\ \mathbf{0} & \alpha \end{bmatrix}$$

where the last row is indexed by v and the last column by $x_n \geq 0$. Therefore, $\alpha > 0$ and by Proposition 3.2.6, the psd rank of S_P is $k+1$. □

The following result will be used in Section 3.4.

Proposition 3.3.8. *If a polytope* P *has a facet of psd rank* k*, then* P *has psd rank at least* $k+1$*. In particular, if* $\operatorname{rank}_{\mathrm{psd}} P = n+1$ *where* $P \subset \mathbb{R}^n$ *is a* n*-polytope, then* $\operatorname{rank}_{\mathrm{psd}} F = i+1$ *for every* i*-dimensional face of* P*.*

Proof. The first fact is an immediate consequence of the proof of Proposition 3.3.2 where we saw that if F is a facet of psd rank k, then Proposition 3.2.6 can be used to construct a submatrix S'_F of the slack matrix S_P that has psd rank at least $k+1$. The second statement then follows from Proposition 3.3.2. □

3.4 Families of polytopes of minimum psd rank

Recall that if P is an n-dimensional polytope in $\mathbb{R}^n$ then $\operatorname{rank}_+ P \geq n+1$. It is straightforward to see that the only n-dimensional polytopes of nonnegative rank $n+1$ are simplices. The psd situation is much richer with many more classes of polytopes achieving the minimum possible psd rank as we show in this section.

Definition 3.4.1. A n-dimensional polytope $P \subset \mathbb{R}^n$ is said to be 2-*level* if it has a slack matrix all of whose entries are zero or one. Geometrically, P is 2-level if and only if for each facet of the polytope, all vertices of P lie on the union of this facet and exactly one other parallel translate of the hyperplane spanning this facet.

It follows from [GPT10] that a 2-level polytope in $\mathbb{R}^n$ admits an $\mathcal{S}_+^{n+1}$-lift which can be constructed explicitly using sums of squares polynomials. In the language of the current chapter, it follows that n-dimensional 2-level polytopes have psd rank $n+1$. We can also see this directly from Theorem 3.3.5.

Corollary 3.4.2. *Let P be an n-dimensional 2-level polytope in $\mathbb{R}^n$. Then the psd rank of P is exactly $n+1$. Further, all the factors in any $\mathcal{S}_+^{n+1}$-factorization of P have rank one.*

Proof. Since a 2-level polytope has a 0/1 slack matrix S_P, rank $\sqrt[+]{S_P} = \operatorname{rank} S_P = n+1$. Therefore, $\operatorname{rank}_{\sqrt{}}\, S_P = n+1$, and by Theorem 3.3.5, the psd rank of a 2-level polytope equals $n+1$. The second statement follows from Proposition 3.3.2. □

Since any n-polytope with $n+1$ vertices is a simplex which is 2-level, its psd rank is $n+1$. In fact, Theorem 3.3.5 implies the following stronger result.

Theorem 3.4.3. *Any full-dimensional polytope in $\mathbb{R}^n$ with $n+2$ vertices has psd rank $n+1$.*

Proof. Suppose P is a polytope with $n+2$ vertices. Then if f is the number of facets of P, we have that S_P is an $(n+2) \times f$ matrix of rank $n+1$. Let S_i denote the ith row of S_P. Since $\operatorname{rank} S_P = n+1$, we have $\sum_{i=1}^{n+2} a_i S_i = (0,\ldots,0)$ for some $a_i \in \mathbb{R}$. Each column of S_P must have at least n zeros, so when we consider the above equation component-wise, all but at most two of the summands must be zero. Thus, for each $j = 1,\ldots,f$, $a_{i_0}\left(S_{i_0}\right)_j + a_{i_1}\left(S_{i_1}\right)_j = 0$ for some $1 \le i_0, i_1 \le n+2$. For each a_i define $b_i := \operatorname{sgn}(a_i)\sqrt{|a_i|}$. Then $b_{i_0}\sqrt{\left(S_{i_0}\right)_j} + b_{i_1}\sqrt{\left(S_{i_1}\right)_j} = 0$. Since this holds for each component, we have $\sum_{i=1}^{n+2} b_i \sqrt{S_i} = (0,\ldots,0)$. Thus, $\sqrt[+]{S_P}$ must have rank $n+1$ and the result follows from Theorem 3.3.5. □

There are $\lfloor n^2/4 \rfloor$ distinct combinatorial types of n-dimensional polytopes with $n+2$ vertices [Grü67]. In the plane, we get that all quadrilaterals have psd rank three. In

$\mathbb{R}^3$, the two combinatorial types of polytopes with five vertices are the pyramid over a quadrilateral and a double simplex (bipyramid over a triangle). A quadrilateral pyramid need not be 2-level but it is combinatorially equivalent to a pyramid over a square which is 2-level. By Theorem 3.4.3, a n-dimensional double simplex (bipyramid over a simplex of dimension $n-1$) has psd rank $n+1$. They are polytopes of minimum psd rank that are not combinatorially equivalent to 2-level polytopes.

Proposition 3.4.4. *There is no 2-level polytope that is combinatorially equivalent to a double simplex except in the plane.*

Proof. Let $P \subset \mathbb{R}^n$ be an n-dimensional double simplex. Then the support of any $(n{+}2)\times 2n$ slack matrix of P where the first and last rows correspond to the vertices acquired when taking the bipyramid over a $(n-1)$-dimensional simplex is

$$M := \left[\begin{array}{ccc|ccc} 0 & \cdots & 0 & 1 & \cdots & 1 \\ \hline & I_n & & & I_n & \\ \hline 1 & \cdots & 1 & 0 & \cdots & 0 \end{array}\right].$$

The rank of M is $n+1$ and hence the left kernel of M has dimension one and is generated by the vector $z := (1, -1, -1, \ldots, -1, -1, 1) \in \mathbb{R}^{n+2}$ with all entries equal to -1 except the first and last. Also, P is combinatorially equivalent to a 2-level polytope if and only if there is a (2-level) polytope with slack matrix M.

Suppose M is the slack matrix of a n-dimensional polytope. Then we should be able to factorize M as in the proof of Lemma 3.3.1 into the form

$$M = \begin{bmatrix} 1 & p_1 \\ \vdots & \vdots \\ 1 & p_{n+2} \end{bmatrix} \begin{bmatrix} \beta_1 & \cdots & \beta_f \\ -a_1 & \cdots & -a_{2n} \end{bmatrix}.$$

Call the two factors V and F. The left kernel of V is non-trivial since V is a $(n{+}2)\times(n{+}1)$ matrix. Let z' be a non-zero element in the left kernel of V. Then since $z'VF = 0$, it must also be that $z'M = 0$. This implies that z' is a scalar multiple of z and hence z is in the left kernel of V. But looking at the first column of V, which is all ones, we see that z can be in the left kernel of V only if $n = 2$. ∎

On the other hand, being combinatorially equivalent to a 2-level polytope does not imply minimal psd rank. The regular octahedron in $\mathbb{R}^3$ is a 2-level polytope but we now show an octahedron whose psd rank is five.

Example 3.4.5. Consider the octahedron with vertices

$$(0,0,0),(2,0,0),(0,2,0),(2,2,0),(1,1,-1),(1,2,1)$$

which has slack matrix:

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 2 & 2 & 2 & 2 \\ 0 & 2 & 0 & 2 & 0 & 0 & 2 & 2 \\ 2 & 0 & 2 & 0 & 2 & 2 & 0 & 0 \\ 2 & 2 & 2 & 2 & 0 & 0 & 0 & 0 \\ 0 & 2 & 3 & 0 & 0 & 2 & 1 & 0 \\ 3 & 0 & 0 & 2 & 2 & 0 & 0 & 1 \end{bmatrix}.$$

It can be checked algebraically as in Example 3.3.6 that no Hadamard square root of this slack matrix has rank four. However, the positive Hadamard square root has rank five and hence the psd rank of this octahedron is five.

Remark 3.4.6. We have seen that having the combinatorial type of a 2-level polytope is not enough for minimal psd rank, while being the image under a projective transformation of a 2-level polytope is enough. Proposition 3.4.4 shows that not all polytopes of minimal psd rank are projectively equivalent to 2-level polytopes. Strictly weaker than being projectively equivalent to a 2-level polytope is the existence of a positive scaling of each row and column of S_P that turns it into a 0/1-matrix. This clearly implies minimal psd rank, and includes double simplices. So one could suppose this to be a necessary and sufficient condition for having minimal psd rank. This turns out to be false. Consider the prism with vertices

$(0,0,0)$, $(1,0,0)$, $(0,1,0)$, $(1,2,0)$, $(0,0,1)$, $(1,0,1)$, $(0,1,1)$, $(1,2,1)$ which has slack matrix

$$\begin{bmatrix} 0 & 0 & 2 & 1 & 0 & 1 \\ 1 & 0 & 0 & 2 & 0 & 1 \\ 0 & 1 & 2 & 0 & 0 & 1 \\ 1 & 2 & 0 & 0 & 0 & 1 \\ 0 & 0 & 2 & 1 & 1 & 0 \\ 1 & 0 & 0 & 2 & 1 & 0 \\ 0 & 1 & 2 & 0 & 1 & 0 \\ 1 & 2 & 0 & 0 & 1 & 0 \end{bmatrix}.$$

The positive square root of this matrix has rank four, so the polytope has minimal psd rank, but it is easy to see that we can never turn the submatrix from the first two rows and the fourth and sixth columns into a 0/1-matrix by any scaling.

In the plane we can fully characterize the polytopes of psd rank three.

Theorem 3.4.7. *A convex polygon P in the plane has psd rank three if and only if it has at most four vertices.*

Proof. The "if" direction was discussed after Theorem 3.4.3.

Now suppose that P is a convex polygon with 5 or more vertices. By an affine transformation we can suppose P has facets given by $x \geq 0$ and $y \geq 0$ with vertices on $(0,0)$, $(1,0)$ and $(0,1)$. Let (a,b) be the vertex sharing an edge with $(0,1)$ and (c,d) the one sharing an edge with $(1,0)$. These facets are then given by the two inequalities $(b-1)x - ay + a \geq 0$ and $(c-1)y - dx + d \geq 0$ respectively, so we can take the 5×4 submatrix of the slack matrix of P indexed by these vertices and facets, which is then

$$S'_P = \begin{bmatrix} 0 & 0 & a & d \\ 0 & 1 & 0 & d+c-1 \\ 1 & 0 & a+b-1 & 0 \\ a & b & 0 & cb-b-da+d \\ c & d & bc-c-ad+a & 0 \end{bmatrix}.$$

It is then enough to show that every possible Hadamard square root of the 4×4 upper left portion of this matrix has rank four. This matrix is given by

$$\begin{bmatrix} 0 & 0 & \pm\sqrt{a} & \pm\sqrt{d} \\ 0 & \pm 1 & 0 & \pm\sqrt{d+c-1} \\ \pm 1 & 0 & \pm\sqrt{a+b-1} & 0 \\ \pm\sqrt{a} & \pm\sqrt{b} & 0 & \pm\sqrt{cb-b-da+d} \end{bmatrix}.$$

Assume this matrix has rank three. Since the first three rows are independent, we can write the fourth row as a combination of the first three. In such a combination, the coefficients for the first three rows must be $\pm\sqrt{a+b-1}$, $\pm\sqrt{b}$ and $\pm\sqrt{a}$, respectively. For ease of notation, let $\alpha = b(d+c-1)$ and $\beta = d(a+b-1)$. Then $\alpha, \beta > 0$ and $\alpha \geq \beta$. Looking at the last column, we see that

$$\pm\sqrt{\alpha - \beta} = \pm\sqrt{\alpha} \pm \sqrt{\beta}.$$

Out of these eight possible equations, the only four that are feasible are $\pm\sqrt{\alpha-\beta} = \sqrt{\alpha} - \sqrt{\beta}$ and $\pm\sqrt{\alpha-\beta} = -\sqrt{\alpha} + \sqrt{\beta}$, all of which imply $\alpha = \beta$. Hence, $cb - b = ad - d$ and we have that $b/(a-1) = d/(c-1)$. Thus, the slope of the line between (a, b) and $(1, 0)$ equals the slope between (c, d) and $(1, 0)$, implying that the three are collinear and cannot all be vertices unless $(a, b) = (c, d)$. □

In $\mathbb{R}^3$, it is more difficult to classify the convex polytopes of minimum psd rank. We have seen that all polytopes with four or five vertices have psd rank four. Additionally, we can say precisely which octahedra in $\mathbb{R}^3$ have psd rank four. Let $O \subset \mathbb{R}^3$ be a (combinatorial) octahedron. We say that O is planar with respect to a plane E if $O \cap E$ contains four vertices of O. For example, the regular octahedron is planar to the xy, xz, and yz planes. A combinatorial octahedron can be planar with respect to at most three planes. We say O is *biplanar* if it is planar with respect to at least two distinct planes.

Theorem 3.4.8. *An octahedron $O \subset \mathbb{R}^3$ has psd rank four if and only if O is biplanar.*

Proof. First, assume O is biplanar. Then, by applying an affine transformation, we can assume that O is planar with respect to the xy plane and has vertices $(0, 0, 0)$, $(1, 0, 0)$, $(0, 1, 0)$, $(a, b, 0)$, (z_1, z_2, z_3), and (w_1, w_2, w_3) where $z_3 > 0$, $w_3 < 0$, and $a + b > 1$.

For ease of notation, let $\alpha = z_3 - w_3$, $\beta = w_1 z_3 - z_1 w_3$, and $\gamma = w_2 z_3 - z_2 w_3$. Then $(0,0,0)$, $(a,b,0)$, (z_1,z_2,z_3), (w_1,w_2,w_3) are coplanar if and only if $b\beta = a\gamma$ and $(1,0,0)$, $(0,1,0)$, (z_1,z_2,z_3), (w_1,w_2,w_3) are coplanar if and only if $\alpha = \beta + \gamma$. The combinatorics of O dictates that these are the only possible further planarities, and since O is biplanar, at least one of these conditions must be satisfied.

Now O has slack matrix S_O:

$$\begin{bmatrix} 0 & 0 & b & a & 0 & 0 & b & a \\ 1 & 0 & 0 & a+b-1 & 1 & 0 & 0 & a+b-1 \\ 0 & 1 & a+b-1 & 0 & 0 & 1 & a+b-1 & 0 \\ a & b & 0 & 0 & a & b & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{-\beta}{w_3} & \frac{-\gamma}{w_3} & \frac{b(\beta-\alpha)+(1-a)\gamma}{w_3} & \frac{a(\gamma-\alpha)+(1-b)\beta}{w_3} \\ \frac{\beta}{z_3} & \frac{\gamma}{z_3} & \frac{b(\alpha-\beta)+(a-1)\gamma}{z_3} & \frac{a(\alpha-\gamma)+(b-1)\beta}{z_3} & 0 & 0 & 0 & 0 \end{bmatrix}.$$

In the case $b\beta = a\gamma$ or the case $\alpha = \beta + \gamma$, row reduction shows that $\sqrt[+]{S_O}$ has rank four. Hence, O has psd rank four.

For the converse, suppose O is planar to either one or zero planes. If a planar condition is satisfied, assume it is by the vertices v_1, v_2, v_3, v_4. By applying an affine transformation, we can assume that $v_1 = (0,0,1)$, $v_2 = (0,0,0)$, $v_3 = (1,0,0)$, and $v_5 = (0,1,0)$. Let $v_4 = (z_1, z_2, z_3)$ and $v_6 = (w_1, w_2, w_3)$ where we must have

$$z_1 < 0,\ w_3 > 0,\ 1 - z_1 - z_2 - z_3 > 0, \text{ and } 1 - w_1 - w_2 - w_3 > 0 \tag{3.1}$$

to preserve the combinatorial structure. (These are not all of the required conditions, but we will use these particular ones below.)

Since O cannot satisfy planarity conditions on the set of vertices $\{v_1, v_2, v_5, v_6\}$ or $\{v_3, v_4, v_5, v_6\}$, we must have that

$$w_1 \neq 0 \text{ and } w_1 z_3 + w_2 z_3 - z_1 w_3 - z_2 w_3 + w_3 - z_3 \neq 0. \tag{3.2}$$

We calculate the slack matrix S_O and consider its 5×5 submatrix M indexed by the vertices v_1, v_2, v_3, v_5, v_6 in the rows and the facets $F_{1,3,5}$, $F_{2,3,6}$, $F_{2,4,5}$, $F_{1,3,6}$, $F_{1,4,5}$ in the columns where $F_{i,j,k}$ is the facet defined by the vertices v_i, v_j, v_k. After multiplying the rows and columns by nonnegative constants, M has the form:

$$\begin{bmatrix} 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & z_3 & 0 & 1-z_1-z_2-z_3 \\ 0 & w_3 & 0 & 1-w_1-w_2-w_3 & 0 \\ -z_1(1-w_1-w_2-w_3) & 0 & -z_1w_3+w_1z_3 & 0 & -z_1(1-w_2-w_3)+w_1(1-z_2-z_3) \end{bmatrix}.$$

Now consider an arbitrary Hadamard square root $\sqrt{M}$. For the purposes of calculating rank of $\sqrt{M}$, we can assume that the $(1,2)$, $(1,3)$, $(2,1)$, $(2,4)$, and $(2,5)$ entries of $\sqrt{M}$ are all 1. Let

$$S = \begin{bmatrix} 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & s_1 & 0 & s_2 \\ 0 & s_3 & 0 & s_4 & 0 \\ s_5 & 0 & s_6 & 0 & s_7 \end{bmatrix}.$$

be a symbolic matrix corresponding to a $\sqrt{M}$ and let $\tilde{z}_1, \ldots, \tilde{w}_3$ be variables corresponding to $z_1, \ldots, w_3$. Consider the ideal I generated by the polynomials:

$$\left\{\det S,\, s_1^2 - \tilde{z}_3, \ldots, s_7^2 + \tilde{z}_1(1-\tilde{w}_2-\tilde{w}_3) - \tilde{w}_1(1-\tilde{z}_2-\tilde{z}_3)\right\}.$$

Now if $\operatorname{rank}_{\text{psd}} O = 4$, then $\operatorname{rank}_{\sqrt{}} M \leq 4$ and, hence, there must exist real numbers $x_1, \ldots, x_7$ such that $(x_1, \ldots, x_7, z_1, \ldots, w_3)$ lies in $V(I)$, the variety of I.

The three possible planarity conditions on O are given by the equations:

$$\tilde{w}_1 = 0,\ \tilde{z}_2 = 0,\ \text{and}\ \tilde{w}_1\tilde{z}_3 + \tilde{w}_2\tilde{z}_3 - \tilde{z}_1\tilde{w}_3 - \tilde{z}_2\tilde{w}_3 + \tilde{w}_3 - \tilde{z}_3 = 0.$$

Let J_1, J_2, J_3 be the ideals generated by two each of the three polynomials defining the above planarity conditions. Then the product ideal $J := J_1 * J_2 * J_3$ has variety $V(J) = V(J_1) \cup V(J_2) \cup V(J_3)$. By our planarity assumption on O, $(x_1, \ldots, x_7, z_1, \ldots, w_3)$ is not contained in $V(J)$. Now $V(I)\backslash(V(J)$ is contained in the variety of the colon ideal $I : J$ [CLO92, Chapter 4.4, Theorem 7] and, hence, $(x_1, \ldots, x_7, z_1, \ldots, w_3)$ vanishes on every polynomial in $I : J$. Using Macaulay2 [GS], we can compute a set of generators of $I : J$ and by elimination one sees that

$$f = \tilde{z}_1\tilde{w}_1\tilde{w}_3(\tilde{w}_1 + \tilde{w}_2 + \tilde{w}_3 - 1)(\tilde{z}_1 + \tilde{z}_2 + \tilde{z}_3 - 1)(\tilde{w}_1\tilde{z}_3 + \tilde{w}_2\tilde{z}_3 - \tilde{z}_1\tilde{w}_3 - \tilde{z}_2\tilde{w}_3 + \tilde{w}_3 - \tilde{z}_3)$$

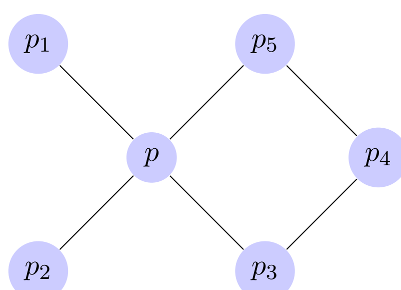
lies in $I : J$. However, no choice of $z_1, \ldots, w_3$ that is required to satisfy (3.1) and (3.2) can vanish on f. Hence, we must have $\operatorname{rank}_{\text{psd}} O \geq 5$. $\square$

A *cuboid*, or combinatorial cube, is a polytope in $\mathbb{R}^3$ that is combinatorially equivalent to a cube. Since the polars of cuboids are octahedra and psd rank is preserved under polarity, the cuboids of minimal psd rank are precisely the polars of biplanar octahedra. We call these biplanar cuboids. More explicitly, these are the cuboids for which there exists two sets of four facets whose supporting hyperplanes intersect in a point (possibly at infinity).

We will now argue that there are no polytopes in $\mathbb{R}^3$ of psd rank four beyond the ones we have considered above (and their polars). Let P be a polytope in $\mathbb{R}^3$ of psd rank four. By Proposition 3.3.8, all the facets of P must be triangles or quadrilaterals. Further, since $\operatorname{rank}_{\text{psd}} P^\circ = 4$, each vertex of P must be of *degree* three or four. Recall that the degree of a vertex of P is the number of edges of P incident to that vertex.

Lemma 3.4.9. *Let $P \subset \mathbb{R}^3$ be a three-dimensional polytope with $\operatorname{rank}_{\text{psd}} P = 4$. If p is a vertex of P of degree four, then the four facets incident to p must be triangles.*

Proof. Let P and p be as above and suppose that the four facets incident to p are not all triangles. By Proposition 3.3.8, one of the facets surrounding p must be a quadrilateral and P contains the following structure (with $p_1, \ldots, p_5$ vertices of P):



Let S_P be a slack matrix of P. Then S_P is of rank four. Further, since P has minimum psd rank, there exists a Hadamard square root $\sqrt{S_P}$ of rank four. Let M be the 5×4 submatrix of $\sqrt{S_P}$ indexed by p, p_1, p_2, p_3, p_4 in the rows and by the four facets incident to p in the columns. By scaling the columns of $\sqrt{S_P}$ by nonzero scalars, we may assume that M is of the following form, with a, b, c, d, e nonzero:

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & a \\ 1 & 0 & b & 0 \\ c & d & e & 0 \end{bmatrix}.$$

The four rows of $\sqrt{S_P}$ and S_P corresponding to the first four rows of M are linearly independent by the structure of M. Hence, we can write the row of $\sqrt{S_P}$ and S_P corresponding to the fifth row of M as a linear combination of the other four. Thus, we can write the fifth row of M and M^2 as a linear combination of the first four. This results in two necessary equations: $d + ae = abc$ and $d^2 + a^2e^2 = (abc)^2$, which implies that $ade = 0$, a contradiction. □

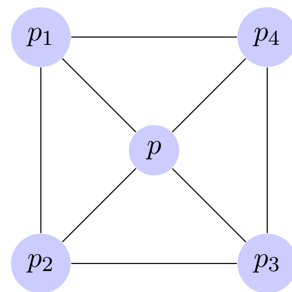
Proposition 3.4.10. *A polytope in $\mathbb{R}^3$ of psd rank four has the combinatorial type of a simplex, quadrilateral pyramid, bisimplex, triangular prism, octahedron, or cube.*

Proof. Let P be a polytope in $\mathbb{R}^3$ of psd rank four with v vertices, e edges, and f facets. Let v_t and v_q denote the number of vertices of degree three and four in P, and let f_t and f_q denote the number of triangular and quadrangular facets of P.

By double counting edges, $2e = 3f_t + 4f_q$, and by considering P°, we also see that $2e = 3v_t + 4v_q$. Now using Euler's formula, $v - e + f = 2$, it is easy to deduce that v_t and f_t are even and that $v_t + f_t = 8$. Hence, we only need to consider polytopes where (v_t, f_t) equals $(0, 8)$, $(2, 6)$, $(4, 4)$, $(6, 2)$, or $(8, 0)$. Further, by taking polars we need only consider the cases where (v_t, f_t) equals $(0, 8)$, $(2, 6)$, or $(4, 4)$.

When $(v_t, f_t) = (0, 8)$, we have that every vertex is of degree four. Thus, by Lemma 3.4.9, every facet must be triangular. The only polytope in $\mathbb{R}^3$ that satisfies these conditions is the octahedron.

Now suppose $(v_t, f_t) = (4, 4)$. If there are no degree four vertices, then there are only four total vertices and the polytope must be the simplex. If there is a degree four vertex, then by Lemma 3.4.9 the polytope must contain the following configuration:



If vertex p_1, p_2, p_3, or p_4 has degree four, then we will be forced to include too many triangular facets. Thus, they all must have degree three and the polytope must be a quadrilateral pyramid.

Finally, suppose $(v_t, f_t) = (2, 6)$. Then P must have a degree four vertex (call it p) and the configuration above is again included in the boundary complex of P with the four triangles shown being facets of P. Since P has only two vertices of degree three, at least two of the vertices surrounding p must have degree four. Suppose two adjacent vertices among p_1, p_2, p_3, p_4 have degree four. Then each of them must be contained in four triangular facets which means that each such vertex is incident to two triangular facets that are not shown in the figure. But since these degree four adjacent vertices already share a facet, they can share at most one of these four extra triangular facets. This creates a total of seven triangular facets in P contradicting $f_t = 6$. Therefore, the two vertices of degree four among p_1, p_2, p_3, p_4 must be nonadjacent. As before, each is adjacent to two triangular facets that are not shown and since $f_t = 6$, it must be that the two vertices share these two triangular facets. Therefore, P is a bisimplex.

Now the facts that the polar of a bisimplex is combinatorially a triangular prism, and the polar of an octahedron is a cube completes the proof. $\square$

We now immediately obtain the following theorem which gives a complete classification of polytopes in $\mathbb{R}^3$ of psd rank four.

Theorem 3.4.11. *The polytopes in $\mathbb{R}^3$ of psd rank four are precisely simplices, quadrilateral pyramids, bisimplicies, combinatorial triangular prisms, biplanar octahedra, and biplanar cuboids.*

A major catalyst for the use of semidefinite programming in combinatorial optimization was the *Lovász theta body of a graph* [Lov79, GLS88], denoted as $\mathrm{TH}(G)$, which is a convex relaxation of the stable set polytope of a graph. Let $G = ([n], E)$ be a graph with vertex set $[n] := \{1, \ldots, n\}$ and edge set E. Recall that a *stable set* of G is a subset $S \subseteq [n]$ such that for all $i, j \in S$, the pair $\{i, j\}$ is not in E. The *characteristic vector* of a stable set S is $\mathcal{X}^S \in \{0,1\}^n$ defined as $(\mathcal{X}^S)_i = 1$ if $i \in S$ and 0 otherwise. The *stable set polytope* of G is the n-dimensional polytope

$$\mathrm{STAB}(G) := \text{convex hull}(\mathcal{X}^S \,:\, S \text{ stable set in } G) \subset \mathbb{R}^n,$$

and $\mathrm{TH}(G)$ is the following projection of an affine slice of $\mathcal{S}_+^{n+1}$:

$$\left\{ x \in \mathbb{R}^n \,:\, \exists \begin{bmatrix} 1 & x^T \\ x & U \end{bmatrix} \succeq 0 \text{ s.t. } U_{ii} = x_i \;\forall\; i = 1, \ldots, n \text{ and } U_{ij} = 0 \;\forall\; \{i, j\} \in E \right\}.$$

Further, $\mathrm{TH}(G) = \mathrm{STAB}(G)$ if and only if G is a *perfect graph* [GLS88, Chapter 9]. Hence if G is perfect, $\mathrm{rank}_{\mathrm{psd}}\, \mathrm{STAB}(G) = n + 1$ and the description of $\mathrm{TH}(G)$ gives a $\mathcal{S}_+^{n+1}$-lift of $\mathrm{STAB}(G)$. In the context of this chapter, it is natural to ask if there are non-perfect graphs for which $\mathrm{rank}_{\mathrm{psd}}\, \mathrm{STAB}(G) = n + 1$, via other $\mathcal{S}_+^{n+1}$-lifts.

Theorem 3.4.12. *Let G be a graph with n vertices. Then* $\mathrm{STAB}(G)$ *has psd rank $n + 1$ if and only if G is perfect.*

Proof. We saw that $\mathrm{rank}_{\mathrm{psd}}\, \mathrm{STAB}(G) = n + 1$ when G is a perfect graph with n vertices. Suppose G is not perfect. By Proposition 3.3.8, it is enough to show that $\mathrm{STAB}(G)$ has a face that is not of minimal psd rank. By the perfect graph theorem [CRST06], G contains a *odd hole* or *odd anti-hole* H. Since $\mathrm{STAB}(H)$ forms a face of $\mathrm{STAB}(G)$, we just need to show that $\mathrm{STAB}(H)$ is not of minimal psd rank.

Let $H = ([2m + 1], E)$ and assume H is an odd hole. The anti-hole case is exactly analogous and is omitted here. Now $\mathrm{STAB}(H)$ is a $(2m + 1)$-dimensional polytope with facet inequalities:

1. $x_i \geq 0$ for each $i \in [2m + 1]$

2. $\mathbf{x}_e \leq 1$ for each $e \in E$

3. $\mathbf{x}_{[2m+1]} \leq m$

where $\mathbf{x}_T := \sum_{i \in T} x_i$ for every subset T of $[2m+1]$ and $\mathbf{x}_e := x_i + x_j$ for $e = \{i, j\} \in E$. Let S be the slack matrix of STAB(H) and let S' be the $(2m+3) \times (2m+3)$ submatrix of S where S' is indexed by the stable sets

$$\{\ \}, \{1\}, \{2\}, \ldots, \{2m+1\}, \{1,3\}$$

in the rows and the facets $\mathbf{x}_{\{1,2\}} \leq 1, x_1 \geq 0, \ldots, x_{2m+1} \geq 0, \mathbf{x}_{[2m+1]} \leq m$ in the columns. Then S' has the form:

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & \cdots & 0 & m \\ 0 & 1 & 0 & 0 & 0 & \cdots & 0 & m-1 \\ 0 & 0 & 1 & 0 & 0 & \cdots & 0 & m-1 \\ 1 & 0 & 0 & 1 & 0 & \cdots & 0 & m-1 \\ 1 & 0 & 0 & 0 & 1 & \cdots & 0 & m-1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 0 & 0 & 0 & 0 & \cdots & 1 & m-1 \\ 0 & 1 & 0 & 1 & 0 & \cdots & 0 & m-2 \end{bmatrix}.$$

Let $\sqrt{S'}$ be an arbitrary Hadamard square root and suppose that rank $\sqrt{S'} \leq 2m+2$. Then since the first $2m+2$ columns are linearly independent, we must have that the final column is a linear combination of the first $2m+2$. Let $\alpha_1, \ldots, \alpha_{2m+2}$ be coefficients in such a combination. By looking at the first, second, fourth, and last columns, we see that $\alpha_1 = \pm\sqrt{m}$, $\alpha_2 = \pm\sqrt{m-1}$, and $\alpha_4 = \pm\sqrt{m} \pm \sqrt{m-1}$. Now by looking at the last row, we must have $\pm\alpha_2 \pm \alpha_4 = \pm\sqrt{m-2}$, which is a contradiction. Hence, $\text{rank}_{\sqrt{}}\, S > 2m+2$ and we have that STAB(H) is not of minimal psd rank. $\square$

Chapter 4

RATIONAL AND REAL POSITIVE SEMIDEFINITE RANK CAN BE DIFFERENT [FGR14]

In this chapter, we answer a basic structural question about the psd rank: if a nonnegative matrix M has only rational entries, can the psd rank of M always be achieved by a factorization using only rational matrices? We answer this question negatively by providing an example of a rational matrix with psd rank four such that every psd factorization of size four uses irrational numbers. Note that the analogous question for the *nonnegative rank* of a matrix was posed by Cohen and Rothblum in [CR93]. It was shown in [CR93] that all rational matrices with nonnegative rank two admit a rational nonnegative factorization, but the question for general nonnegative matrices remains open.

The proof of our example will require a lemma about rational psd matrices of rank one. Any rank one psd matrix has the form $\mathbf{v}\mathbf{v}^T$ for some vector $\mathbf{v}$. Let ϕ denote the map taking the vector $\mathbf{v}$ to the psd matrix $\mathbf{v}\mathbf{v}^T$. Then we have the following.

Lemma 4.0.1. *If the matrix $\phi(\mathbf{v})$ is composed of only rational entries, then $\mathbf{v}$ has the form $\alpha\mathbf{q}$ where α is a real scalar and $\mathbf{q}$ is a rational vector.*

Proof. Suppose that $\mathbf{v}$ is a nonzero vector (else the conclusion is immediate). Then, without loss of generality, we may assume that the first coordinate v_1 is nonzero. Since v_1^2 is an entry in the matrix $\phi(\mathbf{v})$, it must be rational. Hence, the matrix $\frac{1}{v_1^2}\phi(\mathbf{v})$ is also rational. By looking at the first row of this matrix, we see that the vector $\left(1, \frac{v_2}{v_1}, \frac{v_3}{v_1}, \ldots, \frac{v_r}{v_1}\right)$ is rational. Now we just scale this rational vector by v_1 to finish the proof. $\square$

Our candidate matrix M is the 8×6 matrix shown in Figure 4.1. This matrix arises as a slack matrix of the polytope with vertices $(0, 0, 0)$, $(1, 0, 0)$, $(0, 1, 0)$, $(1, 2, 0)$, $(0, 0, 1)$, $(1, 0, 1)$, $(0, 1, 1)$, and $(1, 2, 1)$. We will refrain, however, from using any results about slack matrices in the proofs.

$$M = \begin{pmatrix} 0 & 0 & 2 & 1 & 0 & 1 \\ 1 & 0 & 0 & 2 & 0 & 1 \\ 0 & 1 & 2 & 0 & 0 & 1 \\ 1 & 2 & 0 & 0 & 0 & 1 \\ 0 & 0 & 2 & 1 & 1 & 0 \\ 1 & 0 & 0 & 2 & 1 & 0 \\ 0 & 1 & 2 & 0 & 1 & 0 \\ 1 & 2 & 0 & 0 & 1 & 0 \end{pmatrix}$$

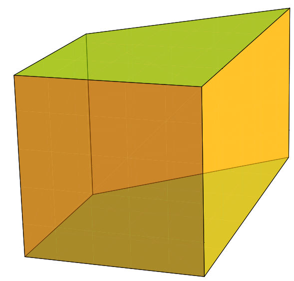


Figure 4.1. Our example matrix is a slack matrix of a three dimensional polytope.

During our analysis of this example, we will require a few known results about psd rank. We summarize these results in the following proposition.

Proposition 4.0.2.

1. *[Proposition 3.2.2], [GPT13, Prop. 5], If $\sqrt{A}$ is an entry-wise square root of A, then* $\operatorname{rank}_{\text{psd}} A \leq \operatorname{rank} \sqrt{A}$.

2. *[Proposition 3.2.6], [LT12, Cor. 4.8], If A contains a $k \times k$ triangular submatrix T, then* $\operatorname{rank}_{\text{psd}} A \geq k$. *Furthermore, in a psd factorization of A of size k, the factor corresponding to the row (or column) of T with $k-1$ zeros must have rank one.*

Now we begin our analysis of the matrix M.

Lemma 4.0.3. *We have that* $\operatorname{rank}_{\text{psd}} M = 4$ *and any psd factorization of M of size four uses only rank one factors.*

Proof. One can verify that the all-nonnegative entry-wise square root of M has usual rank four. Thus Proposition 4.0.2 says that $\operatorname{rank}_{\text{psd}} M \leq 4$. Consider the submatrix of M indexed by rows 1, 5, 7, and 8 and columns 1, 2, 5, and 6. This submatrix is triangular so Proposition 4.0.2 tells us two things: First, we have that $\operatorname{rank}_{\text{psd}} M \geq 4$ and, hence,

$\operatorname{rank}_{\mathrm{psd}} M = 4$. Second, the factors corresponding to the first row and first column in a psd factorization of M of size four must always be rank one. It is easy to verify by inspection that for every row and column of M we can find a 4×4 triangular submatrix such that the row or column in question has three zeros in that submatrix. Thus, repeatedly applying the proposition completes the proof. □

Remark 4.0.4. Note that Lemma 4.0.3 is actually a consequence of Proposition 3.3.2 since our polytope has minimal psd rank (equal to the ambient dimension plus one) and thus any psd factorization must consist entirely of rank-one factors.

The next proposition shows that no rational psd factorization of M can have size four.

Proposition 4.0.5. *We have that* $\operatorname{rank}_{\mathrm{psd}} M = 4$*, but there does not exist a psd factorization of size four using only rational matrices.*

Proof. Suppose, by way of contradiction, that $(A_1, \ldots, A_8, B_1, \ldots, B_6)$ is a psd factorization of M of size four that uses only rational matrices. By Lemma 4.0.3, each matrix must be rank one. Thus, there exist vectors $\mathbf{a_1}, \ldots, \mathbf{a_8}$ and $\mathbf{b_1}, \ldots, \mathbf{b_6}$ such that $A_i = \phi(\mathbf{a_i})$ and $B_j = \phi(\mathbf{b_j})$. Furthermore, by the properties of the trace, we must have that $M_{ij} = \langle A_i, B_j \rangle = \langle \mathbf{a_i}, \mathbf{b_j} \rangle^2$. Thus, the matrix whose (i,j)th entry is given by $\langle \mathbf{a_i}, \mathbf{b_j} \rangle$ is an entry-wise square root of M, which we denote by S. By looking at the submatrix generated by the first two rows and the fourth and sixth columns, we see that S contains a submatrix $\widetilde{S}$ of the form

$$\begin{pmatrix} \pm 1 & \pm 1 \\ \pm\sqrt{2} & \pm 1 \end{pmatrix}$$

where there is ambiguity on the sign of each entry.

Now by Lemma 4.0.1, each $\mathbf{a_i}$ and $\mathbf{b_j}$ must be a rational vector scaled by a nonzero real number. Hence, there must exist nonzero real numbers $\alpha_1, \alpha_2, \beta_1, \beta_2$ such that the matrix resulting from the product

$$\begin{pmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{pmatrix} \cdot \widetilde{S} \cdot \begin{pmatrix} \beta_1 & 0 \\ 0 & \beta_2 \end{pmatrix} = \begin{pmatrix} \pm\alpha_1\beta_1 & \pm\alpha_1\beta_2 \\ \pm\sqrt{2}\alpha_2\beta_1 & \pm\alpha_2\beta_2 \end{pmatrix}$$

is rational. It is easy to see that if $\alpha_1\beta_1$, $\alpha_1\beta_2$, and $\alpha_2\beta_2$ are rational, then $\alpha_2\beta_1$ must also be rational, which results in a contradiction. □

Chapter 5

WORST-CASE RESULTS FOR POSITIVE SEMIDEFINITE RANK OF POLYGONS [GRT13b]

5.1 A lower bound on the psd rank of generic polytopes

In this section we will focus on lower bounds for the psd ranks of generic polytopes. A polytope $P \subset \mathbb{R}^n$ is said to be *generic* if the coordinates of its vertices form an algebraically independent set over the rationals, i.e. the vertex coordinates do not satisfy any non-trivial polynomial equation with rational coefficients. It is clear that no simple description of such a polytope can be expected. It was shown in [FRT12] that the nonnegative rank of a generic polygon with v vertices is at least $\sqrt{2v}$. Their proof in fact extends to showing that the nonnegative rank of a generic n-dimensional polytope with v vertices is at least $\sqrt{nv}$. We adapt the philosophy of their proof to the psd case to prove the lower bound, $\text{rank}_{\text{psd}}(P) \geq (nv)^{\frac{1}{4}}$.

Theorem 5.1.1. *If $P \subset \mathbb{R}^n$ is a generic polytope with v vertices, then its psd rank is at least $(nv)^{\frac{1}{4}}$.*

Proof. Let $d_k := \frac{k(k+1)}{2}$ be the dimension of the $\mathbb{R}$-vector space $\mathcal{S}^k$ of $k \times k$ real symmetric matrices. Suppose $P \subset \mathbb{R}^n$ is a generic polytope with v vertices and $\text{rank}_{\text{psd}}(P) = k$. Then P is the image under a linear map of a spectrahedron living in $\mathcal{S}^k_+$. Without loss of generality, we may assume that this linear map is the projection onto the first n coordinates and that P can be written as:

$$P = \{(x_1, \ldots, x_n) \,|\, \exists\, x_{n+1}, \ldots, x_{d_k-1} \text{ with } g(x_1, \ldots, x_{d_k-1}) \succeq 0\}$$

where $g(x_1, \ldots, x_{d_k-1}) = x_1 G_1 + \ldots + x_{d_k-1} G_{d_k-1} + G_{d_k}$, each $G_i \in \mathcal{S}^k$.

Let Γ be the set of distinct real entries in the matrices G_i. Then $|\Gamma| \leq d_k^2 \leq k^4$. Consider the extension field $\mathbb{Q}(\Gamma)$ and its real closure $\overline{\mathbb{Q}(\Gamma)}$ (this is simply the real part of the algebraic closure of $\mathbb{Q}(\Gamma)$). The transcendence degree of $\overline{\mathbb{Q}(\Gamma)}$ is at most $|\Gamma|$ (see [Hun74, Chap 6] for the definition of transcendence degree and its basic properties). We now show that the vertex coordinates of P are all contained in $\overline{\mathbb{Q}(\Gamma)}$.

Let $p = (p_1, \ldots, p_n)$ be a vertex of P and $\omega \in \mathbb{Q}^n$ a vector such that the linear program $(\mathcal{L})$: $\max\{\omega^T x : x \in P\}$ has p as its unique optimal point. Let $\widetilde{\omega} := (\omega, 0, \ldots, 0) \in \mathbb{Q}^{d_k - 1}$ and $S := \{(x_1, \ldots, x_{d_k-1}) \,|\, g(x_1, \ldots, x_{d_k-1}) \succeq 0\}$. Then the semidefinite program $(\mathcal{SP})$: $\max\{\widetilde{\omega}^T x : x \in S\}$ has the same optimal value as $(\mathcal{L})$ and the solutions of $(\mathcal{SP})$ are the points in S that project to p.

If $S \cap \mathrm{int}(\mathcal{S}_+^k) = \emptyset$, then S could be written as an affine slice of a proper face of $\mathcal{S}_+^k$. Since proper faces of the psd cone are isomorphic to smaller psd cones [WSV00, Chap 3], this would mean that P could be written as a projection of an affine slice of $\mathcal{S}_+^l$ for some $l < k$, which would contradict our original assumption of $\mathrm{rank}_{\mathrm{psd}}(P) = k$. Hence, we must have that $S \cap \mathrm{int}(\mathcal{S}_+^k) \neq \emptyset$. Thus, Slater's condition (i.e., the feasible set has an interior point; see [WSV00, Chap 4] for details) is satisfied for $(\mathcal{SP})$ and we can apply strong duality. Let $(\mathcal{SD})$ denote the semidefinite program dual to $(\mathcal{SP})$. From semidefinite programming duality theory [WSV00, Chap 4], we know that a pair $(x, Y) \in \mathbb{R}^{d_k-1} \times \mathcal{S}^k$ is an optimal primal-dual pair for the programs $(\mathcal{SP})$, $(\mathcal{SD})$ if and only if they satisfy the following first order conditions:

$$g(x) \succeq 0, \;\; Y \succeq 0, \;\; \langle G_i, Y\rangle = -\widetilde{\omega}_i, \;\; \langle g(x), Y\rangle = 0.$$

These conditions are a series of polynomial equations and inequalities with coefficients in $\mathbb{Q}(\Gamma)$. By the Tarski-Seidenberg Theorem [BCR98, Chap 5], we have that there exists a solution to these equations over $\mathbb{R}$ if and only if there exists a solution to these equations over $\overline{\mathbb{Q}(\Gamma)}$. Since Slater's condition was satisfied, strong duality holds for $(\mathcal{SP})$, and there are solutions (x, Y) over $\mathbb{R}$. By our choice of $(\mathcal{SP})$, we have that each of these solutions (x, Y) is of the form $(p_1, \ldots, p_n, x_{n+1}, \ldots, x_{d_k-1}, Y)$. Hence, the coordinates $p_1, \ldots, p_n$ are contained in $\overline{\mathbb{Q}(\Gamma)}$.

By repeating this procedure for each vertex of P, we see that $\overline{\mathbb{Q}(\Gamma)}$ contains all n coordinates for each of the v vertices. Hence, $\overline{\mathbb{Q}(\Gamma)}$ contains nv algebraically independent elements. Thus, the transcendence degree of $\overline{\mathbb{Q}(\Gamma)}$ is at least nv. Hence, we have that $nv \leq |\Gamma| \leq k^4$. □

In [FRT12], the authors also prove that for each $v \geq 3$, there is a v-gon with integer vertices lying in $[2v] \times [4v^2]$ whose nonnegative rank is $\Omega((v/\log v)^{\frac{1}{2}})$. The same statement also holds in the psd setting with the bound changing to $\Omega((v/\log v)^{\frac{1}{4}})$ as recently shown in [BDP13].

5.2 An upper bound on the psd rank of polygons

The result in the previous section implies that the psd rank of a generic v-gon is at least $(2v)^{\frac{1}{4}}$, while on the other hand, v is a trivial upper bound on the psd rank of any v-gon since its slack matrix has size $v \times v$. This tells us that the worst case rank of a v-gon lies somewhere between the two. In this section we will use some simple geometric tools to show that the trivial upper bound can be slightly improved by a constant to $4\left\lceil \frac{v}{6} \right\rceil$. This result can be stated more generally for matrices of rank three. We will show that if M is a nonnegative matrix of rank three of size $p \times q$, then $\text{rank}_{\text{psd}}(M) \leq 4\left\lceil \frac{\min\{p,q\}}{6} \right\rceil$. These results are analogous to recent results on nonnegative rank of polygons and rank three marices. In [Shi14], Shitov proved that the nonnegative rank of a v-gon is at most $\left\lceil \frac{6v}{7} \right\rceil$, and more generally, that the nonnegative rank of a rank three nonnegative matrix of size $p \times q$ is at most $\left\lceil \frac{6\min\{p,q\}}{7} \right\rceil$.

We begin with a general lemma about psd rank of polytopes.

Lemma 5.2.1. *Let P be a polytope with $\text{rank}_{\text{psd}}(P) = k$, and let $\widetilde{P}$ be a polytope obtained from P by adding either a single inequality to the facet description of P or a single point to the vertex description of P. Then $\text{rank}_{\text{psd}}(\widetilde{P}) \leq k+1$.*

Proof. First, suppose that $\widetilde{P}$ arises by adding a single inequality to the facet description of P. Then there exists some affine halfspace A such that $\widetilde{P} = P \cap A$. Write A in the form $\{x \in \mathbb{R}^n \mid a_0 + a_1x_1 + \ldots + a_nx_n \geq 0\}$. As we saw in the proof of Theorem 5.1.1, we can write P in the form:

$$P = \{(x_1, \ldots, x_n) \,|\, \exists\, x_{n+1}, \ldots, x_{d_k-1} \text{ with } g(x_1, \ldots, x_{d_k-1}) \succeq 0\} \tag{5.1}$$

where $g(x_1, \ldots, x_{d_k-1}) = x_1G_1 + \ldots + x_{d_k-1}G_{d_k-1} + G_{d_k}$, each $G_i \in \mathcal{S}^k$. Now define a vector $\widetilde{a} \in \mathbb{R}^{d_k}$ with $\widetilde{a} = (a_1, \ldots, a_n, 0, \ldots, 0, a_0)$ and define matrices $\widetilde{G}_i \in \mathcal{S}^{k+1}$ where the upper left block is G_i, the lower right diagonal entry is $\widetilde{a}_i$, and all other entries are 0. If we let the $\widetilde{G}_i$'s play the role of the G_i's in (5.1), then this new set will be equal to $\widetilde{P}$. Hence, $\widetilde{P}$ has a lift into $\mathcal{S}_+^{k+1}$ and we have that $\operatorname{rank}_{\text{psd}}(\widetilde{P}) \leq k+1$.

The case when $\widetilde{P}$ arises by adding a point to the vertex description of P follows from the fact that a polytope and its polar both have the same psd rank [GPT13]. □

Example 5.2.2. By Theorem 3.4.7, all triangles and quadrilaterals have psd rank three and any polygon with at least five sides has psd rank at least four. Since a pentagon can be obtained by adding an inequality to the facet description of a quadrilateral, Lemma 5.2.1 implies that all pentagons have psd rank exactly four.

The following lemma is a direct consequence of the definition of psd rank.

Lemma 5.2.3. *Let P be a polytope and suppose there exists a polyhedron Q and a linear map π such that $P = \pi(Q)$. Then* $\operatorname{rank}_{\text{psd}}(P) \leq \operatorname{rank}_{\text{psd}}(Q)$.

Theorem 5.2.4. *Every hexagon has psd rank exactly four.*

Proof. Let H be a hexagon. We know that $\operatorname{rank}_{\text{psd}}(H) \geq 4$ by Theorem 3.4.7. Since psd rank is invariant under invertible affine transformations, we may assume that H has vertices $(1,0)$, (a,b), $(0,1)$, (c,d), $(0,0)$, and (e,f) where (a,b), (c,d), and (e,f) lie in the first, second, and fourth quadrants, respectively, and these points also satisfy $a+b>1$, $c+d<1$, and $e+f<1$.

Consider the polytope O in $\mathbb{R}^3$ with vertices $(0,0,0)$, $(1,0,0)$, $(0,1,0)$, $(0,0,1)$, $(v_1,0,v_3)$, and $(0,w_2,w_3)$, where

$$v_1 = c - \frac{ad}{b},\ v_3 = \frac{d}{b},\ w_2 = f - \frac{be}{a},\ w_3 = \frac{e}{a}.$$

With this choice of coordinates, we see that $v_1 < 0$, $v_3 > 0$, $w_2 < 0$, $w_3 > 0$, $v_1 + v_3 < 1$, and $w_2 + w_3 < 1$. These conditions imply that O is a combinatorial octahedron. In Chapter 3, an octahedron O was defined to be *biplanar* if there exist two distinct planes E_1 and E_2 such that $O \cap E_i$ contains four vertices of O for $i = 1,2$. By intersecting the O defined above with the xz and yz-planes, we see that it is biplanar. Thus by Theorem 3.4.8, we have that

$\operatorname{rank}_{\text{psd}}(O) = 4$. Define a linear map $\pi : \mathbb{R}^3 \to \mathbb{R}^2$ by the matrix $\begin{pmatrix} 1 & 0 & a \\ 0 & 1 & b \end{pmatrix}$. Then $\pi(O) = H$ and by Lemma 5.2.3, $\operatorname{rank}_{\text{psd}}(H) = 4$. This lift of a hexagon to an octahedron is shown in Figure 5.1. □

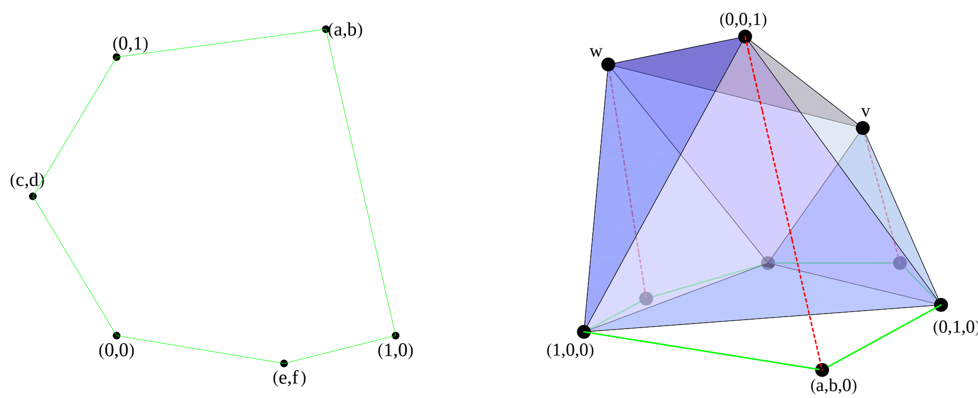


Figure 5.1. This figure depicts the lift of a hexagon to an octahedron shown in the proof of Theorem 5.2.4. The left picture shows a hexagon in a normalized form. The right picture shows the octahedron and its linear projection onto the hexagon. The projection map is the identity on the xy-plane and is depicted by the red dashed lines for the vertices of the octahedron not in the xy-plane.

To obtain our results for nonnegative matrices of rank three, we define a *generalized slack matrix* and an interpretation of its psd rank.

Definition 5.2.5. Let $P \subset \mathbb{R}^n$ be a full-dimensional polytope and $Q \subset \mathbb{R}^n$ be a polyhedron with $P \subseteq Q$. Suppose $P = \operatorname{conv}(p_1, \ldots, p_v)$ is represented as a convex hull of points and $Q = \left\{ x \in \mathbb{R}^n \mid c_j^T x \leq d_j,\ j = 1, \ldots, f \right\}$ is represented by inequalities where $c_j \in \mathbb{R}^n$ and $d_j \in \mathbb{R}$. Then the generalized slack matrix of the pair P, Q is the $v \times f$ nonnegative matrix $S_{P,Q}$ whose (i, j)-entry is $d_j - c_j^T p_i$.

It is a well-known result in the community that the nonnegative rank of $S_{P,Q}$ is the smallest nonnegative rank of a polyhedron R such that $P \subseteq R \subseteq Q$ ([BFPS12]; see [GG12],[Pas12] for related statements). The same result also holds for psd rank, and in fact for any cone rank in the sense of [GPT13]. We include a proof of the psd case, as we will use the result more than once and it does not seem to be written anywhere.

Proposition 5.2.6. *Let P and Q be polyhedra as in Definition 5.2.5, and suppose Q does not contain any lines. Then $\operatorname{rank}_{\text{psd}}(S_{P,Q})$ is equal to the smallest k such that there exists an affine slice L of $\mathcal{S}_+^k$ and a linear map π such that $P \subseteq \pi(L) \subseteq Q$. (We call k the psd rank of the pair P, Q and it measures the smallest possible psd rank of a convex set sandwiched between P and Q.)*

Proof. After translation and rescaling we may assume that

$$Q = \left\{ x \in \mathbb{R}^n \mid c_j^T x \leq 1,\ j = 1, \ldots, f \right\}.$$

Let ℓ denote the psd rank of $S_{P,Q}$. Then $S_{P,Q}$ has a psd factorization through $\mathcal{S}_+^\ell$. Thus there exist matrices $U_1, \ldots, U_v, V_1, \ldots, V_f \in \mathcal{S}_+^\ell$ such that $(S_{P,Q})_{ij} = \langle U_i, V_j \rangle$. Define an affine set

$$A = \left\{ (x, M) \in \mathbb{R}^n \times \mathcal{S}^\ell \mid 1 - c_j^T x = \langle M, V_j \rangle \text{ for all } j = 1, \ldots, f \right\}.$$

Let A_M be the projection of A onto the M coordinates and define $L = A_M \cap \mathcal{S}_+^\ell$. Define π to be the map on L that sends M to any element $x \in \mathbb{R}^n$ where $(x, M) \in A$. This map is well-defined since if (x, M) and (y, M) are both in A, then the points x and y both evaluate to the same values under the inequalities defining Q. Since π is defined on L, we know that $M \succeq 0$ and these slack values must be nonnegative. If x were not equal to y, this would imply that the line containing x and y would be contained in Q, which contradicts our assumptions. Hence, π is well-defined. Since $(p_i, U_i) \in A$, we see that $P \subseteq \pi(L)$. Also, for $z \in \pi(L)$, we have that $c_j^T z \leq 1$ for all $j = 1, \ldots, f$. Thus, $\pi(L) \subseteq Q$. Hence, ℓ is greater than k, the psd rank of the pair P, Q.

For the converse, note that there exists a convex set C with $P \subseteq C \subseteq Q$ such that C has psd rank k. By [GPT13, Theorem 2.4], the slack operator S_C is factorizable through $\mathcal{S}_+^k$, i.e. there exist maps $\sigma : C \to \mathcal{S}_+^k$ and $\tau : C^\circ \to \mathcal{S}_+^k$ such that $1 - \langle x, y \rangle = \langle \sigma(x), \tau(y) \rangle$ for $(x, y) \in C \times C^\circ$. Here C° denotes the polar of C. Then $\sigma(p_1), \ldots, \sigma(p_v), \tau(c_1), \ldots, \tau(c_f)$ give a $\mathcal{S}_+^k$-factorization of $S_{P,Q}$, and so $k \geq \ell$. □

Now suppose we are given a nonnegative $p \times q$ matrix M with $\operatorname{rank}(M) = 3$ and we are interested in $\operatorname{rank}_{\text{psd}}(M)$. First, we may assume that M has no zero rows, since adding or removing zero rows from M will not affect its psd rank. Therefore, if $\mathbb{1}$ denotes the vector of all ones, then $M\mathbb{1}$ is a strictly positive vector. Since scaling the rows of M by positive scalars does not affect the psd rank, we can then assume that $\mathbb{1}$ is in the column span of M. Now consider a *rank factorization* $M = UV$ with $U \in \mathbb{R}^{p\times 3}$ having rows $U_i = (1, u_i^T)$ for $u_i \in \mathbb{R}^2$ and $V \in \mathbb{R}^{3\times q}$. Let

$$P := \operatorname{conv}(u_1, \ldots, u_p) \text{ and } Q := \left\{x \in \mathbb{R}^2 \ : \ (1, x^T)V \geq 0\right\}.$$

Then the pair P, Q satisfies the conditions of Proposition 5.2.6 and $M = S_{P,Q}$. Hence,

$$\operatorname{rank}_{\text{psd}}(M) = \operatorname{rank}_{\text{psd}}(S_{P,Q}) \leq \operatorname{rank}_{\text{psd}}(P)$$

where the inequality follows from Proposition 5.2.6. In particular, a $6\times q$ nonnegative matrix of rank three is the generalized slack matrix of a hexagon inside a q-gon and so has psd rank at most four. Since $\operatorname{rank}_{\text{psd}}(M) = \operatorname{rank}_{\text{psd}}(M^T)$, the psd rank of a $p \times 6$ nonnegative matrix of rank three is also at most four.

Theorem 5.2.7. *Let M be a nonnegative $p \times q$ matrix with usual rank three. Then $\operatorname{rank}_{\text{psd}}(M) \leq 4\left\lceil \frac{\min\{p,q\}}{6} \right\rceil$. In particular, the psd rank of an v-gon is at most $4\left\lceil \frac{v}{6}\right\rceil$.*

Proof. We can write M as the concatenation of $\lceil q/6 \rceil$ matrices with p rows and at most six columns, each of which therefore has psd rank at most four. The result now follows by noting that the psd rank of the concatenation of two matrices is at most the sum of the psd ranks of the individual matrices. Indeed, if $\{A_i\}, \{B_j\}$ factorize M and $\{A_i'\}, \{C_k\}$ factorize M' then the following block diagonal matrices factorize $[M_1 \ M_2]$:

$$\left\{\begin{pmatrix} A_i & 0 \\ 0 & A_i' \end{pmatrix}\right\}, \left\{\begin{pmatrix} B_j & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & C_k \end{pmatrix}\right\}.$$

□

Very little is known about the ranks of v-gons for $v \geq 7$. For example, we know that all 7-gons have psd rank either four or five by Theorem 5.2.4 and Lemma 5.2.1, but there is no concrete heptagon whose psd rank is actually known. We know some 8-gons with psd rank four, but we have no idea how high their psd rank can be, apart from the trivial upper bound of six, obtained again by Lemma 5.2.1. In fact the smallest "concrete" polygons known to have psd rank greater than four are the generic polytopes whose lower bounds are guaranteed by Theorem 5.1.1, which in this case is a generic 129-gon.

Chapter 6

COMPLEXITY OF COMPUTING POSITIVE SEMIDEFINITE RANK [GRT13b, FGP+14]

6.1 Geometry of minimal psd rank

Given a nonnegative matrix M of rank three, a dimension count immediately shows that $\text{rank}_{\text{psd}}(M) \geq 2$. We now derive a geometric characterization of when $\text{rank}_{\text{psd}}(M) = 2$ which will generalize to higher values of rank and yield a complexity result for psd rank. Let our rank three matrix M have size $p \times q$. We may assume, without loss of generality, that M has no all-zero rows or columns, and that $\mathbb{1}$ is in the column span of M. Let $M = UV$ be a rank factorization of M with $\mathbb{1}$ as the first column of U. Let the rows of U be $(1, u_1^T), \ldots, (1, u_p^T)$ and define polyhedra $P := \text{conv}(u_1, \ldots, u_p)$ and $Q := \{x \in \mathbb{R}^2 \mid (1, x^T)V \geq 0\}$ as before. Then $P \subseteq Q$ and $M = S_{P,Q}$.

By Proposition 5.2.6, we know that $\text{rank}_{\text{psd}}(M) = 2$ if and only if there exists a linear map π and an affine space L such that $P \subseteq \pi(L \cap \mathcal{S}^2_+) \subseteq Q$. Since translating P and Q will not affect the slack matrix M, we may assume that $0 \in \text{int}(P)$. Under this assumption, we see that the affine space L cannot be all of $\mathcal{S}^2$. Hence, L must be a two-dimensional slice of $\mathcal{S}^2$ and $\pi|_L$ must be invertible. Since $\mathcal{S}^2_+$ is linearly equivalent to the positive half of the three-dimensional second order cone, $\{(x, y, z) \mid x^2 + y^2 \leq z^2 \text{ and } z \geq 0\}$, we see that $L \cap \mathcal{S}^2_+$ is the linear image of the convex hull of a "half-conic" where half-conics are all ellipses, parabolas, and connected components of hyperbolas in $\mathbb{R}^2$. Finally, we use the fact that the set of conics is invariant under invertible linear maps to see the following.

Proposition 6.1.1. *Let M be a nonnegative rank three matrix. Let $P \subseteq Q \subseteq \mathbb{R}^2$ be the polytope and polyhedron arising from a rank factorization of M as above. Then* $\text{rank}_{\text{psd}}(M) = 2$ *if and only if there exists a half-conic such that its convex hull C satisfies $P \subseteq C \subseteq Q$. In particular if Q is bounded, then* $\text{rank}_{\text{psd}}(M) = 2$ *if and only if we can fit an ellipse between P and Q.*

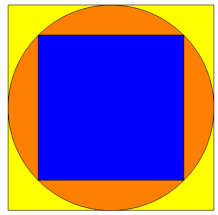


Figure 6.1. Disk nested between P and $(\sqrt{2}/2)P$ where P is the unit square.

Example 6.1.2. Consider the one-parameter family of matrices

$$M_\varepsilon = \begin{bmatrix} 2-\varepsilon & 2-\varepsilon & \varepsilon & \varepsilon \\ \varepsilon & 2-\varepsilon & 2-\varepsilon & \varepsilon \\ \varepsilon & \varepsilon & 2-\varepsilon & 2-\varepsilon \\ 2-\varepsilon & \varepsilon & \varepsilon & 2-\varepsilon \end{bmatrix},$$

with $\varepsilon \in [0,1]$. For $\varepsilon \neq 1$ this matrix has rank 3, and we would like to know for which (if any) values of ε we get $\operatorname{rank}_{\text{psd}}(M) = 2$. Note that $M_\varepsilon = S_{(1-\varepsilon)P,P}$, where P is the ± 1 square. It is easy to see that we can put a half-conic between $(1-\varepsilon)P$ and P if and only if $1-\varepsilon \leq \sqrt{2}/2$ as seen in Figure 6.1. Since it is known that the square itself has psd rank three, Proposition 5.2.6 allows us to completely determine the psd ranks of this matrix family:

$$\operatorname{rank}_{\text{psd}} M_\varepsilon = \begin{cases} 1 & \text{if } \varepsilon = 1; \\ 2 & \text{if } \varepsilon \in [1-\sqrt{2}/2, 1); \\ 3 & \text{if } \varepsilon \in [0, 1-\sqrt{2}/2). \end{cases}$$

The geometric techniques used above generalize to higher rank matrices. Let $M \in \mathbb{R}^{p \times q}$ be a nonnegative matrix of rank $d = \binom{k+1}{2}$. Then by a dimension count, $\operatorname{rank}_{\text{psd}}(M) \geq k$. Thus we can ask the following decision problem about M:

Definition 6.1.3. MIN PSD RANK: Given a nonnegative matrix M of rank $\binom{k+1}{2}$, is $\operatorname{rank}_{\text{psd}}(M) = k$?

For ease in working with higher dimensions, we will switch from the polytope viewpoint used above to a conic viewpoint. In the remainder of this section $d = \binom{k+1}{2} = \operatorname{rank}(M)$.

Let $M = UV$ be a rank factorization and let P, Q be the cones $P = \text{cone}(u_1, \ldots, u_p)$ and $Q = \{x \in \mathbb{R}^d \,|\, x^T V \geq 0\}$ where u_i are the rows of U. Then P and Q are d-dimensional cones with $P \subseteq Q$ and $M = S_{P,Q}$ where $S_{P,Q}$ is a generalized slack matrix of the pair of cones P, Q, defined analogously to that for pairs of polyhedra. Using Proposition 5.2.6 and counting dimensions, we get the following geometric characterization of the MIN PSD RANK problem:

Proposition 6.1.4. *The psd rank of M is k if and only if there is an invertible linear map $\pi : \mathcal{S}^k \to \mathbb{R}^d$ such that $P \subseteq \pi(\mathcal{S}^k_+) \subseteq Q$.*

In [Vav09], Vavasis defined EXACT NMF (*Nonnegative Matrix Factorization*) as the problem of determining whether the nonnegative rank of a given matrix M equals its rank. He also defined INTERMEDIATE SIMPLEX which asks, given two nested polyhedra $P \subseteq Q$, if there is a simplex T such that $P \subseteq T \subseteq Q$. He proceeded to show that EXACT NMF is equivalent to INTERMEDIATE SIMPLEX. The above reduction of MIN PSD RANK to the geometric condition of Proposition 6.1.4 can be thought of as the psd analog to the equivalence shown by Vavasis.

Now we will reduce the geometric criterion into a semialgebraic set feasibility problem. Consider the basis of $\mathcal{S}^k$ given by the elementary symmetric matrices E_{ij} defined as follows. Let E_{ii} be the matrix with a one in position (i, i) and zeros everywhere else. For $i < j$, let E_{ij} be the matrix with $\frac{1}{\sqrt{2}}$ in positions (i, j) and (j, i) and zeros everywhere else. This basis allows a natural bijection between $\mathcal{S}^k$ and $\mathbb{R}^d$ by identifying a symmetric matrix $Y = \sum_{1 \leq i \leq j \leq d} E_{ij} y_{ij}$ with the vector $y = (y_{ij}) \in \mathbb{R}^d$. Note that this bijection preserves the inner product in $\mathcal{S}^k$ (this is the reason for the $\sqrt{2}$ factors). Let L be the $r \times r$ nonsingular matrix representing the invertible linear map π with respect to the above basis. Then $\pi(Y) = Ly$, and $\pi^{-1} : \mathbb{R}^d \to \mathcal{S}^k$ sends $z \mapsto L^{-1} z$ where $L^{-1} z$ corresponds to a matrix in $\mathcal{S}^k$ under the bijection discussed above. We can now write down the conditions given by Proposition 6.1.4 in terms of L and L^{-1}.

The condition that $P \subseteq \pi(\mathcal{S}^k_+)$ is equivalent to $\pi^{-1}(u_i) \in \mathcal{S}^k_+$ for every generator u_i of P. Thus we need $L^{-1} u_i \succeq 0$ for each row u_i of U. Note that each entry in the symmetric matrix corresponding to $L^{-1} u_i$ is a linear polynomial in the entries of L^{-1}. The condition

$\pi(\mathcal{S}_+^k) \subseteq Q$ says that for each column v_j of V, the linear inequality $v_j^T x \geq 0$ is valid on $\pi(\mathcal{S}_+^k)$, or equivalently, that for every matrix $A \in \mathcal{S}_+^k$, $v_j^T(\pi(A)) \geq 0$. Therefore, we get that for every column v_j of V, the symmetric matrix corresponding to $v_j^T L$ is psd. Putting all this together we get the following reduction of the MIN PSD RANK problem.

Proposition 6.1.5. *The matrix M has psd rank k if and only if there are two matrices $L, K \in \mathbb{R}^{d\times d}$ such that*

1. *L is the inverse of K, i.e., $LK = KL = I$,*

2. *The $k \times k$ linear matrix inequality $Ku_i \succeq 0$ holds for each row u_i of U,*

3. *The $k \times k$ linear matrix inequality $v_j^T L \succeq 0$ holds for each column v_j of V.*

Further, the above system can be written down in polynomial time from M.

Proof. The equivalence of MIN PSD RANK and the feasibility of the above system was argued in the discussion before the proposition. The scalars in the system come from a rank factorization of M which can be done in polynomial time. □

The number of variables in the above semialgebraic system depends only on k and not on the size of the input matrix M. In [Ren92], Renegar showed that the feasibility of a system of m polynomial inequalities and equalities in ℓ variables with degree at most j can be determined in time $(mj)^{O(\ell)}$. Here, Renegar used the Blum-Shub-Smale model of complexity for computing with real numbers, so the only requirement on the coefficients of the polynomials is that they are real numbers. We use this to get a complexity result for MIN PSD RANK.

Theorem 6.1.6. *Using the Blum-Shub-Smale model of complexity, the problem MIN PSD RANK can be solved in time $(pq)^{O(d^{2.5})}$ where $p \times q$ is the dimension of the input matrix M and $d = \binom{k+1}{2}$ is the rank of M. In particular, for fixed rank, the problem MIN PSD RANK can be solved in polynomial time.*

Proof. First, we consider the problem formulated in Proposition 6.1.5. This problem can be formulated as the existence of a solution to a system of $d^2 + 2^k(p+q)$ polynomial equalities and inequalities in $2d^2$ variables with each polynomial having degree less than or equal to k. By applying [Ren92] and noting that $d \sim k^2$, we see that this problem can be solved in time $(pq)^{O(d^{2.5})}$. We conclude by noting that MIN PSD RANK can be reduced to the above problem in time polynomial in pq. □

In [Vav09], Vavasis showed that EXACT NMF is NP-Hard. The corresponding question for MIN PSD RANK is still open. We can consider the more general problem: given a nonnegative $p \times q$ matrix M and a number k, determine if $\text{rank}_{\text{psd}}(M) \leq k$. For the analogous problem with nonnegative rank, Moitra [Moi13] showed an algorithm that runs in time $(pq)^{O(k^2)}$. Theorem 6.1.6 can be seen as a restricted psd analog of Moitra's result.

6.2 Computing square root rank is NP-hard

While we strongly suspect that it is NP-hard to compute psd rank, there is no proof of this fact at the moment. The situation is clearer for square root rank.

Theorem 6.2.1. *The square root rank of a nonnegative matrix is NP-hard to compute.*

Proof. Recall that given a list of n positive integers $a_1, \ldots, a_n$, the partition problem asks whether there exist sign choices $s_1, \ldots, s_n \in \{-1, 1\}$ such that $\sum_{i=1}^n s_i a_i = 0$. This problem is known to be NP-complete [GJ79].

Given the integers $a_1, \ldots, a_n$, define an $(n+1) \times (n+1)$ matrix A of the form:

$$\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & a_1^2 \\ 0 & 1 & 0 & \cdots & 0 & a_2^2 \\ 0 & 0 & 1 & \cdots & 0 & a_3^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & a_n^2 \\ 1 & 1 & 1 & \cdots & 1 & 0 \end{pmatrix}.$$

Since A contains the $n \times n$ identity matrix as a submatrix, the square root rank of A must be either n or $n+1$. If $\sqrt{A}$ is a Hadamard square root, then we may scale rows and columns of $\sqrt{A}$ by -1 and not affect the rank. Thus, we may assume that the first n columns of $\sqrt{A}$ are composed of zeros and positive ones. With this assumption, we see immediately that there exists a Hadamard square root of rank n if and only if the partition problem for $a_1, \ldots, a_n$ is satisfiable. $\square$

Remark 6.2.2. Although the partition problem is NP-complete, it is only weakly NP-complete and admits a pseudo-polynomial time algorithm. Thus, the above theorem does not rule out the existence of an algorithm for deciding $\text{rank}_{\sqrt{}}$ that runs in time polynomial in the problem dimension and the *magnitude* (not encoding length) of the matrix entries. Furthermore, this embedding of the partition problem cannot hope to show that psd rank is NP-hard to compute. To see this, consider the matrix A corresponding to the partition problem with integers 5, 12, and 13:

$$\begin{pmatrix} 1 & 0 & 0 & 25 \\ 0 & 1 & 0 & 144 \\ 0 & 0 & 1 & 169 \\ 1 & 1 & 1 & 0 \end{pmatrix}.$$

This instance of the partition problem is not satisfiable, yet the matrix A has a 3×3 psd factorization. Such a factorization is obtained by placing the matrices

$$\begin{pmatrix} 1 & 0 & -\frac{5}{13} \\ 0 & 1 & -\frac{12}{13} \\ -\frac{5}{13} & -\frac{12}{13} & 1 \end{pmatrix}, \begin{pmatrix} 25 & 60 & 65 \\ 60 & 144 & 156 \\ 65 & 156 & 169 \end{pmatrix}$$

on the fourth row and the fourth column, respectively, of A and by placing the standard basis factorization of the identity in the first three rows and columns.

Chapter 7

SPACES OF POSITIVE SEMIDEFINITE MATRIX FACTORIZATIONS [FGP+14]

In this chapter, we fix a nonnegative matrix M and consider the set of all valid psd factorizations of M as a topological space. In the special case where rank $(M) = \binom{\text{rank}_{\text{psd}}(M)+1}{2}$, we show in Proposition 7.0.2 and Corollary 7.0.3 that this topological space is closely related to the space of all linear images of the psd cone that nest between two polyhedral cones coming from M. An extension of this result to general M is not possible, as seen in Example 7.0.4. In Examples 7.0.6 and 7.0.7, we use this machinery to construct psd factorizations from the linear embedding of the psd cone. Finally in Proposition 7.0.8, we show that for rank three matrices with psd rank two, the space of psd factorizations is connected. This contrasts with the nonnegative rank case where it is known that the space of nonnegative factorizations can be disconnected for rank three matrices with nonnegative rank three [MSvS03].

For this section, let $M \in \mathbb{R}_+^{p\times q}$ be a nonnegative matrix with psd rank k. As before, we define a *psd factorization* to be a set of matrices $(A_1, \ldots, A_p, B_1, \ldots, B_q) \in \left(\mathcal{S}^k\right)^{p+q}$ such that each of the component matrices is psd and $M_{ij} = \langle A_i, B_j\rangle$ for each entry in M. We define the set of all such psd factorizations to be the *space of psd factorizations* associated to M and denote it by $\mathcal{SF}(M)$. Note that this definition only considers matrices whose size is equal to the psd rank of M.

As a warm-up, it is straightforward to see that $\mathcal{SF}(M)$ is closed and infinite. To see that $\mathcal{SF}(M)$ is infinite, simply note that for any psd factorization $(A_1, \ldots, B_q)$ and any matrix $L \in GL(k)$ (the group of invertible $k \times k$ matrices), the matrices $\left(L^T A_1 L, \ldots, L^{-1} B_q L^{-T}\right)$ also form a psd factorization of M. We refer to the set of all such psd factorizations as the *orbit* of $(A_1, \ldots, B_q)$ in $\mathcal{SF}(M)$. In some cases the entire space of psd factorizations is equivalent to a single orbit.

Example 7.0.1. Let M be the 3×3 derangement matrix (i.e. $M_{ii} = 0$ and $M_{ij} = 1$ for $i \neq j$). This matrix has usual rank three and psd rank two as shown by the factorization

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}.$$

Let $\mathcal{X}$ denote an arbitrary psd factorization of M. The zero pattern of M implies that the matrices composing $\mathcal{X}$ must all be rank one. Now it is straightforward to see that there exists an invertible matrix such that conjugation by this matrix will send $\mathcal{X}$ to the explicit factorization above. Hence, $\mathcal{SF}(M)$ is composed of a single orbit.

The next proposition gives a geometric picture of the space of factorizations in the special case where $\text{rank}_{\text{psd}}(M) = k$ and $\text{rank}(M) = \binom{k+1}{2}$. By taking a rank factorization, it is easy to see that any nonnegative matrix M can be viewed as $S_{P,Q}$ for some polyhedral cones P and Q whose dimension is equal to $\text{rank}(M)$. Furthermore, Theorem 1.2.8 tells us that $\text{rank}_{\text{psd}}(S_{P,Q})$ is the smallest integer k for which there exists a subspace $\mathcal{L}$ and a linear map π such that $P \subset \pi(\mathcal{S}^k_+ \cap \mathcal{L}) \subset Q$. In the special case where the cones P and Q come from a matrix M with $\text{rank}_{\text{psd}}(M) = k$ and $\text{rank}(M) = \binom{k+1}{2}$, we can count dimensions to see that the map π is invertible and the subspace $\mathcal{L}$ is all of $\mathcal{S}^k$. If we define $\Delta_k(P,Q)$ to be the space of all linear maps $\pi : \mathcal{S}^k \to \mathbb{R}^{\binom{k+1}{2}}$ such that $P \subset \pi(\mathcal{S}^k_+) \subset Q$, then $\Delta_k(P,Q)$ is nonempty with M, P, and Q as above. We can actually say much more about $\Delta_k(P,Q)$ in this special case.

Proposition 7.0.2. *Let $M \in \mathbb{R}^{p \times q}_+$ with $\text{rank}(M) = \binom{k+1}{2}$ and $\text{rank}_{\text{psd}}(M) = k$. Fix a rank factorization of $M = UV$ where u_i is the ith row of U and v_j is the jth column of V. Let $P = \text{cone}(u_1, \dots, u_p)$ and $Q = \left\{x \mid v_j^T x \geq 0 \text{ for all } j\right\}$ be the cones generated by this rank factorization so that $M = S_{P,Q}$. Then $\mathcal{SF}(M)$ is homeomorphic to $\Delta_k(P,Q)$.*

Proof. Suppose $(A_1, \dots, B_q)$ is a psd factorization of M. The set $(A_1, \dots, A_p)$ spans $\mathcal{S}^k$ (else we could find a lower dimensional rank factorization of M), so we can define a linear map π by making $\pi(A_i) = u_i$. This map is well-defined since if $\sum_i \alpha_i A_i$ and $\sum_j \beta_j A_j$ are two representations of the same matrix in $\mathcal{S}^k$, then we have that $\left(\sum_i \alpha_i u_i - \sum_j \beta_j u_j\right)^T V = 0$. Since V has full row rank, this implies that $\sum_i \alpha_i u_i = \sum_j \beta_j u_j$. By the definition of π, it is immediate that $P \subset \pi(\mathcal{S}_+^k)$. Since π has the property that $\langle \pi(L), v_j \rangle = \langle L, B_j \rangle$ for each j, we also have that $\pi(\mathcal{S}_+^k) \subset Q$. Thus, we have defined a map from $\mathcal{SF}(M)$ to $\Delta_k(P, Q)$.

Next, suppose that we have $\pi \in \Delta_k(P, Q)$. Define $A_i = \pi^{-1}(u_i)$ and $B_j = \pi^*(v_j)$ where π^* is the adjoint map. Then $A_i \in \mathcal{S}_+^k$ and $B_j \in (\mathcal{S}_+^k)^* = \mathcal{S}_+^k$ and these matrices form a psd factorization of M. This map is the inverse of the one defined above and both of the maps are continuous. Hence, the spaces are homeomorphic. □

Both of the spaces in the previous proposition permit a natural action by $GL(k)$. The action on $\mathcal{SF}(M)$ was mentioned above when we discussed the orbits of $\mathcal{SF}(M)$. The action on $\Delta_k(P, Q)$ takes the form $g \cdot \pi(L) = \pi(gLg^T)$, i.e. we compose the map π with an automorphism of the psd cone. The homeomorphism in the previous proposition respects these group actions so we can descend to the quotient to see the following.

Corollary 7.0.3. *Under the same assumptions as the previous proposition, the spaces $\mathcal{SF}(M)/GL(k)$ and $\Delta_k(P, Q)/GL(k)$ are homeomorphic. Furthermore, $\Delta_k(P, Q)/GL(k)$ is homeomorphic to the space of all linear images C of $\mathcal{S}_+^k$ such that $P \subset C \subset Q$.*

Proof. The first statement is shown by descending to the quotient as discussed prior to the corollary. The second homeomorphism is just given by $[\pi] \mapsto \pi(\mathcal{S}_+^k)$. It is straightforward to check that this map is a well-defined homeomorphism. □

The next example shows that Corollary 7.0.3 cannot hold for general M.

Example 7.0.4. Let M be the slack matrix of the regular hexagon, i.e. M is the 6×6 circulant matrix defined by the vector $(0, 1, 2, 2, 1, 0)$. It was shown in [GRT13a] that M has rank three, psd rank four, and at least two distinct factorization orbits (since there exists a factorization consisting entirely of rank one matrices and another factorization using both rank one and rank two matrices). Since this matrix is a slack matrix of a polytope, however, the cones P and Q must be equal and the only image nested between them must be P itself. Hence, there cannot exist a bijection between factorization orbits and images nested between P and Q.

In the following, we apply our machinery to matrices with rank three and psd rank two.

Remark 7.0.5. In light of Corollary 7.0.3, we can gain new insight on Example 7.0.1. By taking the trivial rank factorization of the 3×3 derangement matrix, we obtain the cones

$$P = \text{cone}\left((0,1,1),(1,0,1),(1,1,0)\right) \subset \mathbb{R}^3_+ = Q.$$

Dehomogenizing these cones gives us the two triangles seen in Figure 7.1. In this dehomogenized picture, linear images of the psd cone correspond to ellipses and it is straightforward to see that there is a unique ellipse that fits between the two triangles. Hence, $\Delta_2(P,Q)/GL(2)$ consists of a single point and by the corollary, the space of psd factorizations is composed of a single orbit.

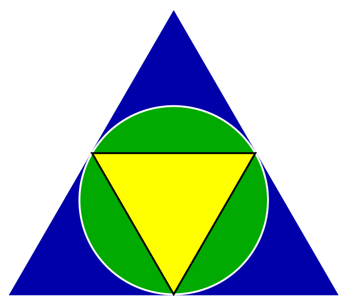


Figure 7.1. A space of psd factorizations consisting of a single orbit: The yellow and blue triangles correspond to the dehomogenized cones coming from a rank factorization of the 3×3 derangement matrix. The green circle is the unique ellipse that can be nested between the two triangles.

We now show how to apply Proposition 7.0.2 to find different psd factorizations of a matrix.

Example 7.0.6. In this example, we consider the following matrix M of rank three (shown along with a rank factorization):

$$\begin{pmatrix} 3 & 3 & 1 & 1 \\ 1 & 3 & 3 & 1 \\ 1 & 1 & 3 & 3 \\ 3 & 1 & 1 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & -1 \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 2 & 2 & 2 \\ 0 & 1 & 0 & -1 \\ 1 & 0 & -1 & 0 \end{pmatrix}.$$

By forming the cones P and Q corresponding to this rank factorization and then dehomogenizing through the plane $\{(1, x_2, x_3)\}$, we see that P corresponds to the square centered at the origin of side length two and that Q corresponds to the same square scaled by a factor of two (see Figure 7.2). Now any linear image of $\mathcal{S}_+^2$ corresponds to an ellipse in this dehomogenized picture. So to get a psd factorization of M, we just need to pick an ellipse, figure out a linear image of $\mathcal{S}_+^2$ that corresponds to it, and apply the homeomorphism discussed in Proposition 7.0.2.

For the circle centered at the origin with radius $\sqrt{2}$, we get the following (where $\alpha = \frac{1}{\sqrt{2}}$):

$$\begin{pmatrix} 1+\alpha & \alpha \\ \alpha & 1-\alpha \end{pmatrix}, \begin{pmatrix} 1-\alpha & \alpha \\ \alpha & 1+\alpha \end{pmatrix}, \begin{pmatrix} 1-\alpha & -\alpha \\ -\alpha & 1+\alpha \end{pmatrix}, \begin{pmatrix} 1+\alpha & -\alpha \\ -\alpha & 1-\alpha \end{pmatrix},$$

$$\begin{pmatrix} 1+\alpha & 0 \\ 0 & 1-\alpha \end{pmatrix}, \begin{pmatrix} 1 & \alpha \\ \alpha & 1 \end{pmatrix}, \begin{pmatrix} 1-\alpha & 0 \\ 0 & 1+\alpha \end{pmatrix}, \begin{pmatrix} 1 & -\alpha \\ -\alpha & 1 \end{pmatrix}.$$

For the ellipse centered at the origin with horizontal axis of length four and vertical axis of length three, we get the factorization:

$$\begin{pmatrix} 5/3 & 1/2 \\ 1/2 & 1/3 \end{pmatrix}, \begin{pmatrix} 1/3 & 1/2 \\ 1/2 & 5/3 \end{pmatrix}, \begin{pmatrix} 1/3 & -1/2 \\ -1/2 & 5/3 \end{pmatrix}, \begin{pmatrix} 5/3 & -1/2 \\ -1/2 & 1/3 \end{pmatrix},$$

$$\begin{pmatrix} 7/4 & 0 \\ 0 & 1/4 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 1/4 & 0 \\ 0 & 7/4 \end{pmatrix}, \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}.$$

It is interesting to note how the ranks of the factors change depending on whether the ellipse contacts the vertices of P or the facets of Q. For example, in the second factorization, the matrices corresponding to the columns are rank one exactly when the corresponding facet of the outer square is tight to the ellipse. Of course, this is not a coincidence, but due to how we construct the factorization once we know the linear embedding of the psd cone.

In every example of a psd factorization that has been presented so far, either the matrices corresponding to the rows or those corresponding to the columns can be chosen to be rank one matrices. Initial attempts to construct a matrix without this property proved fruitless. With the machinery of this section, finding such an example becomes almost trivial.

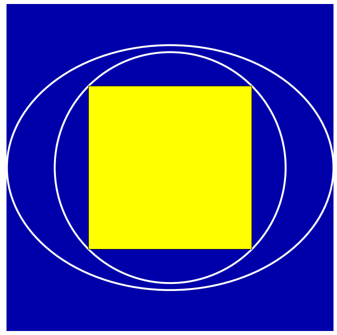


Figure 7.2. This shows the situation described in Example 7.0.6. The inner and outer squares correspond to the dehomogenized cones P and Q and any ellipse nested between the two squares corresponds to an orbit of psd factorizations. In the example, we showed factorizations corresponding to both the circle and the ellipse drawn in the figure.

Example 7.0.7. In this example, we present a 4×4 matrix with psd rank two such that every 2×2 psd factorization must have a rank two matrix on a row and a rank two matrix on a column. To construct this example, we start with the 3×3 derangement matrix as in Example 7.0.1, which corresponds to the picture shown in Figure 7.1. Now we add an extra vertex to the inner triangle and an extra facet to the outer triangle so that neither the new vertex nor the new facet touch the circle, as shown in Figure 7.3. This corresponds to a new matrix

$$M = \begin{bmatrix} 0 & 1 & 1 & 2 \\ 1 & 0 & 1 & 2 \\ 1 & 1 & 0 & 6 \\ 1 & 1 & 3 & 3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 3 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & -1 \end{bmatrix}.$$

The same circle as before is still the unique ellipse nested between the two polytopes so the space of factorizations consists of a single orbit. When we construct a psd factorization in this orbit, the matrix corresponding to the new vertex must lie in the interior of $\mathcal{S}^2_+$ and the matrix corresponding to the new facet must lie in the interior of $(\mathcal{S}^2_+)^*$. Hence, they must have rank two. Such a factorization may be obtained by augmenting our previous factorization for the derangement matrix with the matrix $\begin{bmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{bmatrix}$ for the new vertex and the matrix $\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$ for the new facet.

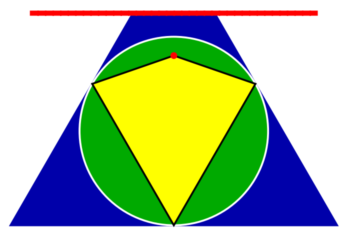


Figure 7.3. Here we take the picture corresponding to the 3×3 derangement matrix and add an additional facet to the outer polytope and an additional vertex to the inner polytope (shown in red). Since the new facet and vertex are not incident to the boundary of the linear embedding of $\mathcal{S}_+^2$, their corresponding matrices in the psd factorization must have full rank.

For the special case where M has rank three and we are considering 2×2 psd factorizations, we can show that the space of factorization orbits is connected.

Proposition 7.0.8. *Let M be a nonnegative $p \times q$ matrix with psd rank two and usual rank three. Then $\mathcal{SF}(M)/GL(2)$ is connected.*

Proof. Let P and Q be the cones in $\mathbb{R}^3$ arising from a rank factorization as above. By Corollary 7.0.3, it is enough to show that $\Delta_2(P, Q)/GL(2)$ is connected. To do this, we will dehomogenize the cones so that we can work with polytopes and ellipses.

The cone Q must be pointed since it was formed from a full-rank matrix. Thus, we can find an affine hyperplane such that the dehomogenization of Q through this hyperplane is bounded. We dehomogenize through this hyperplane to get polytopes $\widetilde{P} \subset \widetilde{Q}$. Any element of $\Delta_2(P, Q)/GL(2)$ corresponds to an ellipse nested between $\widetilde{P}$ and $\widetilde{Q}$. Thus, to finish the proof, it is enough to show that any two ellipses that are nested between $\widetilde{P}$ and $\widetilde{Q}$ can be connected by a path of ellipses.

Suppose E_0 and E_1 are ellipses nested between the two polytopes. Then there exist quadratic polynomials q_0 and q_1 such that $E_i = \{x \mid q_i(x) \geq 0\}$. For $t \in [0,1]$, define a quadratic polynomial $q_t = (1-t)q_0 + tq_1$ and the corresponding ellipse E_t. Since q_0 and q_1 are nonnegative on the points of $\widetilde{P}$, so is q_t and we have that $\widetilde{P} \subset E_t$. Similarly, since q_0 and q_1 are negative on $(E_0 \cup E_1)^c$, we have that $E_t \subset E_0 \cup E_1 \subset \widetilde{Q}$. Thus, E_t gives the desired path of ellipses. □

We are not sure if Proposition 7.0.8 extends to matrices M with $\operatorname{rank}_{\text{psd}}(M) = k$ and $\operatorname{rank}(M) = \binom{k+1}{2}$ for $k > 2$. The proof in the $k = 2$ case relied on the fact that bounded spectrahedra in $\mathcal{S}_+^2$ can be represented by a single polynomial inequality. Higher dimensional spectrahedra require several polynomial inequalities and it is not clear if the proof can be extended to this setting. Searching for a counterexample has also been difficult, since the next case involves linear images of $\mathcal{S}_+^3$ nested between six-dimensional cones.

BIBLIOGRAPHY

[AGH+12] S. Arora, R. Ge, Y. Halpern, D. Mimno, A. Moitra, D. Sontag, Y. Wu, and M. Zhu. A practical algorithm for topic modeling with provable guarantees. *arXiv:1212.4777*, 2012.

[AGKM12] S. Arora, R. Ge, R. Kannan, and A. Moitra. Computing a nonnegative matrix factorization–provably. In *Proceedings of the 44th Symposium on Theory of Computing (STOC)*, pages 145–162. ACM, 2012.

[BCR98] J. Bochnak, M. Coste, and M. Roy. *Real Algebraic Geometry*. Springer-Verlag, 1998.

[BDP13] J. Briët, D. Dadush, and S. Pokutta. On the existence of 0/1 polytopes with high semidefinite extension complexity. In *Algorithms, ESA 2013*, volume 8125 of *Lecture Notes in Computer Science*, pages 217–228. Springer Berlin Heidelberg, 2013.

[BFPS12] G. Braun, S. Fiorini, S. Pokutta, and D. Steurer. Approximation limits of linear programs (beyond hierarchies). *2013 IEEE 54th Annual Symposium on Foundations of Computer Science*, 0:480–489, 2012.

[Bjö95] A. Björner. Topological methods. In *Handbook of Combinatorics (Vol. 2)*, pages 1819–1872. MIT Press, Cambridge, MA, USA, 1995.

[Bre93] G.E. Bredon. *Topology and Geometry*. Graduate Texts in Mathematics. Springer, 1993.

[BTN01] A. Ben-Tal and A. Nemirovski. On polyhedral approximations of the second-order cone. *Mathematics of Operations Research*, 26(2):193–205, 2001.

[CLO92] D. Cox, J. Little, and D. O'Shea. *Ideals, Varieties, and Algorithms*. Springer-Verlag New York, Inc., 1992.

[CR93] J.E. Cohen and U.G. Rothblum. Nonnegative ranks, decompositions, and factorizations of nonnegative matrices. *Linear Algebra and its Applications*, 190:149–168, 1993.

[CRST06] M. Chudnovsky, N. Robertson, P. Seymour, and R. Thomas. The strong perfect graph theorem. *Annals of Mathematics*, 164:51–229, 2006.

[Dye83] M.E. Dyer. The complexity of vertex enumeration methods. *Mathematics of Operations Research*, 8(3):381–402, 1983.

[FGP+14] H. Fawzi, J. Gouveia, P. Parrilo, R.Z. Robinson, and R.R. Thomas. Positive semidefinite rank. *arXiv:1407.4095*, 2014.

[FGR14] H. Fawzi, J. Gouveia, and R. Z. Robinson. Rational and real positive semidefinite rank can be different. *arXiv:1404.4864*, 2014.

[FKPT13] S. Fiorini, V. Kaibel, K. Pashkovich, and D. O. Theis. Combinatorial bounds on nonnegative rank and extended formulations. *Discrete Mathematics*, 313(1):67 – 83, 2013.

[FMP+12] S. Fiorini, S. Massar, S. Pokutta, H.R. Tiwary, and R. de Wolf. Linear vs. semidefinite extended formulations: Exponential separation and strong lower bounds. In *Proceedings of the Forty-fourth Annual ACM Symposium on Theory of Computing*, STOC '12, pages 95–106. ACM, 2012.

[FO85] R. Freund and J. Orlin. On the complexity of four polyhedral set containment problems. *Mathematical Programming*, 33(2):139–145, 1985.

[FRT12] S. Fiorini, T. Rothvoss, and H.R. Tiwary. Extended formulations for polygons. *Discrete Comput. Geom.*, 48(3):658–668, 2012.

[GG12] N. Gillis and F. Glineur. On the geometric interpretation of the nonnegative rank. *Linear Algebra and its Applications*, 437(11):2685–2712, 2012.

[GGK+13] J. Gouveia, R. Grappe, V. Kaibel, K. Pashkovich, R.Z. Robinson, and R.R. Thomas. Which nonnegative matrices are slack matrices? *Linear Algebra and its Applications*, 439(10):2921 – 2933, 2013.

[GJ79] M. R. Garey and D. S. Johnson. *Computers and Intractability: A guide to the theory of NP-completeness.* W. H. Freeman and Company, 1979.

[GLS88] M. Grötschel, L. Lovász, and A. Schrijver. *Geometric algorithms and combinatorial optimization.* Algorithms and combinatorics. Springer-Verlag, 1988.

[Goe14] M. Goemans. Smallest compact formulation for the permutahedron. *Mathematical Programming Series B*, pages 1–7, 2014.

[GPT10] J. Gouveia, P.A. Parrilo, and R.R. Thomas. Theta bodies for polynomial ideals. *SIAM J. Optim.*, 20(4):2097–2118, 2010.

[GPT13] J. Gouveia, P.A. Parrilo, and R.R. Thomas. Lifts of convex sets and cone factorizations. *Mathematics of Operations Research*, 38(2):248–264, 2013.

[GRT13a] J. Gouveia, R. Z. Robinson, and R. R. Thomas. Polytopes of minimum positive semidefinite rank. *Discrete & Computational Geometry*, 50(3):679–699, 2013.

[GRT13b] J. Gouveia, R. Z. Robinson, and R. R. Thomas. Worst-case results for positive semidefinite rank. *arXiv:1305.4600*, 2013.

[Grü67] B. Grünbaum. *Convex Polytopes*. Graduate Texts in Mathematics. Interscience, 1967.

[GS] D. Grayson and M. Stillman. Macaulay2, a software system for research in algebraic geometry. Available at http://www.math.uiuc.edu/Macaulay2/.

[Hun74] T. Hungerford. *Algebra*. Holt, Rinehart, and Winston, 1974.

[Jos04] M. Joswig. Software. In *Handbook of Discrete and Computational Geometry, Second Edition*, Discrete and Combinatorial Mathematics Series. Taylor & Francis, 2004.

[JSWZ13] R. Jain, Y. Shi, Z. Wei, and S. Zhang. Efficient protocols for generating bipartite classical distributions and quantum states. *IEEE Transactions on Information Theory*, 59(8):5171–5178, 2013.

[JZ04] M. Joswig and G. Ziegler. Convex hulls, oracles, and homology. *Journal of Symbolic Computation*, 38(4):1247 – 1259, 2004.

[KP03] V. Kaibel and M. Pfetsch. Some algorithmic problems in polytope theory. In *Algebra, Geometry, and Software Systems*, pages 23–47. Springer-Verlag, 2003.

[KP11] V. Kaibel and K. Pashkovich. Constructing extended formulations from reflection relations. In *Proceedings of the 15th International Conference on Integer Programming and Combinatoral Optimization*, IPCO'11, pages 287–300. Springer-Verlag, 2011.

[Lov79] L. Lovász. On the shannon capacity of a graph. *IEEE Trans. Inf. Theor.*, 25(1):1–7, 1979.

[LS91] L. Lovsz and A. Schrijver. Cones of matrices and set-functions and 0-1 optimization. *SIAM Journal on Optimization*, 1:166–190, 1991.

[LS99] D. D. Lee and H. S. Seung. Learning the parts of objects by non-negative matrix factorization. *Nature*, 401(6755):788–791, 1999.

[LT12] T. Lee and D. O. Theis. Support-based lower bounds for the positive semidefinite rank of a nonnegative matrix. *arXiv:1203.3961*, 2012.

[Moi13] A. Moitra. An almost optimal algorithm for computing nonnegative rank. In *SODA*, pages 1454–1464. SIAM, 2013.

[MR80] T.H. Matheiss and D.S. Rubin. A survey and comparison of methods for finding all vertices of convex polyhedral sets. *Mathematics of Operations Research*, 5(2):167–185, 1980.

[MSvS03] D. Mond, J. Smith, and D. van Straten. Stochastic factorizations, sandwiched simplices and the topology of the space of explanations. *Royal Soc. Proc., Math. Phys. and Engineering Sciences*, 459(2039):2821–2845, 2003.

[Pas12] K. Pashkovich. *Extended formulations for combinatorial polytopes.* PhD thesis, Otto-von-Guericke-Universität, Magdeburg, 2012.

[Ren92] J. Renegar. On the computational complexity and geometry of the first-order theory of the reals. Part I: Introduction. preliminaries. The geometry of semi-algebraic sets. The decision problem for the existential theory of the reals. *Journal of Symbolic Computation*, 13(3):255 – 299, 1992.

[RG96] J. Richter-Gebert. *Realization Spaces of Polytopes.* Lecture Notes in Mathematics. Springer, 1996.

[Sch86] A. Schrijver. *Theory of Linear and Integer Programming.* John Wiley & Sons, Inc., New York, NY, USA, 1986.

[Sei97] R. Seidel. Convex hull computations. In *Handbook of Discrete and Computational Geometry*, pages 361–375. CRC Press, 1997.

[Sey94] P.D. Seymour. A note on hyperplane generation. *J. Comb. Theory Ser. B*, 61(1):88–91, 1994.

[Shi14] Y. Shitov. An upper bound for nonnegative rank. *J. Comb. Theory Ser. A*, 122:126–132, February 2014.

[SR34] E. Steinitz and H. Rademacher. *Vorlesungen über die Theorie der Polyeder.* Springer, 1934.

[Vav09] S. A. Vavasis. On the complexity of nonnegative matrix factorization. *SIAM Journal on Optimization*, 20(3):1364–1377, 2009.

[WSV00] H. Wolkowicz, R. Saigal, and L. Vandenberghe. *Handbook of Semidefinite Programming: Theory, Algorithms, and Applications.* International Series in Operations Research & Management Science. Springer US, 2000.

[Yan91] M. Yannakakis. Expressing combinatorial optimization problems by linear programs. *J. Comput. System Sci.*, 43(3):441–466, 1991.