Nearly positive matrices

Bryan Shader, Naomi Shaked-Monderer
and Daniel B. Szyld

Research Report 13-07-17
July 2013. Revised January 2014

Department of Mathematics
Temple University

This report is available in the World Wide Web at
`http://www.math.temple.edu/szyld`

Nearly positive matrices*

Bryan Shader† Naomi Shaked-Monderer‡ Daniel B. Szyld§

Abstract

Nearly positive matrices are nonnegative matrices which, when premultiplied by orthogonal matrices as close to the identity as one wishes, become positive. In other words, all columns of a nearly positive matrix are mapped simultaneously to the interior of the nonnegative cone by mutiplication by a sequence of orthogonal matrices converging to the identity. In this paper, nearly positive matrices are analyzed and characterized in several cases. Some necessary and some sufficient conditions for a nonnegative matrix to be nearly positive are presented. A connection to completely positive matrices is also presented.

Keywords:
Nonnegative matrices, Positive matrices, Completely positive matrices.

AMS subject classification: 15F10

1 Introduction

Consider the cone of nonnegative vectors in the m-dimensional space $\mathbb{R}^m_+ = \{\mathbf{x} \in \mathbb{R}^m, \mathbf{x} \geq \mathbf{0}\}$. Any nonzero vector $\mathbf{v}$ in its boundary is nonnegative, but not positive. It is not hard to see that an infinitesimal rotation in the appropriate direction can bring this vector into the interior of the cone $\mathbb{R}^m_+$. In other words, one can build a sequence of orthogonal matrices $Q(\ell)$ such that $\lim_{\ell\to\infty} Q(\ell) = I$ with the property that $Q(\ell)\mathbf{v} > \mathbf{0}$ for every ℓ.

For two non-orthogonal nonnegative vectors $\mathbf{u}$, $\mathbf{v}$, one can also build a sequence of orthogonal matrices, such that *both* $Q(\ell)\mathbf{u} > \mathbf{0}$ and $Q(\ell)\mathbf{v} > \mathbf{0}$ for every ℓ [7, Theore 6.12]. The existence of such a sequence was used in [7] to study topological properties of the set of matrices having a Perron-Frobenius property; see also [6].

*This version dated 6 January 2014.

†Department of Mathematics, University of Wyoming, 1000 E. University Avenue, Laramie, WY 82071-3036, USA (bshader.uwyo.edu).

‡Department of Economics and Management, The Max Stern Academic College of Yezreel Valley, Yezreel Valley, 19300 Israel (nomi@tx.technion.ac.il).

§Department of Mathematics, Temple University (038-16), 1805 N Broad Street, Philadelphia, PA 19122-6094, USA (szyld@temple.edu). Supported in part by the U.S. National Science Foundation under grant DMS-1115520.

Several natural questions arise from the above-mentioned results. The first of such questions which we address in this paper is: Can one build such a sequence to bring any set of *more than two* non-orthogonal vectors in the boundary of the nonnegative cone into its interior simultaneously? As we shall see, the answer is 'yes' for up to three vectors, but 'no' for four or more vectors. More specifically, let us call an $m \times n$ matrix A *nearly positive* provided there exists a sequence of orthogonal matrices $Q(\ell)$ such that

$$\lim_{\ell \to \infty} Q(\ell) = I \text{ and } Q(\ell)A > O \text{ for every } \ell,$$

where the last inequality is understood entrywise. Note that the condition $Q(\ell)A > \mathbf{0}$ is equivalent to $Q(\ell)\mathbf{a}_i > \mathbf{0}$ for every i, where $\mathbf{a}_1, \ldots, \mathbf{a}_n \in \mathbb{R}^m$ are the columns of A. In this paper we characterize nearly positive matrices, and study their properties. In particular, we present some necessary and some sufficient conditions for a nonnegative matrix to be nearly positive. A connection to completely positive matrices is also presented, and used to deduce certain results on nearly positive matrices.

We use the following notation: O denotes a zero matrix, I the identity matrix, and J a matrix of all ones. When we want to stress the order or size of the matrix we add it as a subscript; e.g. I_n stands for the $n \times n$ identity matrix and $O_{m\times n}$ is the $m \times n$ zero matrix. A vector of all ones is denoted by $\mathbf{1}$, and a zero vector is denoted by $\mathbf{0}$. The Hadamard (entrywise) product of two matrices A and B of the same order is denoted by $A \circ B$, and the direct sum of two matrices by $A \oplus B$. The inner product of two matrices of the same order is the Frobenius inner product $\langle A, B\rangle = \mathrm{trace}(AB^T)$. Whenever we consider a norm of a matrix, we mean the Frobenius norm $||A|| = \sqrt{\mathrm{trace}(AA^T)}$.

2 A Necessary Condition

In this section, we present a simple necessary condition, together with some sufficient conditions for a nonnegative matrix to be nearly positive.

We begin with a few simple observations.

Proposition 2.1 *Let A be an $m \times n$ nonnegative matrix.*

(a) *If P is an $m \times m$ permutation matrix, then A is nearly positive if and only if PA is nearly positive.*

(b) *If Q is an $n \times n$ permutation matrix, then A is nearly positive if and only if AQ is nearly positive.*

(c) *If D is an $n \times n$ diagonal matrix with a positive diagonal, then A is nearly positive if and only if AD is nearly positive.*

Proof. Let $Q(\ell)$ $(\ell = 1, 2, \ldots)$ be a sequence of orthogonal matrices. Part (a) is easy, since P is an orthogonal matrix, and one can consider the sequence $PQ(\ell)P^T$. We have that $Q(\ell)A > O$ for every ℓ and $\lim_{\ell\to\infty} Q(\ell) = I$ if and

only if $(PQ(\ell)P^T)PA > 0$ for every ℓ and $\lim_{\ell\to\infty} PQ(\ell)P^T = I$
For part (b), it suffices to note that each column of $Q(\ell)A$ is a positive vector if and only if any permutation of these columns is positive.
Part (c) is similar to (b), this time considering a scaling of the columns of A. ■

It is natural to ask whether a statement similar to (c) of Proposition 2.1 exists for left multiplication of A by a diagonal matrix with a positive diagonal. This question is less trivial, and we will consider it in the final section.

The next proposition contains another basic observation: If A is "nearly-nearly positive" then A is nearly positive:

Proposition 2.2 *Let A be an $m \times n$ nonnegative matrix. If there exists a sequence $U(\ell)$ of orthogonal matrices such that*

$$\lim_{\ell\to\infty} U(\ell) = I \text{ and } U(\ell)A \text{ is nearly positive for every } \ell,$$

then A is nearly positive.

Proof. For every ℓ let $P_\ell(i)$ be a sequence of orthogonal matrices such that $\lim_{i\to\infty} P_\ell(i) = I$ and $P_\ell(i)U(\ell)A > O$ for every i. Then $P_\ell(\ell)U(\ell)$ $(\ell = 1, 2, \ldots)$ is a sequence of orthogonal matrices converging to I such that $P_\ell(\ell)U(\ell)A > 0$ for every ℓ. ■

Next, we present a simple necessary condition for being nearly positive.

Proposition 2.3 *Let A be a nonnegative matrix. If $A \geq O$ is nearly positive, then $A^TA > O$.*

Proof. Let Q be an orthogonal matrix such that $QA > O$, then $A^TA = A^TQ^TQA = (QA)^T(QA) > O$. ■

Observe that the nonnegative matrix A satisfies $A^TA > O$ if and only if each column of A is nonzero and no pair of columns of A are orthogonal. Geometrically, one can see that if two nonnegative vectors are orthogonal, no single orthogonal matrix can bring both vectors to the interior of the nonnegative cone simultaneously.

One of the questions we ask here is: in which cases is this necessary condition also sufficient? To that end, we present next a simple sufficient condition.

Proposition 2.4 *Let $A \geq O$ have a positive row. Then A is nearly positive.*

Proof. Without loss of generality A is an $m \times n$ matrix whose first row is positive. For $j = 2, \ldots, m$ and $\theta > 0$, let $R_j(\theta)$ be the Givens rotation by θ radians involving coordinates 1 and j; i.e. $R_j(\theta)$ fixes the coordinates axes in $\mathbb{R}^m$ other than 1 and j and acts as the rotation

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

on the plane containing the 1st and jth coordinate axes. Then for θ sufficiently small

$$R_2(\theta)R_3(\theta)\cdots R_m(\theta)A > O.$$

It follows that A is nearly positive. ■

Corollary 2.5 *Let A be an $m \times n$ nearly positive matrix, and let C be a $k \times n$ nonnegative matrix. Then $B = \begin{bmatrix} A \\ C \end{bmatrix}$ is nearly positive.*

Proof. If $Q(\ell)A > O$ and $Q(\ell)$ converges to I, then $U(\ell) = Q(\ell) \oplus I_k$ converges to I and $U(\ell)B$ is nearly positive by Proposition 2.4. Thus B is nearly positive by Proposition 2.2. ■

Proposition 2.6 *Let A be an $m \times n$ nonnegative matrix, and let C be an $m \times k$ positive matrix. Then $B = \left[\begin{array}{c|c} A & C \end{array}\right]$ is nearly positive if and only if A is nearly positive.*

Proof. If Q is an orthogonal matrix which is close enough to the identity, then $QC > O$. ■

Remark 2.7 *If A is not nearly positive, then no matrix obtained from A by appending nonnegative columns is nearly positive.*

Combining Propositions 2.3 and 2.4, we fully characterize below the $m \times n$ matrices with $\min m, n \leq 2$ are nearly positive. For $n = 2$ the result is the same as [7, Theorem 6.12], but is set here in a different context and with a cleaner simpler proof. As we shall see, a similar result holds for $m \times 3$ matrices, but this is postponed until section 5. We show later in Example 4.5 in section 4 that the same result does not hold for $n \geq 4$. We also discuss the case $n \geq 5$ in section 7.

Theorem 2.8 *Let A be an $m \times n$ nonnegative matrix, where $\min(m, n) \leq 2$. Then A is nearly positive if and only if $A^T A > O$.*

Proof. The only if part is given by Proposition 2.3.

Conversely assume that $A^T A > O$. If $n \leq 2$, then Proposition 2.4 implies that A is nearly positive. If $m = 1$, then $A^T A > O$ implies that $A > O$ and hence A is nearly positive. If $m = 2$, then since no two columns are perpendicular, a zero in one row of A implies the other row is positive and hence by Proposition 2.4, A is nearly positive. ■

In section 4 we shall see an example of a $3 \times n$ nonnegative matrix A, which satisfies the condition $A^T A > O$, but is not nearly positive; see Example 4.3. Therefore, for $m \times n$ nonnegative matrices A, the necessary condition $A^T A > O$ is sufficient for $m = 1, 2$, but not for $m \geq 3$.

3 A Sufficient Condition

The next theorem gives a sufficient condition for a nonnegative matrix to be nearly positive, and depends on the following characterization of orthogonal matrices with no eigenvalue equal to -1; see, e.g., [9, Ch. IX, § 14]. Note that continuity of eigenvalues implies no orthogonal matrix which is sufficiently close to I has eigenvalue -1.

Proposition 3.1 (The Cayley transform) *The set of $n \times n$ orthogonal matrices with no eigenvalue equal to -1 is the same as*

$$\{(I+K)^{-1}(I-K) : K \text{ is an } n \times n \text{ skew-symmetric matrix}\}. \tag{3.1}$$

Theorem 3.2 *Let $A = [a_{ij}]$ be a nonnegative $m \times n$ matrix and K be a skew-symmetric matrix such that*

$$(KA)_{ij} > 0 \text{ for all } (i,j) \text{ with } a_{ij} = 0.$$

Then A is nearly positive.

Proof. For each positive integer ℓ, define $Q(\ell)$ by the Cayley formula (3.1), using the skew-symmetric matrix $-\frac{1}{\ell}K$. That is, set

$$Q(\ell) = (I - \frac{1}{\ell}K)^{-1}(I + \frac{1}{\ell}K).$$

For ℓ sufficiently large, $||\frac{1}{\ell}K|| < 1$ (where $||\cdot||$ denotes the Frobenius norm), and thus

$$\begin{aligned} Q(\ell) &= (I + \frac{1}{\ell}K + \frac{1}{\ell^2}K^2 + \frac{1}{\ell^3}K^3 + \cdots)(I + \frac{1}{\ell}K) \\ &= I + \frac{2}{\ell}K + \frac{2}{\ell^2}K^2 + \frac{2}{\ell^3}K^3 + \cdots \quad . \end{aligned}$$

For ℓ large enough so that $\frac{1}{\ell}||K||$ is bounded away from 1, we get that for every $1 \le p \le m$ and $1 \le q \le n$

$$\left|\left(Q(\ell)A - (I + \frac{2}{\ell}K)A\right)_{pq}\right| \le \sum_{j=2}^{\infty} \frac{2}{\ell^j}||K||^j||A|| < \gamma\frac{1}{\ell}$$

for some positive constant γ. Thus for ℓ sufficiently large the entries of $Q(\ell)A$ are positive (negative) when those of $(I+\frac{2}{\ell}K)A = A+\frac{2}{\ell}KA$ are positive (negative). The assumptions on KA imply that $A+\frac{2}{\ell}KA > O$ for ℓ sufficiently large. Hence A is nearly positive. ■

We illustrate Theorem 3.2 with several examples.

Example 3.3 *Let A be an $m \times n$ nonnegative matrix with positive first row, and*

$$K = \left[\begin{array}{c|c} 0 & -\mathbf{1}^T \\ \hline \mathbf{1} & O_{(m-1)\times(m-1)} \end{array}\right].$$

Each row of KA other that the first is positive. Hence, by Theorem 3.2, A is nearly positive.

Example 3.3 gives a different proof of Proposition 2.4.

Example 3.4 *Let*

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

For the skew-symmetric matrix

$$K = \begin{bmatrix} 0 & -5 & 3 \\ 5 & 0 & -4 \\ -3 & 4 & 0 \end{bmatrix},$$

the $(1,3)$*,* $(2,2)$ *and* $(3,1)$ *entries of* KB *are positive. Thus Theorem 3.2 implies that* B *is nearly positive.*

Example 3.5 *Let*

$$C = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}.$$

For an arbitrary 3×3*, skew-symmetric matrix*

$$K = \begin{bmatrix} 0 & r & -s \\ -r & 0 & t \\ s & -t & 0 \end{bmatrix},$$

if each diagonal entry of KC *is positive, then* $r > s$*,* $t > r$ *and* $s > t$*, which cannot occur. Hence the hypotheses of Theorem 3.2 are not satisfied.*

However, the matrix C *is in fact nearly positive, as small, nonzero rotations about the axis containing* $(1,1,1)$ *make each column of* C *positive (see also Example 4.6). Thus, we see that the converse of Theorem 3.2 is not true.*

The following theorem shows that something slightly weaker than the converse of Theorem 3.2 is true.

Theorem 3.6 *Let* $A \geq O$ *be a nearly positive matrix. Then there exists a nonzero skew-symmetric matrix* K *such that* $(KA)_{ij} \geq 0$ *for all* (i,j) *such that* $a_{ij} = 0$*.*

Proof. If $A > O$, then there is nothing to show.

Assume that A is not positive. Let $Q(\ell)$ $(\ell = 1, 2, \ldots)$ be a sequence of orthogonal matrices such that $\lim_{\ell\to\infty} Q(\ell) = I$ and $Q(\ell)A > O$ for all ℓ. As A is not positive, $Q(\ell) \neq I$.

For ℓ sufficiently large, $Q(\ell)$ does not have eigenvalue -1. Hence, by Proposition 3.1, for ℓ sufficiently large there exists a skew-symmetric matrix L_ℓ and a positive real number ϵ_ℓ such that $\|L_\ell\| = 1$ and $Q(\ell) = (I + \epsilon_\ell L_\ell)^{-1}(I - \epsilon_\ell L_\ell)$ (here too the norm stands for the Frobenius norm). Furthermore, by compactness, we may assume without loss of generality that $\lim_{\ell\to\infty} L_\ell$ exists and is equal to some skew-symmetric matrix L of norm 1.

Note that

$$Q(\ell) = I - 2\epsilon_\ell L_\ell + 2\epsilon_\ell^2 L_\ell^2 - 2\epsilon_\ell^3 L_\ell^3 + \cdots \quad .$$

Thus

$$\frac{Q(\ell)A - A}{\epsilon_\ell} + 2LA = -2(L_\ell - L)A + 2\epsilon_\ell L_\ell^2 A - 2\epsilon_\ell^2 L_\ell^3 A + \cdots \quad . \tag{3.2}$$

Since $L_\ell \to L$ and $\epsilon_\ell \to 0$ as $\ell \to \infty$, (3.2) implies that

$$\lim_{\ell \to \infty} \frac{Q(\ell)A - A}{\epsilon_\ell} = -2LA.$$

As $(Q(\ell)A - A)_{ij} > 0$ for all (i,j) such that $a_{ij} = 0$, we conclude that $(-LA)_{ij} \geq 0$ for all (i,j) such that $a_{ij} = 0$. We may now take $K = -L$. ■

We next give equivalent conditions to those in Theorem 3.2 that are sometimes simpler to apply.

Theorem 3.7 *Let $A = [a_{ij}]$ be an $m \times n$ nonnegative matrix. Set*

$$\mathcal{Z} = \{(i,j) : a_{ij} = 0\}.$$

Then exactly one of the following occurs:

(a) *There exists a nonzero skew-symmetric matrix K such that*

$$(KA)_{ij} > 0 \text{ for all } (i,j) \in \mathcal{Z}$$

(b) *There exists an nonzero, nonnegative $m \times n$ matrix $Y = [y_{ij}]$ such that $A \circ Y = O$, and AY^T is symmetric.*

Proof. Let $K = [k_{ij}]$ be an $n \times n$ skew-symmetric matrix whose entries above the main diagonal are distinct indeterminates. Consider the $|\mathcal{Z}| \times \binom{n}{2}$ matrix M whose rows are indexed by elements of $\mathcal{Z}$ and whose columns are indexed by $E = \{(i,j) : 1 \leq i < j \leq n\}$, such that the entry corresponding to row (u,v) and column (i,j) is the coefficient of k_{ij} in $(KA)_{uv}$, i.e.,

$$m_{(u,v),(i,j)} = \begin{cases} -a_{iv}, & \text{if } i < j = u \\ a_{jv}, & \text{if } i = u < j \\ 0, & \text{otherwise} \end{cases} .$$

For every $n \times n$ skew-symmetric matrix X let $\mathbf{x}$ be the vector of its entries above the diagonal x_{ij}, $(i,j) \in E$. Then for every (u,v) in $\mathcal{Z}$, $(XA)_{uv}$ is equal to the (u,v)th entry of $M\mathbf{x}$. Condition (a) is therefore equivalent to the statement: There exists a nonzero $\mathbf{x} \in \mathbb{R}^{\binom{n}{2}}$ such that $M\mathbf{x} > \mathbf{0}$.

With every $\mathbf{y} \in \mathbb{R}^{|\mathcal{Z}|}$ we associate an $m \times n$ matrix Y such that $y_{ij} = 0$ if $(i,j) \notin \mathcal{Z}$, and y_{ij} is the entry of $\mathbf{y}$ corresponding to the index $(i,j) \in \mathcal{Z}$ otherwise. Then $A \circ Y = O$ and the matrix AY^T is symmetric if and only if AY^T is in the orthogonal complement (with respect to the Frobenius inner product) of the $n \times n$ skew-symmetric matrices, i.e., if and only if $\text{trace}(KAY^T) = 0$ for every skew-symmetric K. But

$$\begin{aligned}
\text{trace}(KAY^T) = \text{trace}(Y^T KA) &= \sum_{i,j} y_{ij}(KA)_{ij} \\
&= \sum_{(i,j)\in\mathcal{Z}} y_{ij}(KA)_{ij} \\
&= \sum_{(i,j)\in\mathcal{Z}} y_{ij}(M\mathbf{k})_{(i,j)} \\
&= \sum_{(i,j)\in\mathcal{Z}} y_{ij} \sum_{(u,v)\in E} m_{(i,j),(u,v)} k_{uv} \\
&= \sum_{(u,v)\in E} \left(\sum_{(i,j)\in\mathcal{Z}} y_{ij} m_{(i,j),(u,v)} \right) k_{uv}.
\end{aligned}$$

That is, $\text{trace}(KAY^T) = 0$ for every skew-symmetric K is equivalent to $\mathbf{y}^T M = \mathbf{0}^T$. Condition (b) is therefore equivalent to the statement: There exists a nonzero nonnegative $\mathbf{y} \in \mathbb{R}^{|\mathcal{Z}|}$ such that $\mathbf{y}^T M = \mathbf{0}^T$.

Let $P = \{\mathbf{x} \in \mathbb{R}^{|\mathcal{Z}|} : \mathbf{x} > 0\}$, and $\mathcal{R}(M)$ be the column space or range of M. Note that P is an open convex set and $\mathcal{R}(M)$ is a subspace.

Then condition (a) is equivalent to the statement $P \cap \mathcal{R}(M) \neq \emptyset$. By the separating hyperplane theorem (see, e.g., [11, Theorem 11.2]), $P \cap \mathcal{R}(M) = \emptyset$ if and only if there exists $\mathbf{y} \in \mathbb{R}^{|\mathcal{Z}|}$ such that $\mathbf{y}^T\mathbf{x} > 0$ for every $\mathbf{x} \in P$ and $\mathbf{y}$ is orthogonal to $\mathcal{R}(M)$. That is, $P \cap \mathcal{R}(M) = \emptyset$ if and only if there exists a nonnegative nonzero $\mathbf{y}$ such that $\mathbf{y}^T M = \mathbf{0}^T$, which is equivalent to condition (b). ■

We point out that Theorem 3.7 can also be shown using Farkas' Lemma [8, 1].

Note that by the nonnegativity of A and Y in Theorem 3.7(b), $A \circ Y = O$ is equivalent to $\text{trace}(AY^T) = 0$, or to the diagonal of AY^T being zero. That is, $y_{ij} = 0$ for every $(i,j) \notin \mathcal{Z}$.

Combining Theorems 3.2 and 3.7 we have the following sufficient condition.

Corollary 3.8 *Let A be an $m \times n$ nonnegative matrix. If $Y = O$ is the only $m \times n$ nonnegative matrix such that AY^T is symmetric with zero diagonal, then A is nearly positive.*

Both Example 3.4 and Example 3.4 could be treated via Corollary 3.8:

Example 3.9 *Let B be as in Example 3.4. We now use Corollary 3.8 to show that B is nearly positive. A nonnegative matrix Y such that each diagonal entry of BY^T is 0 has the form*

$$Y^T = \begin{bmatrix} 0 & 0 & c \\ 0 & b & 0 \\ a & 0 & 0 \end{bmatrix}.$$

In addition, BY^T is symmetric only if $a = b$, $c = 2a$ and $b = c$. This requires $Y = O$. Hence, by Corollary 3.8, B is nearly positive.

Example 3.10 *Let C be as in Example 3.5. Then for $Y = I$ the matrix $CY^T = C$ is symmetric, and therefore there is no nonzero skew-symmetric matrix K such that the diagonal entries of KC are all positive.*

4 More Necessary Conditions

We now turn our attention to establishing additional necessary conditions for a matrix to be nearly positive. Let A be a nonnegative matrix. Theorem 3.7 provides a sufficient condition for A to be nearly positive; namely, the non-existence of a nonzero nonnegative matrix $Y = [y_{ij}]$ such that $y_{ij} > 0$ only if $a_{ij} = 0$, and AY^T is symmetric. However, the existence of such a Y does not determine whether or not A is nearly positive (see Examples 3.5 and 4.4).

The following theorem gives a necessary condition for A to be nearly positive, in the case that such a Y exists with the additional constraint that $y_{ij} > 0$ if and only if $a_{ij} = 0$.

Theorem 4.1 *Let A be a nearly positive $m \times n$ matrix for which there exists a nonnegative $m \times n$ matrix $Y = [y_{ij}]$ such that $y_{ij} > 0$ if and only if $a_{ij} = 0$ and AY^T is symmetric. Then there exists a skew-symmetric matrix K with $(KA)_{ij} \geq 0$ for all (i, j) with $a_{ij} = 0$, and each such K satisfies $(KA)_{ij} = 0$ for all (i, j) with $a_{ij} = 0$.*

Proof. By Theorem 3.6 there exists a nonzero skew-symmetric matrix K with $(KA)_{ij} \geq 0$ for all (i, j) with $a_{ij} = 0$. Let K be any such matrix, and let M be as in the proof of Theorem 3.7. The matrix Y corresponds to a positive vector $\mathbf{y}$ such that $\mathbf{y}^T M = 0$. The matrix K corresponds to a nonzero vector $\mathbf{k}$ such that $M\mathbf{k} \geq \mathbf{0}$. Then $\mathbf{y}^T M\mathbf{k} = 0$, which requires $M\mathbf{k} = \mathbf{0}$. This translates to the condition that $(KA)_{ij} = 0$ for all (i, j) with $a_{ij} = 0$. ■

We use this corollary in the next example.

Example 4.2 *Consider*

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

The nonnegative matrix

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

has the properties that AY^T is symmetric with zero trace and $y_{ij} > 0$ if and only if $a_{ij} = 0$. Hence, by Theorem 4.1, either A is not nearly positive or there exists a nonzero skew-symmetric matrix K such that $(KA)_{ij} = 0$ for all

$(i,j) \in \{(1,1),(2,2),(3,3),(3,4)\}$. *We claim that the latter cannot occur. To see this, let*

$$K = \begin{bmatrix} 0 & r & -s \\ -r & 0 & t \\ s & -t & 0 \end{bmatrix}.$$

The conditions on KA *would require* $r-s=0$, $-r+t=0$, $s-2t=0$ *and* $s-(1/2)t=0$. *Evidently, these require* $K=O$.

Therefore, by Theorem 4.1, A *is not nearly positive (although each* 3×3 *submatrix of* A *is nearly positive, by Proposition 2.4, Example 3.4 and Proposition 2.1. See also Theorem 5.6 in the next section).*

Note that for A of the last example $A^TA > O$ holds. Thus this example shows that one can have a 3×4 nonnegative matrix with mutually non-orthogonal columns which is not nearly positive.

Example 4.3 *Let* A *be as in Example 4.2. By Remark 2.7 any nonnegative* $3\times n$ *matrix* $\hat{A} = [\ A \mid E\ ]$ *is not nearly positive.*

This example shows that for $n\geq 4$, having mutually non-orthogonal columns is not a sufficient condition for an $3\times n$ nonnegative matrix to be nearly positive. The same statement holds for square nonnegative matrices with more than three rows, as the following example shows.

Example 4.4 *Let* $n\geq 4$.

$$A = \left[\begin{array}{c|c} 0 & \mathbf{1}^T \\ \hline \mathbf{1} & I_{n-1} \end{array}\right].$$

The matrix

$$Y = \left[\begin{array}{c|c} n-2 & \mathbf{0}^T \\ \hline \mathbf{0} & J_{n-1}-I_{n-1} \end{array}\right]$$

satisfies AY^T *is symmetric with zero trace and* $y_{ij}>0$ *if and only if* $a_{ij}=0$. *Suppose*

$$K = \left[\begin{array}{c|c} 0 & \mathbf{u}^T \\ \hline -\mathbf{u} & L \end{array}\right]$$

is a skew-symmetric matrix such that $(KA)_{ij}=0$ *whenever* $a_{ij}=0$. *Then* $l_{ij}=-u_i$ *and* $l_{ji}=-u_j$ *for all* $i\neq j$. *As* L *is skew-symmetric, this implies that* $u_i=-u_j$ *for* $1\leq i\neq j\leq n-1$. *Since* $n\geq 4$, *this requires* $\mathbf{u}^T=[0,0,\ldots,0]$ *and thus* $L=O$. *Hence* $K=O$, *and by Theorem 4.1,* A *is not nearly positive.*

We can use the matrix of Example 4.4 to generate matrices with more columns that are not nearly positive, as in Example 4.3:

Example 4.5 *For* $m\geq 4$ *let* A *be the* $m\times m$ *matrix of the previous example. By Remark 2.7, any nonnegative* $m\times n$ *matrix* $\hat{A}=[\ A\mid E\ ]$ *is not nearly positive.*

Example 4.5 shows that for every $n \geq m \geq 4$ there exists a nonnegative $m \times n$ matrix with mutually non-orthogonal columns which is not nearly positive.

In the final example of this section we go back to Example 3.5, and use our results to prove that the matrix D defined there is nearly positive.

Example 4.6 *Let C be the 3×3 matrix in Example 3.5. Then $Y = I$ satisfies CY^T is symmetric with trace zero, and $y_{ij} > 0$ if and only if $c_{ij} = 0$. Since the skew-symmetric matrix*

$$K = \begin{bmatrix} 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \end{bmatrix}$$

has the property that $(KC)_{ij} = 0$ for all (i, j) such that $c_{ij} = 0$, the second conclusion in Theorem 4.1 is satisfied. This is not enough to conclude that C is nearly positive. However note that

$$K^2C = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}.$$

Let $Q_\epsilon = (I + \epsilon K)^{-1}(I - \epsilon K)$. It follows, by the series expansion of $Q_\epsilon C$, that $Q_\epsilon C > O$ for for ϵ sufficiently small and positive, and hence that C is nearly positive.

5 The $3 \times n$ and $m \times 3$ nearly positive matrices

In this section we determine which matrices with three rows or three columns are nearly positive. As we saw earlier, not all nonnegative matrices with three rows (with mutually non-orthogonal columns) are nearly positive, but we characterize here those that are. On the other hand, we show that all nonnegative matrices with three mutually non-orthogonal columns are nearly positive. In other words, we show that for any three mutually non-orthogonal nonnegative vectors, there exists a sequence of unitary matrices converging to the identity matrix that simultaneously map the three vectors to the interior of the nonnegative orthant.

We begin by studying a certain family of $3 \times n$ matrices, namely those of the form

$$A = \left[\begin{array}{c|c|c} \mathbf{0}^T & \mathbf{y}^T & \mathbf{1}^T \\ \hline \mathbf{1}^T & \mathbf{0}^T & \mathbf{z}^T \\ \hline \mathbf{x}^T & \mathbf{1}^T & \mathbf{0}^T \end{array}\right], \tag{5.1}$$

where $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{z}$ are positive $p_x \times 1$, $p_y \times 1$ and $p_z \times 1$ vectors, respectively.

To study whether or not the matrix A in (5.1) is nearly positive we consider a nonnegative matrix of the form

$$Y = \left[\begin{array}{c|c|c} \mathbf{u}^T & \mathbf{0}^T & \mathbf{0}^T \\ \hline \mathbf{0}^T & \mathbf{v}^T & \mathbf{0}^T \\ \hline \mathbf{0}^T & \mathbf{0}^T & \mathbf{w}^T \end{array}\right],$$

where $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ are $p_x \times 1$, $p_y 1$ and $p_z \times 1$ vectors, respectively.

Evidently AY^T is symmetric if and only if

$$\begin{array}{rcl} \mathbf{v}^T\mathbf{y} & = & \mathbf{u}^T\mathbf{1} \\ \mathbf{w}^T\mathbf{z} & = & \mathbf{v}^T\mathbf{1} \\ \mathbf{u}^T\mathbf{x} & = & \mathbf{w}^T\mathbf{1}. \end{array} \tag{5.2}$$

Suppose that (5.2) holds with $\mathbf{v}^T$, $\mathbf{w}^T$ and $\mathbf{u}^T$ nonnegative. Set $\bar{x}$ and $\underline{x}$ to be the largest and smallest entry of $\mathbf{x}$, respectively. Define $\bar{y}$, $\underline{y}$, $\bar{z}$ and $\underline{z}$ analogously. Then

$$\begin{array}{rcccl} \bar{y}\mathbf{v}^T\mathbf{1} & \geq & \mathbf{u}^T\mathbf{1} & \geq & \underline{y}\mathbf{v}^T\mathbf{1}, \\ \bar{z}\mathbf{w}^T\mathbf{1} & \geq & \mathbf{v}^T\mathbf{1} & \geq & \underline{z}\mathbf{w}^T\mathbf{1}, \\ \bar{x}\mathbf{u}^T\mathbf{1} & \geq & \mathbf{w}^T\mathbf{1} & \geq & \underline{x}\mathbf{u}^T\mathbf{1}. \end{array} \tag{5.3}$$

It follows that either each of $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ is a zero vector or

$$\bar{x}\bar{y}\bar{z} \geq 1 \geq \underline{x}\underline{y}\underline{z}. \tag{5.4}$$

Corollary 3.8 now implies the following.

Proposition 5.1 *Let A be a $3 \times n$ nonnegative matrix of the form* (5.1)*. If*

$$1 > \bar{x}\bar{y}\bar{z} \textit{ or } \underline{x}\underline{y}\underline{z} > 1,$$

then A is nearly positive.

We now turn our attention to matrices of the form (5.1) for which (5.4) holds. First, consider the case $\bar{x}\bar{y}\bar{z} > 1 > \underline{x}\underline{y}\underline{z}$. We claim that there exist positive $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ satisfying (5.2). To see this, note that in this case, there exist positive numbers a, b, c such that

$$\begin{array}{lllllllllll} a = \underline{y}b & \text{if} & \bar{y} = \underline{y}, & \text{and} & \bar{y}b & > & a & > & \underline{y}b & \text{otherwise} \\ b = \underline{z}c & \text{if} & \bar{z} = \underline{z}, & \text{and} & \bar{z}c & > & b & > & \underline{z}c & \text{otherwise} \\ c = \underline{x}a & \text{if} & \bar{x} = \underline{x}, & \text{and} & \bar{x}a & > & c & > & \underline{x}a & \text{otherwise.} \end{array} \tag{5.5}$$

The existence of a, b, c satisfying (5.5) follows from the fact that there exist positive α, β, γ such that

$$\begin{array}{lllllllllll} \alpha = \underline{y} & \text{if} & \bar{y} = \underline{y}, & \text{and} & \bar{y} & > & \alpha & > & \underline{y} & \text{otherwise} \\ \beta = \underline{z} & \text{if} & \bar{z} = \underline{z}, & \text{and} & \bar{z} & > & \beta & > & \underline{z} & \text{otherwise} \\ \gamma = \underline{x} & \text{if} & \bar{x} = \underline{x}, & \text{and} & \bar{x} & > & \gamma & > & \underline{x} & \text{otherwise.} \end{array}$$

and $\alpha\beta\gamma = 1$. Take, e.g., $b = 1$, $a = \alpha$, $c = \alpha\gamma$. Thus, it suffices to show that there exist positive $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ such that $\mathbf{u}^T\mathbf{1} = a$ and $\mathbf{u}^T\mathbf{x} = c$, $\mathbf{v}^T\mathbf{1} = b$ and $\mathbf{v}^T\mathbf{y} = a$, $\mathbf{w}^T\mathbf{1} = c$ and $\mathbf{w}^T\mathbf{z} = b$. This is shown by using the following lemma.

Lemma 5.2 *Let $\mathbf{y}$ be a positive vector and $\mathbf{1}$ be the vector of all ones in $\mathbb{R}^n$, a and b positive numbers. If either $\bar{y} > \underline{y}$ and $\bar{y}b > a > \underline{y}b$, or $\bar{y} = \underline{y}$ and $a = \underline{y}b$, then there exists a positive vector $\mathbf{v} \in \mathbb{R}^n$ such that $\mathbf{1}^T\mathbf{v} = b$ and $\mathbf{y}^T\mathbf{v} = a$.*

Proof. If $\bar{y} = \underline{y}$ and $a = \underline{y}b$, then $\mathbf{y} = \underline{y}\mathbf{1}$, and $\mathbf{v}$ satisfies $\mathbf{1}^T\mathbf{v} = b$ if and only if $\mathbf{y}^T\mathbf{v} = a$. Obviously, there exists a positive vector $\mathbf{v}$ such that $\mathbf{1}^T\mathbf{v} = b$. (Note that this case includes the case that $n = 1$.)

We prove the case that $\bar{y} > \underline{y}$ and $\bar{y}b > a > \underline{y}b$ by induction on n: If $n = 2$ then the system of equations in $\mathbf{v}$

$$\begin{cases} \mathbf{1}^T\mathbf{v} &= b \\ \mathbf{y}^T\mathbf{v} &= a \end{cases}$$

has the solution $\mathbf{v} = [v_1 \ v_2]^T$, with

$$v_1 = \frac{a - y_2 b}{y_1 - y_2}, \quad v_2 = \frac{y_1 b - a}{y_1 - y_2},$$

both of which are positive.

For $n \geq 3$, assuming the claim holds for $n - 1$. We may assume $\mathbf{y}^T = [\underline{y} \ \bar{y} \ \ldots \ y_n]$. Let $\epsilon > 0$ be small enough, so that

$$(b - \epsilon)\bar{y} > a - \epsilon y_n > (b - \epsilon)\underline{y}.$$

Let $v_n = \epsilon$. Let $\tilde{\mathbf{y}}$ be the vector $\mathbf{y}$ with the nth entry omitted. By the induction hypothesis there exists a positive $\tilde{\mathbf{v}} \in \mathbb{R}^{n-1}$ such that $\mathbf{1}^T\tilde{\mathbf{v}} = b - \epsilon$ and $\tilde{\mathbf{y}}^T\tilde{\mathbf{v}} = a - \epsilon y_n$. The desired positive vector $\mathbf{v}$ is obtained by appending v_n to $\tilde{\mathbf{v}}$. ■

It is easy to verify that if $\bar{x}\bar{y}\bar{z} > 1 > \underline{x}\underline{y}\underline{z}$, then a skew-symmetric matrix K such that $(KA)_{ij} = 0$ whenever $a_{ij} = 0$ is necessarily the zero matrix. Indeed, in this case at least one of $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{z}$ is not constant. Suppose, for example, that $\bar{x} > \underline{x}$. If

$$K = \begin{bmatrix} 0 & -r & s \\ r & 0 & -t \\ -s & t & 0 \end{bmatrix},$$

then the fact that entries $(1,1)$ and $(1,2)$ of

$$K \begin{bmatrix} 0 & 0 & \bar{y} \\ 1 & 1 & 0 \\ \bar{x} & \underline{x} & 1 \end{bmatrix}$$

are zero implies that $\bar{x}s = \underline{x}s = r$, and thus $s = 0$ and $r = 0$. This, together with the fact that entry $(2,3)$ of the product is also zero, implies that also $t = 0$. Thus, $K = O$. By Theorem 4.1 we have proven the following result.

Proposition 5.3 *If A is a matrix of the form* (5.1) *for which*

$$\bar{x}\bar{y}\bar{z} > 1 > \underline{x}\underline{y}\underline{z},$$

then A is not nearly positive.

Next we consider the case that $1 = \underline{x}\underline{y}\underline{z}$. Let

$$K = \begin{bmatrix} 0 & -1 & \underline{y}\underline{z} \\ 1 & 0 & -\underline{y} \\ -\underline{y}\underline{z} & \underline{y} & 0 \end{bmatrix}$$

and let

$$B = \begin{bmatrix} 0 & \beta & 1 \\ 1 & 0 & \gamma \\ \alpha & 1 & 0 \end{bmatrix},$$

where $\underline{x} \le \alpha \le \bar{x}$, $\underline{y} \le \beta \le \bar{y}$ and $\underline{z} \le \gamma \le \bar{z}$. Then

$$KB = \begin{bmatrix} \alpha\underline{y}\underline{z} - 1 & \underline{y}\underline{z} & -\gamma \\ -\alpha\underline{y} & \beta - \underline{y} & 1 \\ \underline{y} & -\beta\underline{y}\underline{z} & \underline{y}(\gamma - \underline{z}) \end{bmatrix}$$

and therefore each diagonal entry of KB is nonnegative.

It is easy to verify that each diagonal entry of K^2B is positive. Let $L = -K$. Then the matrix $(I - 2\epsilon L + 2\epsilon^2 L^2)B$ is positive for $\epsilon > 0$ sufficiently small, and thus we obtain that $((I+\epsilon L)^{-1}(I-\epsilon L))B$ is positive for $\epsilon > 0$ sufficiently small. Since every column of A of the form (5.1) is in the form of one of the columns of B, it follows that $((I+\epsilon L)^{-1}(I-\epsilon L))A$ is positive for ϵ sufficiently small, and hence that A is nearly positive.

A similar argument for the case $1 = \bar{x}\bar{y}\bar{z}$, using the same B and

$$K = \begin{bmatrix} 0 & 1 & -\bar{y}\bar{z} \\ -1 & 0 & \bar{y} \\ \bar{y}\bar{z} & -\bar{y} & 0 \end{bmatrix},$$

shows that such matrices A are nearly positive.

Thus, Propositions 5.1 and 5.3, and the above argument give the following characterization of the family of $3 \times n$ nearly positive matrices of the form (5.1).

Theorem 5.4 *If A is a $3 \times n$ nonnegatvie matrix of the form* (5.1)*, then A is nearly positive if and only if $\bar{x}\bar{y}\bar{z} \le 1$ or $\underline{x}\underline{y}\underline{z} \ge 1$.*

We now study $3 \times n$ general nonnegative matrices. Without loss of generality, we may assume that such matrix has the form

$$B = \begin{bmatrix} P & A & S & \Omega \end{bmatrix}, \tag{5.6}$$

where P is a positive matrix, A is a matrix of the form (5.1), each column of S has exactly one nonzero entry, and Ω is a zero matrix. Some of these submatrices may possibly be vacuous, and A may not have columns of all three types (that is, with all three zero-nonzero patterns). In other words, every column of B has either three, two, one, or no positive entry, and from Proposition 2.1 the columns with two positive entries can be scaled to have the form (5.1), and all columns may be permuted arbitrarily.

Theorem 5.5 *Let B be a $3 \times n$ matrix of the form* (5.6)*. Then B is nearly positive if and only if the following holds:*
(i) *Ω is vacuous and*
(ii) *either* (a) *B has a positive row, or*
(b) *S is vacuous, A contains at least one column of each zero-nonzero pattern, and either $\bar{x}\bar{y}\bar{z} \leq 1$ or $\underline{x}\underline{y}\underline{z} \geq 1$.*

Proof. First suppose that B is nearly positive. Since $B^T B > O$, the matrix Ω is vacuous. Also, $B^T B > O$ implies that if there is a column with exactly one positive entry, then B has a positive row. Thus if B does not have a positive row, S is vacuous, and A has at least one column of each type. By Proposition 2.6, A is nearly positive and hence $\bar{x}\bar{y}\bar{z} \leq 1$ or $\underline{x}\underline{y}\underline{z} \geq 1$, by Theorem 5.4.

Conversely, suppose that Ω is vacuous. If B has a positive row, then B is nearly positive by Proposition 2.4. If S is vacuous, A has columns of all three types, and either $\bar{x}\bar{y}\bar{z} \leq 1$ or $\underline{x}\underline{y}\underline{z} \geq 1$, then by Theorem 5.4, A is nearly positive, and by Proposition 2.6, B is also nearly positive. ∎

We conclude this section by showing that the necessary condition that $A^T A > O$ is also sufficient for $m \times 3$ matrices.

Theorem 5.6 *Let A be an $m \times 3$ nonnegative matrix. Then A is nearly positive if and only if $A^T A > O$.*

Proof. By Proposition 2.3, if A is nearly positive, then $A^T A > O$.

Conversely, suppose that $A^T A > O$. If A has a positive row, then, by Proposition 2.4, A is nearly positive. Otherwise, since $A^T A > O$, A contains (up to column permuatation) a submatrix of the form

$$\begin{bmatrix} 0 & + & + \\ + & 0 & + \\ + & + & 0 \end{bmatrix}.$$

By Proposition 2.1(c) we may scale the columns of this submatrix to be of form (5.1), where $\mathbf{x}, \mathbf{y}, \mathbf{z}$ are each 1×1. In particular, $\bar{x} = \underline{x}$, $\bar{y} = \underline{y}$ and $\bar{z} = \underline{z}$. Thus by Theorem 5.4, this submatrix of A is nearly positive. By Corollary 2.5 it follows that A is nearly positive. ∎

6 Appending rows of zeros

Assume n nonnegative vectors in $\mathbb{R}^m$ cannot be rotated simultaneously into the interior of the nonnegative cone. Now suppose we embed these vectors in some $\mathbb{R}^k$, $k > m$, in the simplest way, by appending zero entries to the vectors. Is it possible that the embedded vectors could be rotated into the interior of the nonnegative cone? That is, if A is an $m \times n$ nonnegative matrix which is not nearly positive, can we generate a nearly positive matrix by appending rows of zeros to A?

We give two examples to show that this is possible.

Example 6.1 *Let*

$$A = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 2 & 1/2 \\ 1 & 1 & 0 & 0 \end{bmatrix}$$

be the matrix of Example 4.2. Let

$$\hat{A} = \left[\frac{A}{O_{3\times 4}} \right].$$

We have shown that A is not nearly positive. However, $\hat{A}$ is nearly positive! To see this, let

$$K = \left[\begin{array}{ccc|ccc} 0 & 0 & 0 & 1 & -2.1 & -2.1 \\ 0 & 0 & 0 & -2.1 & 1 & -2.1 \\ 0 & 0 & 0 & -1.1 & -1.1 & 2 \\ \hline -1 & 2.1 & 1.1 & 0 & 0 & 0 \\ 2.1 & -1 & 1.1 & 0 & 0 & 0 \\ 2.1 & 2.1 & -2 & 0 & 0 & 0 \end{array} \right].$$

Then K is skew-symmetric,

$$K\hat{A} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \mathbf{3.2} & \mathbf{.1} & \mathbf{3.2} & \mathbf{.05} \\ \mathbf{.1} & \mathbf{3.2} & \mathbf{.1} & \mathbf{1.6} \\ \mathbf{.1} & \mathbf{.1} & \mathbf{6.3} & \mathbf{3.15} \end{bmatrix},$$

and

$$K^2\hat{A} = \begin{bmatrix} \mathbf{2.78} & -6.83 & -10.24 & -9.925 \\ -6.83 & \mathbf{2.78} & -19.85 & -5.12 \\ -3.43 & -3.43 & \mathbf{8.97} & \mathbf{4.485} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

It follows that for small, positive ϵ the matrix $\hat{A} + 2\epsilon K\hat{A} + 2\epsilon^2 K^2 \hat{A}$ is positive, and thus $(I - \epsilon K)^{-1}(I + \epsilon K)\hat{A}$ is positive. Therefore, $\hat{A}$ is nearly positive.

Is it possible to append less than three zero rows to A of Example 6.1 to get a nearly positive matrix? We don't know the answer. But in the next example, one additional zero row is enough.

Example 6.2 *Let*

$$A = \begin{bmatrix} 0 & \mathbf{v}^T \\ \mathbf{u} & I \end{bmatrix} \in \mathbb{R}^{n\times n},$$

where $\mathbf{u}$ and $\mathbf{v}$ are positive vectors in $\mathbb{R}^{n-1}$. In Example 4.4 we showed that when $\mathbf{u} = \mathbf{v} = \mathbf{1}$ this matrix is not nearly positive. We now show that appending a row of zeros yields a nearly positive matrix.

Let
$$\hat{A} = \left[\begin{array}{c} A \\ \hline O_{1\times n} \end{array}\right].$$
For every $0 < \epsilon < 1$*, let*
$$\hat{A}_\epsilon = \left[\begin{array}{cc} 0 & \sqrt{1-\epsilon^2}\mathbf{v}^T \\ \mathbf{u}_\epsilon & I \\ \alpha(\epsilon) & \epsilon\mathbf{v}^T \end{array}\right],$$
where
$$\alpha(\epsilon) = 2\epsilon\frac{\mathbf{v}^T\mathbf{u}}{1+\epsilon^2||\mathbf{v}||^2} \quad \text{and} \quad \mathbf{u}_\epsilon = \mathbf{u} - \epsilon\alpha(\epsilon)\mathbf{v}.$$
Then $\hat{A}_\epsilon^T\hat{A}_\epsilon = \hat{A}^T\hat{A}$*. For a sufficiently small* $\epsilon > 0$ $\mathbf{u}_\epsilon > 0$ *and thus* $\hat{A}_\epsilon$ *has a positive last row and is therefore, by Proposition 2.4, nearly positive. For such small* $0 < \epsilon < 1$ *we also have that the matrix*
$$\left[\begin{array}{cc} 0 & \mathbf{u}_\epsilon^T \\ \sqrt{1-\epsilon^2}\,\mathbf{v} & I \end{array}\right]$$
is invertible (the Schur complement of I *in this matrix is* $-\sqrt{1-\epsilon^2}\,\mathbf{u}_\epsilon^T\mathbf{v}$*, so its determinant is* $-\sqrt{1-\epsilon^2}\,\mathbf{u}_\epsilon^T\mathbf{v}\cdot\det(I) \neq 0$*). Thus for every such* ϵ *there exists an* $\mathbf{x}^\epsilon \in \mathbb{R}^{n+1}$ *such that* $\hat{A}_\epsilon^T\mathbf{x}^\epsilon = \mathbf{0}$*,* $||\mathbf{x}_\epsilon|| = 1$*, and* $(\mathbf{x}^\epsilon)_{n+1} > 0$*. Note that* $\lim_{\epsilon\to 0+}\mathbf{x}^\epsilon = \mathbf{e}_{n+1}$*, the vector in* $\mathbb{R}^{n+1}$ *with all zero entries except for 1 in position* $n+1$*. Let* $\bar{A} = \left[\begin{array}{c|c} \hat{A} & \mathbf{e}_{n+1} \end{array}\right]$ *and* $\bar{A}_\epsilon = \left[\begin{array}{c|c} \hat{A}_\epsilon & \mathbf{x}^\epsilon \end{array}\right]$*. Then* $\bar{A}_\epsilon^T\bar{A}_\epsilon = \bar{A}^T\bar{A}$ *and hence* $Q_\epsilon = \bar{A}_\epsilon\bar{A}^{-1}$ *is orthogonal. Obviously,* $Q_\epsilon\hat{A} = \hat{A}_\epsilon$*. Since* $\lim_{\epsilon\to 0+}\bar{A}_\epsilon = \bar{A}$*,* $\lim_{\epsilon\to 0+}Q_\epsilon = I$*.*

By Proposition 2.2 this implies that $\hat{A}$ *is nearly-nearly positive, and thus nearly positive.*

In the next section we discuss the connection to completely positive matrices, and use it to show that it is not always possible to turn a non-nearly positive matrix into a nearly positive one by appending a row of zeros.

7 Connection to Completely Positive Matrices

In this section we make some connections between completely positive matrices, copositive matrices, and nearly positive matrices. We begin by reviewing some definitions and basic notions.

A (symmetric) matrix M is *completely positive* if $M = X^TX$ where X is a (not necessarily square) nonnegative matrix. The set of all $n\times n$ completely positive matrices
$$\mathcal{C}_n^* = \left\{M \in \mathbb{R}^{n\times n} : M = X^TX, X \geq 0\right\}$$
is a closed convex cone in the space $\mathcal{S}_n$ of symmetric $n\times n$ matrices. The * in the notation marks the fact that this cone is the dual of the cone $\mathcal{C}_n$ of $n\times n$

copositive matrices: an $n \times n$ matrix F is *copositive* if $\mathbf{x}^T F \mathbf{x} \geq 0$ for every $\mathbf{x} \in \mathbb{R}^n_+$, and the *dual* of a cone $\mathcal{K}$ in an inner product vector space V is

$$\mathcal{K}^* = \{\mathbf{x} \in V : \langle \mathbf{x}, \mathbf{y} \rangle \geq 0 \text{ for every } \mathbf{y} \in \mathcal{K}\}.$$

Here, the space is $\mathcal{S}_n$ with the Frobenius inner product $\langle X, Y \rangle = \operatorname{trace}(XY)$.

The interior of the completely positive cone $\mathcal{C}_n^*$ was characterized in [5] and the characterization refined in [3] as follows.

Theorem 7.1 *A completely positive matrix M is in the interior of $\mathcal{C}_n^*$ if and only if $M = X^T X$, where X has full column rank and at least one positive row.*

By the duality of $\mathcal{C}_n^*$ and $\mathcal{C}_n$, a completely positive matrix M is on the boundary of the cone $\mathcal{C}_n^*$ if and only if M is orthogonal to at least one (extreme) copositive matrix E, where an *extreme copositive* matrix is a matrix generating an extreme ray of the cone $\mathcal{C}_n$.

For $n \leq 4$, there are two types of extreme copositive matrices: positive semidefinite matrices (of rank 1) and nonnegative matrices (either with a single diagonal positive entry or a single pair of off-diagonal positive entries). The completely positive matrices orthogonal to these extreme copositive matrices are either singular or have a zero. However, for every $n \geq 5$ there exist extreme copositive matrices which are neither positive semidefinite nor nonnegative, and there are completely positive matrices orthogonal to them. Unfortunately, not all the extreme copositive matrices for $n \geq 6$ are known (characterizing those is a major open problem). The 5×5 extreme copositive matrices are, however, fully known. In addition to the positive semidefinite and the nonnegative ones, the others are (up to diagonal scaling and/or simultaneous permutation of rows and columns) the (now famous) Horn matrix [2]

$$\begin{bmatrix} 1 & -1 & 1 & 1 & -1 \\ -1 & 1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 \\ 1 & 1 & -1 & 1 & -1 \\ -1 & 1 & 1 & -1 & 1 \end{bmatrix}, \tag{7.1}$$

and the Hildebrand matrices, explictly described in [10]

$$\begin{bmatrix} 1 & -\cos(\theta_1) & \cos(\theta_1+\theta_2) & \cos(\theta_4+\theta_5) & -\cos(\theta_5) \\ -\cos(\theta_1) & 1 & -\cos(\theta_2) & \cos(\theta_2+\theta_3) & \cos(\theta_5+\theta_1) \\ \cos(\theta_1+\theta_2) & -\cos(\theta_2) & 1 & -\cos(\theta_3) & \cos(\theta_3+\theta_4) \\ \cos(\theta_4+\theta_5) & \cos(\theta_2+\theta_3) & -\cos(\theta_3) & 1 & -\cos(\theta_4) \\ -\cos(\theta_5) & \cos(\theta_5+\theta_1) & \cos(\theta_3+\theta_4) & -\cos(\theta_4) & 1 \end{bmatrix}, \tag{7.2}$$

where each θ_j is positive and $\sum_{i=1}^5 \theta_i < \pi$.

A nonzero nonnegative vector $\mathbf{x} \in \mathbb{R}^n$ is a *zero* of $E \in \mathcal{C}_n$ if $\mathbf{x}^T E \mathbf{x} = 0$. If $M \in \mathcal{C}_n^*$ and $M = X^T X$, $X \geq O$, then M is orthogonal to an extreme copositive

$E \in \mathcal{C}_n$ if and only if each row of X is a zero of E. Some facts are known about zeros of extreme copositive matrices that are not positive semidefinite. In particular, for our purposes it is useful to mention the following result (see e.g. [4, Corollary 4.14]), where $\operatorname{supp} \mathbf{x} = \{1 \leq i \leq n : x_i \neq 0\}$.

Proposition 7.2 *Let $E \in \mathcal{C}_n$ be an extreme copositive matrix, which is indefinite. Let $\mathbf{x}$ be a zero of E. Then $|\operatorname{supp} \mathbf{x}| \leq n - 2$.*

The zeros of the Horn matrix (7.1) are the cyclic permutations of vectors of the form

$$\begin{bmatrix} s \\ s+t \\ t \\ 0 \\ 0 \end{bmatrix}, \quad s, t \geq 0, \quad \text{and } s + t > 0.$$

The supports of the zeros of Hildebrand matrices (7.2) are

$$\{1,2,3\}, \{2,3,4\}, \{3,4,5\}, \{4,5,1\}, \{5,1,2\}.$$

We can now consider implications of the theory of completely positive and copositive matrices to our problem. If $A \geq O$ is nearly positive, then by definition $A^T A$ is completely positive. The characterization of $\operatorname{int} \mathcal{C}_n^*$ given in Theorem 7.1 yields the following addition to Proposition 2.3

Proposition 7.3 *Let $A \geq O$ have full column rank. If A is nearly positive, then $A^T A \in \operatorname{int} \mathcal{C}_n^*$.*

Proof. Assume that A is nearly positive. Then there exists an orthogonal to Q such that $QA > O$. Since $A^T A = (QA)^T (QA)$ and $QA > O$ is of full column rank, $A^T A \in \operatorname{int} \mathcal{C}_n^*$. ∎

The following example shows that this necessary condition for near positivity of a nonnegative matrix of full column rank is not sufficient.

Example 7.4 *Let A be as in Example 4.4, where it was shown that A is not nearly positive. But the matrix*

$$M = A^T A = \left[\begin{array}{c|c} n-1 & \mathbf{1}^T \\ \hline \mathbf{1} & I_{n-1} + J_{n-1} \end{array}\right]$$

is in the interior of the completely positive cone. Indeed,

$$M = D + J_n = D^{1/2}(I + \mathbf{x}\mathbf{x}^T) D^{1/2},$$

where

$$D = \left[\begin{array}{c|c} n-2 & \mathbf{0}^T \\ \hline \mathbf{0} & I_{n-1} \end{array}\right] \quad \text{and} \quad \mathbf{x} = D^{-1/2}\mathbf{1}_n.$$

The matrix $I + \mathbf{x}\mathbf{x}^T$ has a positive square root B ($B = I + \mu \mathbf{x}\mathbf{x}^T$, for $\mu = \frac{\sqrt{1+\mathbf{x}^T\mathbf{x}}-1}{\mathbf{x}^T\mathbf{x}}$). Thus $M = C^T C$, where $C = BD^{1/2} > O$.

In the next example we use Proposition 7.3 to show that a certain nonsingular 5×5 matrix is not nearly positive. We later see that, unlike the previous matrix with this property (of Example 4.4), this matrix cannot be made nearly positive by appending a row of zeros.

Example 7.5 *Let*

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

Then $A^T A > O$, *but* A *is not nearly positive. To see that, note that* $\operatorname{trace}(A^T A H) = 0$, *where* H *is the Horn matrix (7.1). Thus for every orthogonal* Q *such that* $QA > O$, $\operatorname{trace}((QA)^T (QA) H) = \operatorname{trace}(A^T A H) = 0$ *implies that the rows of* QA *are zeros of the matrix* H. *But none of the zeros of* H *is a positive vector.*

The fact that A *is not nearly positive can also be verified by using Theorem 4.1: The matrix*

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

satisfies $A \circ Y = O$, AY^T *is symmetric, and* $A + Y > O$. *The only skew-symmetric matrix* K *such that* $(KA)_{ij} = 0$ *whenever* $a_{ij} = 0$ *is the zero matrix.*

Our final example shows that for every $m, n \geq 5$ there exists a nonsingular nonnegative $m \times n$ matrix A such that $A^T A > O$, such that no matrix obtained by appending a row of zeros to A is not nearly positive.

Example 7.6 *Let* A *be the matrix of Example 7.5. Border it by* $n - 5$ *positive columns, and then append to the matrix* $m - 5$ *zero rows to obtain*

$$\hat{A} = \left[\begin{array}{c|c} A & E \\ \hline \multicolumn{2}{c}{O} \end{array} \right].$$

Then $\hat{A}^T \hat{A} > O$, *but* $\hat{A}$ *is not nearly positive, since* $\hat{A}^T \hat{A}$ *is orthogonal to* $H \oplus O$, *where* H *is the Horn matrix. Thus for any orthogonal matrix* Q *such that* $Q\hat{A} \geq O$, *the rows of* $Q\hat{A}$ *are zeros of* $H \oplus O$, *and cannot be positive.*

Similar examples can be constructed using matrices whose rows are zeros of a Hildebrand matrix.

8 Conclusion and open problems

We have found some necessary conditions for a nonnegative matrix A to be nearly positive (Proposition 2.3 and Theorem 3.6; see also Theorem 4.1), and a sufficient condition for A to be nearly positive (Theorem 3.2). A theorem of the alternative (Theorem 3.7) was proved, providing an additional method for checking whether this sufficient condition holds.

The basic necessary condition is the following: If A is nearly positive, then $A^TA > O$. We were able to show that this condition is also sufficient in the case that A has at most three columns or at most two rows (Theorems 2.8 and 5.6). The nearly positive matrices with exactly three rows were completely characterized in Theorem 5.5. It turns out that for every $n \geq 4$ there exist matrices of order $3 \times n$ that satisfy the necessary condition $A^TA > O$ and are not nearly positive (Example 4.2, Theorem 5.5). In fact, such examples were given in this paper for every order $m \times n$, where $m \geq 3$ and $n \geq 4$, except for the orders $m \times 4$, where $m \geq 5$ — see Examples 4.3, 4.4 and 7.6. This leaves the following question open:

Problem 8.1 *Let A be an $m \times 4$ nonnegative matrix, $m \geq 5$, such that $A^TA > O$. Is A necessarily nearly positive?*

One way one might try to generate a matrix A of order $m \times 4$, $m \geq 5$, which is not nearly positive, is by appending a zero row to a 3×4 or a 4×4 matrix which is not nearly positive. However, we have shown that appending zero rows to a nonnegative matrix which is not nearly positive may result in a nearly positive matrix (Examples and 6.1 and 6.2). On the other hand, there are matrices for which no amount of appended zero rows results in a matrix which is nearly positive (Example 7.6). This brings us to the next open problem.

Problem 8.2 *Which (not nearly positive) matrices can be made nearly positive by appending zero rows? Is there a bound on the number of zero rows required?*

When a nonnegative $m \times n$ matrix A has full column rank, a necessary condition on A to have this property is that A^TA is in the interior of the completely positive cone $\mathcal{C}_n^*$ (Proposition 7.3). Is this necessary condition also sufficient? And what about nonnegative matrices that are not of full column rank?

Finally, in Proposition 2.1 we have stated some simple basic results. According to these results, near-positivity is preserved under row permutations, column permutations and right multiplication by a diagonal matrix with a positive diagonal. It is natural to ask:

Problem 8.3 *Let A be an $m \times n$ nonnegative matrix, and D an $m \times m$ diagonal matrix with positive diagonal. Is it true that A is nearly positive if and only if DA is nearly positive?*

Note that since A is nonnegative, $A^TA > O$ if and only if $(DA)^T(DA) > O$, so A and DA either both satisfy this necessary condition or both don't satisfy it. In particular, the answer to Problem 8.3 is 'yes' for every matrix A with at most 2 rows or at most 3 columns. If A is a $3 \times n$ matrix of the form (5.1) and

$$D = \begin{bmatrix} r & 0 & 0 \\ 0 & s & 0 \\ 0 & 0 & t \end{bmatrix},$$

then DA is nearly positive if and only if DAE is, where

$$E = \begin{bmatrix} \frac{1}{s} I_\ell & O & O \\ O & \frac{1}{t} I_p & O \\ O & O & \frac{1}{r} I_q \end{bmatrix}.$$

But

$$DAE = \left[\begin{array}{c|c|c} \mathbf{0}^T & \frac{r}{t}\mathbf{y}^T & \mathbf{1}^T \\ \hline \mathbf{1}^T & \mathbf{0}^T & \frac{s}{r}\mathbf{z}^T \\ \hline \frac{t}{s}\mathbf{x}^T & \mathbf{1}^T & \mathbf{0}^T \end{array}\right],$$

hence the product of the the maximum (minimum) entries of $\frac{t}{s}\mathbf{x}$, $\frac{r}{t}\mathbf{y}$ and $\frac{s}{r}\mathbf{z}$ is equal to $\bar{x}\bar{y}\bar{z}$ ($\underline{x}\underline{y}\underline{z}$). By Theorem 5.4 DAE (and therefore also DA) is nearly positive if and only if A is. Theorem 5.5 then implies that the answer to Problem 8.3 is also affirmative for matrices with 3 rows.

Moreover, an $m \times n$ nonnegative matrix A satisfies the sufficient condition of Theorem 3.2 if and only if DA does (a skew symmetric matrix K satisfies the condition $(KA)_{ij} > 0$ whenever $a_{ij} = 0$ if and only if the skew symmetric $L = D^{-1}KD^{-1}$ satisfies the condition $(L(DA))_{ij} > 0$ whenever $(DA)_{ij} = 0$). This provides more cases when A and DA are both nearly positive. Yet the full answer to Problem 8.3 is not known to us.

Acknowledgment

Discussions on the subject of this paper commenced while the authors were attending a workshop at the Banff International Research Station for Mathematical Innovation and Discovery in July 2012. BIRS' hospitality and the environment conducive to research they provide is greatly appreciated.

The authors thank the anonymous referee for comments that helped improve the paper.

References

[1] Achyia Dax. An Elementary Proof of Farkas' Lemma. *SIAM Review*, 39:503–507, 1997.

[2] Palahenedi H. Diananda. On non-negative forms in real variables some or all of which are non-negative. *Proceedings of the Cambridge Philosophical Society*, 58:17–25, 1962.

[3] Peter J.C. Dickinson. An improved characterisation of the interior of the completely positive cone. *Electronic Journal on Linear Algebra*, 20:723–729, 2010.

[4] Peter J.C. Dickinson, Mirjam Dür, Luuk Gijben, and Roland Hildebrand. Irreducible elements of the copositive cone. *Linear Algebra and its Applications*, 439:16051626, 2013.

[5] Mirjam Dür and Georg Still, Interior points of the completely positive cone. *Electronic Journal on Linear Algebra*, 17:48–53, 2008.

[6] Abed Elhashash, Uriel G. Rothblum, and Daniel B. Szyld. Paths of matrices with the strong Perron-Frobenius property converging to a given matrix with the Perron-Frobenius property. *Electronic Journal on Linear Algebra*, 19:90–97, 2009.

[7] Abed Elhashash and Daniel B. Szyld. On general matrices having the Perron-Frobenius property. *Electronic Journal on Linear Algebra*, 17:389–413, 2008.

[8] Julius Farkas. Theorie der einfachen Ungleichungen. *Journal für die reine und angewandte Mathematik*, 124:1–27, 1902.

[9] Feliks R. Gantmacher. *The Theory of Matrices*, Vol. 1. Chelsea, New York, 1959.

[10] Roland Hildebrand, The extreme rays of the 5×5 copositive cone. *Linear Algebra and its Applications* 437 (2012), 1538–1547.

[11] R. Tyrrell Rockafellar. *Convex Analysis.* Princeton University Press, Princeton, 1970.