Linear Algebra and its Applications 436 (2012) 2366–2374

Contents lists available at SciVerse ScienceDirect

Linear Algebra and its Applications

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

A characterization of minimal Hermitian matrices

E. Andruchow [a,b], G. Larotonda [a,b], L. Recht [c], A. Varela [a,b,*]

[a] Instituto de Ciencias, Universidad Nacional de General Sarmiento, J.M. Gutiérrez 1150, (1613) Los Polvorines, Argentina
[b] Instituto Argentino de Matemática (IAM) "Alberto P. Calderón", CONICET, Saavedra 15, 3er piso (C1083ACA), Buenos Aires, Argentina
[c] Universidad Simón Bolívar, Apartado 89000, Caracas 1080A, Venezuela

ARTICLE INFO

Article history:
Received 20 April 2011
Accepted 23 August 2011
Available online 17 September 2011

Submitted by T. Ando

AMS classification:
15A12
15B51
15B57
15A60
58B25

Keywords:
Minimal Hermitian matrix
Diagonal matrix
Quotient operator norm
Best approximation

ABSTRACT

We describe properties of a Hermitian matrix $M \in M_n(\mathbb{C})$ having minimal quotient norm in the following sense:

$$\|M\| \leqslant \|M + D\|$$

for all real diagonal matrices $D \in M_n(\mathbb{C})$. Here $\|\ \|$ denotes the operator norm. We show a constructive method to obtain all the minimal matrices of any size.

1. Introduction

Let $M_n(\mathbb{C})$ and $D_n(\mathbb{R})$ be, respectively, the algebras of complex and real diagonal $n \times n$ matrices. In this paper we describe Hermitian matrices $M \in M_n(\mathbb{C})$ such that

$$\|M\| \leqslant \|M + D\|, \quad \text{for all } D \in D_n(\mathbb{R})$$

or equivalently

$$\|M\| = \text{dist}\,(M, D_n(\mathbb{R}))\,,$$

* Corresponding author at: Instituto de Ciencias, Universidad Nacional de General Sarmiento, J.M. Gutiérrez 1150, (1613) Los Polvorines, Argentina.

E-mail addresses: eandruch@ungs.edu.ar (E. Andruchow), glaroton@ungs.edu.ar (G. Larotonda), recht@usb.ve (L. Recht), avarela@ungs.edu.ar (A. Varela).

0024-3795/$ - see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.laa.2011.08.042

where $\| \ \|$ denotes the operator norm. These matrices M will be called minimal. These matrices appeared in the study of minimal length curves in the flag manifold $\mathcal{P}(n) = \mathcal{U}(M_n(\mathbb{C}))/\mathcal{U}(\mathcal{D}_n(\mathbb{C}))$, where $\mathcal{U}(\mathcal{A})$ denotes the unitary matrices of the algebra $\mathcal{A}$, when $\mathcal{P}(n)$ is endowed with the quotient Finsler metric of the operator norm [4]. Minimal length curves δ in $\mathcal{P}(n)$ are given by the left action of $\mathcal{U}(M_n(\mathbb{C}))$ on $\mathcal{P}(n)$. Namely

$$\delta(t) = \left[e^{itM}\right],$$

where M is minimal and $[U]$ denotes the class of U in $\mathcal{P}(n)$.

The following theorem follows ideas in [4], where this problem was also studied in the context of von Neumann and C* algebras. The next result was stated in Theorem 3.3 of [1] for 3×3 matrices. The same proof holds for $n \times n$ matrices.

Theorem 1. *A Hermitian matrix $M \in M_n(\mathbb{C})$ is minimal in the above sense if and only if there exists a positive semidefinite matrix $P \in M_n^h(\mathbb{C})$ such that*

- *$PM^2 = \lambda^2 P$ for $\lambda = \|M\|$.*
- *All the diagonal elements of PM are zero.*

Previous attempts to describe minimal matrices were done in [1] for 3×3 matrices. In that paper, all 3×3 minimal matrices were parametrized. We note that, Theorem 1 does not show how to construct $n \times n$ minimal matrices. Our goal in the present paper is to study some properties of $n \times n$ minimal matrices that allow the construction of them.

Minimal operators were studied in [8] where Theorem 2.2 of [1] was used to relate Leibnitz seminorms with quotient norms in C*-algebras.

2. Preliminaries and notation

Let $M_n(\mathbb{C})$ be the algebra of square complex matrices of $n \times n$, $M_n^h(\mathbb{C})$ the real subspace of Hermitian complex matrices, and $D_n(\mathbb{R})$ the real subalgebra of the diagonal real matrices. We denote with $\|A\|$ the usual operator norm of $A \in M_n(\mathbb{C})$ and with $\|A\|_1 = \mathrm{tr}(|A|) = \mathrm{tr}\left((A^*A)^{1/2}\right)$ the trace norm of A, where tr denotes the usual (non-normalized) trace.

Given a matrix $A \in M_n^h(\mathbb{C})$, $\lambda(A) \subset \mathbb{R}^n$ denotes the set of the eigenvalues of A, in decreasing order and counting multiplicity, that is,

$$\lambda(A) = (\lambda_1, \lambda_2, \dots, \lambda_n),$$

with $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n$. In this context $\lambda_{min}(A)$ and $\lambda_{max}(A)$ denote the smallest and biggest eigenvalues of A respectively.

The symbol $\sigma(A)$ denotes here the set (unordered) of eigenvalues of A.

We denote with $\{e_i\}_{i=1}^n$ the canonical basis of $\mathbb{C}^n$. Given a matrix $A \in M_n^h(\mathbb{C})$, we denote with $a_{i,j}$ the i, j entry of A and we write $A = [a_{i,j}]$ for $i, j = 1, \dots, n$.

Observe that if $M \in M_n^h(\mathbb{C})$ and $D \in D_n(\mathbb{R})$ then $(M + D) \in M_n^h(\mathbb{C})$. Let us consider the quotient $M_n^h(\mathbb{C})/D_n(\mathbb{R})$ and the quotient norm

$$|||\,[M]\,||| = \min_{D \in D_n(\mathbb{R})} \|M + D\| = \mathrm{dist}\,(M, D_n(\mathbb{R}))$$

for $[M] = \{M + D : D \in D_n(\mathbb{R})\} \in M_n^h(\mathbb{C})/D_n(\mathbb{R})$. The minimum is clearly attained.

Definition 1. *A matrix $M \in M_n^h(\mathbb{C})$ is called* **minimal** *if*

$$\|M\| \leqslant \|M + D\| \quad \text{for all } D \in D_n(\mathbb{R}),$$

or equivalently, if $\|M\| = |||\,[M]\,||| = \min_{D \in D_n(\mathbb{R})} \|M + D\| = dist\,(M, D_n(\mathbb{R}))$.

Remark 1. Note that if $M \in M_n^h(\mathbb{C})$ is a minimal matrix then its spectrum is centered, i.e. $\|M\|$, $-\|M\| \in \sigma(M)$. In general, for a given matrix $A \in M_n^h(\mathbb{C})$, $\pm\|A\| \in \sigma(A)$ if and only if $\|A\| = \min_{\lambda\in\mathbb{R}}\|A + \lambda I\|$ if and only if $\lambda_{min}(A) + \lambda_{max}(A) = 0$.

For $a_1, a_2, \ldots, a_n \in \mathbb{R}$ we denote with $\operatorname{diag}(a_1, a_2, \ldots, a_n)$ the diagonal matrix of $D_n(\mathbb{R})$ with $a_1, a_2, \ldots, a_n$ on the diagonal.

Given $v \in \mathbb{C}^n$, $v \otimes v$ denotes the linear map in $\mathbb{C}^n$ defined by $(v \otimes v)(x) = \langle x, v\rangle v$.

Let us denote with Φ the linear map from $M_n^h(\mathbb{C})$ to $D_n(\mathbb{R})$ defined by

$$\Phi(X) = \operatorname{diag}(x_{1,1}, \ldots, x_{n,n}), \quad \text{for } X = [x_{i,j}] \in M_n^h(\mathbb{C}).$$

Note that

$$\Phi(X) = \sum_{j=1}^{n} \langle Xe_j, e_j\rangle \, e_j \otimes e_j.$$

For $M \in M_n^h(\mathbb{C})$ and $v \in \mathbb{C}^n$ we write $\overline{M}$ and $\overline{v}$ to denote the matrix and vector obtained from M and v by conjugation of its coordinates.

If $M, N \in M_n(\mathbb{C})$ we denote with $M \circ N$ the Schur or Hadamard product of these matrices defined by $(M \circ N)_{i,j} = M_{i,j}N_{i,j}$ for $1 \leqslant i, j \leqslant n$. Therefore, if $v \in \mathbb{C}^n$, with coordinates in the canonical basis given by $v = (v_1, v_2, \ldots, v_n)$,

$$v \circ \overline{v} = (|v_1|^2, |v_2|^2, \ldots, |v_n|^2) = \sum_{j=1}^{n} |v_j|^2 e_j \in \mathbb{R}^n_+.$$

Observe that with these notations, if $X \in M_n^h(\mathbb{C})$ and $\{v_i\}_{i=1,\ldots,n}$ is an orthonormal basis of $\mathbb{C}^n$ of eigenvectors of X with corresponding eigenvalues $\lambda(X) = (\lambda_1, \ldots, \lambda_n)$, then $X = \sum_{i=1}^n \langle Xv_i, v_i\rangle \, v_i \otimes v_i = \sum_{i=1}^n \lambda_i \, v_i \otimes v_i$. Direct calculations with the canonical coordinates of these eigenvectors prove that

$$\Phi(X) = \operatorname{diag}\left(\sum_{i=1}^{n} \lambda_i \; (v_i \circ \overline{v_i})\right). \tag{2.1}$$

For $M, N \in M_n(\mathbb{C})$ the usual matrix product will be denoted with MN and $\operatorname{ran}(M)$ will denote the range of the linear transformation M.

3. Minimal matrices

It is apparent that for $X \in M_n^h(\mathbb{C})$

$$\operatorname{tr}(DX) = 0 \quad \forall\, D \in D_n(\mathbb{R}) \iff \Phi(X) = 0. \tag{3.1}$$

Then, from the Banach duality formula for the quotient norm and (3.1), it follows that

$$\max_{\substack{X\in M_n^h(\mathbb{C}),\, \Phi(X)=0 \\ \|X\|_1=1}} |\operatorname{tr}(MX)| = \min_{D\in D_n(\mathbb{R})} \|M + D\|. \tag{3.2}$$

Note that for an orthogonal projection E and $A \in M_n^h(\mathbb{C})$ the condition $EA = A$ is equivalent to $\operatorname{ran}(A) \subset \operatorname{ran}(E)$.

If $X \in M_n^h(\mathbb{C})$, let X^+ and X^- be the positive and negative parts of X, that is,

$$X^+ = \frac{|X| + X}{2} \text{ and } X^- = \frac{|X| - X}{2} \; (\text{with } |X| = (X^2)^{1/2} \geqslant 0).$$

Theorem 2. *Let* $0 \neq M \in M_n^h(\mathbb{C})$ *and* E_+ *(respectively* E_-*) the spectral projection of* M *corresponding to the eigenvalue* $\lambda_{max}(M)$ *(respectively* $\lambda_{min}(M)$*). The following conditions are equivalent:*

(i) *M is minimal.*

(ii) *There is a non-zero* $X \in M_n^h(\mathbb{C})$ *such that*

$$\Phi(X) = 0,\ E_+X^+ = X^+,\ E_-X^- = X^- \quad and \quad tr(MX) = \|M\|\,\|X\|_1.$$

(iii) $\lambda_{max}(M) + \lambda_{min}(M) = 0$*, and for any diagonal* $D \in D_n(\mathbb{R})$ *there exist* $y \in ran(E_+)$ *and* $z \in ran(E_-)$ *such that*

$$\|y\| = \|z\| = 1 \quad and \quad \langle Dy, y\rangle \leqslant \langle Dz, z\rangle.$$

Proof. We may assume that $\|M\| = 1$.

(i) $\Rightarrow$ (ii). Since M is minimal, by Remark 1 it must be $\lambda_{max} = 1$ and $\lambda_{min} = -1$. Consider the projections

$$E_1 = E_+,\ E_2 = E_- \quad \text{and} \quad E_3 = I - E_1 - E_2.$$

Then E_3 is the spectral projection of M corresponding to the open interval $(-1, 1)$, hence $E_3M = ME_3$ and $\|ME_3\| < 1$. Now M is written as

$$M = E_1 - E_2 + ME_3.$$

In view of (3.2) there exists $X \in M_n^h(\mathbb{C})$ such that

$$\Phi(X) = 0,\ \|X\|_1 = 1 \quad \text{and} \quad \text{tr}(MX) = 1. \tag{3.3}$$

In terms of the orthogonal decomposition $\mathbb{C}^n = \text{ran}(E_1) \oplus \text{ran}(E_2) \oplus \text{ran}(E_3)$, we can write

$$M = \begin{pmatrix} I & 0 & 0 \\ 0 & -I & 0 \\ 0 & 0 & M_{3,3} \end{pmatrix} \quad \text{and} \quad X = \begin{pmatrix} X_{1,1} & X_{1,2} & X_{1,3} \\ X_{2,1} & X_{2,2} & X_{2,3} \\ X_{3,1} & X_{3,2} & X_{3,3} \end{pmatrix}.$$

Let us to prove the identities $X_{1,2} = X_{2,1}^* = 0, X_{1,3} = X_{3,1}^* = 0, X_{2,3} = X_{3,2}^* = 0$ and $X_{3,3} = 0$.

The pinching inequality of Chandler Davis [2, IV.52] implies that

$$\left\| \begin{pmatrix} X_{1,1} & X_{1,2} \\ X_{2,1} & X_{2,2} \end{pmatrix} \right\|_1 + \|X_{3,3}\|_1 \leqslant \|X\|_1 = 1 \tag{3.4}$$

and

$$\|X_{1,1}\|_1 + \|X_{2,2}\|_1 \leqslant \left\| \begin{pmatrix} X_{1,1} & X_{1,2} \\ X_{2,1} & X_{2,2} \end{pmatrix} \right\|_1. \tag{3.5}$$

Note that $\|M_{3,3}\| < 1$. First let us show that $X_{3,3} = 0$. Suppose, that $\|X_{3,3}\|_1 \neq 0$. Then by (3.3) and the inequalities (3.4) and (3.5) we have

$$\begin{aligned} \|X\|_1 = 1 = \text{tr}(MX) &= \text{tr}(X_{1,1}) - \text{tr}(X_{2,2}) + \text{tr}(M_{3,3}X_{3,3}) \\ &\leqslant \|X_{1,1}\|_1 + \|X_{2,2}\|_1 + \|M_{3,3}\|\,\|X_{3,3}\|_1 \\ &< \|X_{1,1}\|_1 + \|X_{2,2}\|_1 + \|X_{3,3}\|_1 \leqslant \|X\|_1 = 1, \end{aligned}$$

a contradiction. Hence $X_{3,3} = 0$. Incidentally, we have proved that

$$\mathrm{tr}(X_{1,1}) = \|X_{1,1}\|_1 \quad \text{and} \quad \mathrm{tr}(-X_{2,2}) = \|-X_{2,2}\|_1.$$

Therefore, by the well-known fact that $\mathrm{tr}(Y) = \|Y\|_1$ occurs only if $Y \geqslant 0$, we have that

$$X_{1,1} \geqslant 0 \quad \text{and} \quad -X_{2,2} \geqslant 0. \tag{3.6}$$

Moreover by (3.3)

$$\mathrm{tr}(MX) = \mathrm{tr}\begin{pmatrix} X_{1,1} & X_{1,2} & X_{1,3} \\ -X_{2,1} & -X_{2,2} & -X_{2,3} \\ M_{3,3}X_{3,1} & M_{3,3}X_{3,2} & 0 \end{pmatrix} = 1 = \left\|\begin{pmatrix} X_{1,1} & X_{1,2} & X_{1,3} \\ -X_{2,1} & -X_{2,2} & -X_{2,3} \\ M_{3,3}X_{3,1} & M_{3,3}X_{3,2} & 0 \end{pmatrix}\right\|_1.$$

Then, by the same argument, the matrix MX should be positive semidefinite, which implies that $X_{1,3} = X_{3,1}^* = 0$ and $X_{2,3} = X_{3,2}^* = 0$.

In the same way from the relation

$$\mathrm{tr}\begin{pmatrix} X_{1,1} & X_{1,2} \\ -X_{2,1} & -X_{2,2} \end{pmatrix} = \left\|\begin{pmatrix} X_{1,1} & X_{1,2} \\ -X_{2,1} & -X_{2,2} \end{pmatrix}\right\|_1$$

we can conclude that $\begin{pmatrix} X_{1,1} & X_{1,2} \\ -X_{2,1} & -X_{2,2} \end{pmatrix} \geqslant 0$, and then $X_{1,2} = X_{2,1}^* = 0$.

Therefore

$$X = \begin{pmatrix} X_{1,1} & 0 & 0 \\ 0 & X_{2,2} & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{with } X_{1,1} \geqslant 0 \text{ and } X_{2,2} \leqslant 0,$$

which proves that $X^+ = E_+X_{1,1}E_+$ and $X^- = -E_-X_{2,2}E_-$, hence $E_+X^+ = X^+$ and $E_-X^- = X^-$.

(ii) $\Rightarrow$ (i) is immediate from (3.2).

(ii) $\Rightarrow$ (iii). Take a non-zero $X \in M_n^h(\mathbb{C})$ such that $\Phi(X) = 0$, $E_+ = X^+$, $E_-X^- = X^-$ and $\|X\|_1 = \mathrm{tr}(MX)$. Pick a diagonal $D \in D_n(\mathbb{R})$. Note that $X \neq 0$ and $\Phi(X) = 0$ imply that $\Phi(X^+) = \Phi(X^-) \neq 0$. Since $X^+, X^- \geqslant 0$, it follows that

$$\|X^+\|_1 = \|\Phi(X^+)\|_1 = \|\Phi(X^-)\|_1 = \|X^-\|_1 \neq 0.$$

The inequalities

$$\mathrm{tr}\left(\Phi(X^+)D\right) = \mathrm{tr}\left(X^+D\right) \geqslant \|X^+\|_1 \min_{y\in \mathrm{ran}(E_+),\ \|y\|=1} \langle Dy, y\rangle$$

and

$$\mathrm{tr}\left(\Phi(X^-)D\right) = \mathrm{tr}\left(X^-D\right) \leqslant \|X^-\|_1 \max_{z\in \mathrm{ran}(E_-),\ \|z\|=1} \langle Dz, z\rangle$$

prove (iii).

(iii) $\Rightarrow$ (ii). Suppose that there is no $0 \neq X \in M_n^h(\mathbb{C})$ satisfying the requirements of (ii). Consider the following two compact convex subsets of $M_n^h(\mathbb{C})$

$$\mathcal{A} = \{Y : E_+Y = Y \geqslant 0,\ \mathrm{tr}(Y) = 1\} \quad \text{and} \quad \mathcal{B} = \{Z : E_-Z = Z \geqslant 0,\ \mathrm{tr}(Z) = 1\}.$$

Since the assumption implies that $\Phi(\mathcal{A}) \cap \Phi(\mathcal{B}) = \emptyset$, the compact convex sets $\Phi(\mathcal{A})$ and $\Phi(\mathcal{B})$ in $\mathbb{R}^n$ are separated by a linear form, that is, there is a non-zero vector $d = (d_1, \ldots, d_n) \in \mathbb{R}^n$ such that

$$\min_{Y\in\mathcal{A}}\langle\Phi(Y),d\rangle > \max_{Z\in\mathcal{B}}\langle\Phi(Z),d\rangle.$$

This contradicts the condition (iii): taking $D = \text{diag}(d_1, \dots, d_n)$,

$$\min_{Y\in\mathcal{A}}\langle\Phi(Y),d\rangle = \min_{y\in\text{ran}(E_+),\ \|y\|=1}\langle Dy,y\rangle$$

and

$$\max_{Z\in\mathcal{B}}\langle\Phi(Z),d\rangle = \max_{z\in\text{ran}(E_-),\ \|z\|=1}\langle Dz,z\rangle.$$

This completes the proof. □

Remark 2. Let $M \in M_n^h(\mathbb{C})$ be a minimal matrix and $X \in M_n^h(\mathbb{C})$ be as in (ii) of the previous theorem. The functional $\psi(\cdot) = \text{tr}(X\cdot)$ is a witness for the fact that 0 is a best approximation to M in $D_n(\mathbb{R})$ as defined in [8]. That is, ψ is a norm one functional such that $\psi|_{D_n(\mathbb{R})} = 0$ and $\psi(M) = \|M - 0\|$.

4. An algorithm to construct minimal matrices

It is now clear that Theorem 2 can be used to construct all minimal matrices.

Theorem 3. *(step 1) Take non-zero $X \in M_n^h(\mathbb{C})$ with 0 diagonal (hence $X^+ \neq 0, X^- \neq 0$ and $\text{ran}(X^+) \perp \text{ran}(X^-)$).*
(step 2) Take non-zero orthoprojections E_+ and E_- such that $E_+E_- = 0, E_+X^+ = X^+$ and $E_-X^- = X^-$.
(step 3) Take $R \in M_n^h(\mathbb{C})$ such that $R(E_+ + E_-) = 0$ and $\|R\| < 1$.
Then $M = E_+ - E_- + R$ is a minimal matrix with $\|M\| = 1$.
Conversely every minimal matrix M with $\|M\| = 1$ is obtained in this way.

Remark 3. Note that for different $X \in M_n^h(\mathbb{C})$ with zero diagonal, the construction detailed in Theorem 3 may give the same orthoprojections E_+ and E_- onto $\text{ran}(X^+)$ and $\text{ran}(X^-)$, and therefore the same minimal matrices. Take for example the 3×3 unitary $U = \frac{1}{\sqrt{3}}\begin{pmatrix}1 & 1 & 1\\ 1 & w & w^2\\ 1 & w^2 & w\end{pmatrix}$ with $w = e^{i\frac{2\pi}{3}}$. Then define $X_t = U\text{diag}\,(1, t-1, -t)\,U^*$ for $t \in \mathbb{R}$ and $0 < t < 1$. It is apparent that $X_t \in M_n^h(\mathbb{C})$, $\Phi(X_t) = 0$ and $\|X_t\|_1 = 2$. By construction, if $t_1 \neq t_2$, the matrices X_{t_1} and X_{t_2} are different. However $\text{ran}\left((X_{t_1})^+\right) = \text{ran}\left((X_{t_2})^+\right)$ and $\text{ran}\left((X_{t_1})^-\right) = \text{ran}\left((X_{t_2})^-\right)$ for $t_1, t_2 \in (0, 1)$.

The following corollary is a slight variation of Theorem 1.

Corollary 1. A non-zero matrix $M \in M_n^h(\mathbb{C})$ is minimal if and only if there exists a non-zero positive semidefinite matrix $P \in M_n^h(\mathbb{C})$ such that

- $PM^2 = \lambda^2 P$ for $\lambda = \|M\|$.
- All the diagonal elements of PM are zero.
- P commutes with M.

Proof. If M is minimal and X is as in (ii) of Theorem 2 then $P = X^+ + X^-$ fulfills all the required conditions. That these conditions are necessary follows from Theorem 1. □

Recall that E_+ and E_- are the spectral projections corresponding respectively to the eigenvalues $\lambda_{max}(M)$ and $\lambda_{min}(M)$.

Corollary 2. A non-zero matrix $M \in M_n^h(\mathbb{C})$ is minimal if and only if $\lambda_{min}(M) + \lambda_{max}(M) = 0$ and there exist two non-zero positive semidefinite matrices $P, Q \in M_n^h(\mathbb{C})$ such that

- $\operatorname{ran}(P) \subset \operatorname{ran}(E_+)$ and $\operatorname{ran}(Q) \subset \operatorname{ran}(E_-)$.
- $\Phi(P) = \Phi(Q)$.
- $PQ = 0$.

Proof. If M is minimal and X is as in (ii) of Theorem 2, then $P = X^+$ and $Q = X^-$ satisfy all the required conditions. That these conditions are necessary for M to be minimal follows picking $X = \frac{1}{\|P-Q\|_1}(P - Q)$, which satisfies condition (ii) of Theorem 2. □

5. Spectral eigenspaces corresponding to λ_{min} and λ_{max} for a minimal matrix

In this section we describe some properties of the subspaces $\operatorname{ran}(E_+)$ and $\operatorname{ran}(E_-)$, where E_+ and E_- are the spectral projections of a minimal matrix M corresponding to the eigenvalues $\lambda_{max}(M)$ and $\lambda_{min}(M)$. As seen in Theorem 3 these are the building blocks of all the minimal matrices.

For given vectors $\{w_k\}_{k=1}^m \subset \mathbb{C}^n$ we denote with $\operatorname{co}\left(\{w_k\}_{k=1}^m\right)$ the convex hull generated by them.

Corollary 3. Let $M \in M_n^h(\mathbb{C})$ be a non-zero matrix such that $\lambda_{max}(M) + \lambda_{min}(M) = 0$. Then the following properties are equivalent:

(a) M is minimal.
(b) There exist orthonormal sets $\{v_i\}_{i=1}^r \subset \operatorname{ran}(E_+)$ and $\{v_j\}_{j=r+1}^{r+s} \subset \operatorname{ran}(E_-)$ such that

$$\operatorname{co}(\{v_i \circ \overline{v_i}\}_{i=1}^r) \cap \operatorname{co}\left(\{v_j \circ \overline{v_j}\}_{j=r+1}^{r+s}\right) \neq \emptyset. \tag{5.1}$$

Proof. Suppose that M is minimal. By using Theorem 2 there exists a non-zero $X \in M_n^h(\mathbb{C})$ that satisfies (ii) of that theorem. Fix a basis of $\operatorname{ran}(X^+)$ of orthonormal eigenvectors $\{v_i\}_{i=1}^r$ corresponding to the (strictly) positive eigenvalues $\{a_i\}_{i=1}^r$ of X^+, and a basis of $\operatorname{ran}(X^-)$ of orthonormal eigenvectors $\{v_j\}_{j=r+1}^{r+s}$ corresponding to the (strictly) positive eigenvalues $\{a_j\}_{j=r+1}^{r+s}$ of X^- (note that $\operatorname{ran}(X^+) \perp \operatorname{ran}(X^-)$). Then, since $X^+ = \sum_{i=1}^r a_i(v_i \otimes v_i)$ and $X^- = \sum_{j=r+1}^{r+s} a_j(v_j \otimes v_j)$, using the formula (2.1) for $\Phi(X^+)$ and $\Phi(X^-)$, it can be shown that

$$\Phi(X) = \Phi(X^+) - \Phi(X^-) = \operatorname{diag}\left(\sum_{i=1}^r a_i\ (v_i \circ \overline{v_i})\right) - \operatorname{diag}\left(\sum_{j=r+1}^{r+s} a_j\ (v_j \circ \overline{v_j})\right).$$

Since $\Phi(X) = 0$, it is apparent that $\sum_{i=1}^r a_i\ (v_i \circ \overline{v_i}) = \sum_{j=r+1}^{r+s} a_j\ (v_j \circ \overline{v_j})$ and $\operatorname{tr}(X^+) = \operatorname{tr}(X^-) > 0$, which proves that $\sum_{i=1}^r a_i = \sum_{j=r+1}^{r+s} a_j$. Therefore,

$$\sum_{i=1}^r \frac{a_i}{\sum_{i=1}^r a_i}\ (v_i \circ \overline{v_i}) = \sum_{j=r+1}^{r+s} \frac{a_j}{\sum_{j=r+1}^{r+s} a_j}\ (v_j \circ \overline{v_j})\,.$$

Then, since $\operatorname{ran}(X^+) \subset \operatorname{ran}(E_+)$ and $\operatorname{ran}(X^-) \subset \operatorname{ran}(E_-)$, (b) holds.

Conversely, if (b) holds, there exist $\alpha_i, \beta_j > 0$ satisfying $\sum_{i=1}^r \alpha_i = 1 = \sum_{i=r+1}^{r+s} \beta_j$, and orthonormal sets $\{v_i\}_{i=1}^r \subset \operatorname{ran}(E_+)$ and $\{v_j\}_{j=r+1}^{r+s} \subset \operatorname{ran}(E_-)$, such that

$$\sum_{i=1}^r \alpha_i(v_i \circ \overline{v_i}) = \sum_{j=r+1}^{r+s} \beta_j(v_j \circ \overline{v_j}) \ \in\ \operatorname{co}(\{v_i \circ \overline{v_i}\}_{i=1}^r) \cap \operatorname{co}\left(\{v_j \circ \overline{v_j}\}_{j=r+1}^{r+s}\right).$$

Put

$$X = \frac{1}{2}\left(\sum_{i=1}^{r} \alpha_i \ (v_i \otimes v_i) - \sum_{j=r+1}^{r+s} \beta_j \ (v_j \otimes v_j)\right).$$

It is straightforward that X satisfies condition (ii) of Theorem 2. Therefore M is minimal. □

The previous corollary could have been proved with similar techniques as in the proof of (ii) ⇒ (iii) in Theorem 2. Moreover, define the following subsets of $\mathbb{R}^n_+$

$$\mathcal{P}_+ = \bigcup_{\substack{\text{o.n. set } \{v_i\}_{i=1}^r \\ \{v_i\}_{i=1}^r \subset \text{ran}(E_+)}} \text{co}(\{v_i \circ \overline{v_i}\}_{i=1}^r) \ \text{ and } \ \mathcal{P}_- = \bigcup_{\substack{\text{o.n. set } \{v_j\}_{j=r+1}^{r+s} \\ \{v_j\}_{j=r+1}^{r+s} \subset \text{ran}(E_-)}} \text{co}\left(\{v_j \circ \overline{v_j}\}_{j=r+1}^{r+s}\right).$$

Then $\mathcal{P}_+$ and $\mathcal{P}_-$ induce the subsets $\Phi(\mathcal{A})$ and $\Phi(\mathcal{B}) \subset D_n(\mathbb{R})$, where $\mathcal{A}$ and $\mathcal{B}$ are the compact convex sets defined in the proof of Theorem 2. Then, $\mathcal{P}_+$ and $\mathcal{P}_-$ are compact and convex sets of $\mathbb{R}^n$. Therefore for a matrix M such that $\lambda_{min}(M) + \lambda_{max}(M) = 0$, the property $\mathcal{P}+ \cap \mathcal{P}_- \neq \emptyset$ is equivalent to being minimal.

A different way to construct minimal matrices is the following. Take $a_i > 0$, for $1 \leqslant i \leqslant r$, $a_j > 0$ for $r+1 \leqslant j \leqslant r+s$ with $1 \leqslant r, s, r+s \leqslant n$ and such that $\sum_{i=1}^r a_i = \sum_{j=r+1}^{r+s} a_j$. If we define $\vec{a} = (a_1, \ldots, a_r, -a_{r+1}, \ldots, -a_{r+s}, 0, \ldots, 0) \in \mathbb{R}^n$, it follows that $\vec{a}$ majorizes $\vec{0} = (0, \ldots, 0) \in \mathbb{R}^n$, and we will denote $\vec{0} \prec \vec{a}$ as usual (see [6] for basic facts on majorization). Then a concrete unitary matrix $U \in M_n(\mathbb{C})$ can be found (see [5–7]) such that $(U \circ \overline{U}) \in M_n(\mathbb{R}_+)$ satisfies that $(U \circ \overline{U})\vec{a} = \vec{0}$. This last equality can be written as

$$\sum_{i=1}^{r} a_i(v_i \circ \overline{v_i}) - \sum_{j=r+1}^{r+s} a_j(v_j \circ \overline{v_j}) = \vec{0},$$

where $\{v_k\}_{k=1}^n$ are the columns of the unitary U. Then any matrix of the form

$$M = \lambda \sum_{i=1}^{r} v_i \otimes v_i \ - \ \lambda \sum_{j=r+1}^{r+s} v_j \otimes v_j \ + \sum_{h=r+s+1}^{n} \lambda_h (v_h \otimes v_h) \tag{5.2}$$

is minimal, provided that $\lambda > 0$, $\lambda_h \in \mathbb{R}$, $|\lambda_h| < \lambda$. These computations provide a different way to construct examples of minimal matrices of any size.

In [3] several algorithms are produced to find unitary (or orthogonal) matrices U that satisfy $(U \circ \overline{U})\,\vec{a} = \vec{0}$ for a given $\vec{a}$. Nevertheless, the set of all possible unitaries U that satisfy $(U \circ \overline{U})\vec{a} = 0$ is not known in general. The papers [9] and [10] study this problem.

The method to obtain minimal matrices as in (5.2) has the disadvantage that M relies on the construction of the unitary U.

Remark 4. In [1] a different characterization of minimal 3×3 matrices was obtained. It is shown that given a 3×3 matrix M, with $\lambda(M) = (\lambda, \mu, -\lambda)$, $|\mu| \leqslant \lambda = \|M\|$, then, M is minimal, if and only if, there exists a normalized eigenvector v_λ of the eigenvalue λ and a normalized eigenvector $v_{-\lambda}$ of the eigenvalue $-\lambda$ such that $v_\lambda \circ \overline{v_\lambda} = v_{-\lambda} \circ \overline{v_{-\lambda}}$. The statement remains valid if any of the eigenvalues has multiplicity two ($\mu = \pm\lambda$). The following is an example of a 4×4 minimal Hermitian matrix where this condition does not hold. Let

$$M = \begin{pmatrix} \frac{9}{14} & -\frac{15}{14} - \frac{i}{7} & -\frac{1}{7} + \frac{5i}{7} & \frac{2}{7} + \frac{6i}{7} \\ -\frac{15}{14} + \frac{i}{7} & \frac{13}{14} & -\frac{1}{7} + i & \frac{6i}{7} \\ -\frac{1}{7} - \frac{5i}{7} & -\frac{1}{7} - i & \frac{5}{7} & -1 - \frac{2i}{7} \\ \frac{2}{7} - \frac{6i}{7} & -\frac{6i}{7} & -1 + \frac{2i}{7} & \frac{5}{7} \end{pmatrix}.$$

Then $\lambda(M) = (2, 2, 1, -2)$, and the eigenspace of the eigenvalue 2 is generated by the orthonormal eigenvectors

$$v_1 = \frac{1}{5\sqrt{2}}(-1-2i, 5, -3-i, 1-3i) \quad \text{and}$$
$$v_2 = \frac{1}{10\sqrt{14}}(17-11i, -15+5i, -9+17i, 3-19i).$$

The vector $w = \frac{1}{2\sqrt{2}}(1-i, 1-i, 1+i, 1+i)$ is a normalized eigenvector of eigenvalue -2. A direct calculation shows that for $\alpha = \frac{2}{9}$, $\alpha(v_1 \circ \overline{v_1}) + (1-\alpha)(v_2 \circ \overline{v_2}) = w \circ \overline{w} = (\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$, which proves that M is minimal (using Corollary 3). However, there is not an eigenvector v in the eigenspace of the eigenvalue 2 such that $v \circ \bar{v} = w \circ \overline{w}$. This follows writing $v = \beta v_1 + \gamma v_2$ with $\beta, \gamma \in \mathbb{C}$, and $|\beta|^2 + |\gamma|^2 = 1$, and proving that $v \circ \bar{v} = w \circ \overline{w}$ cannot happen (note that it can be supposed that $\gamma = \sqrt{1-|\beta|^2}$).

Acknowledgements

We are indebted to the editor Professor T. Ando for several significant improvements of the original manuscript.

References

[1] Esteban Andruchow, Luis E. Mata-Lorenzo, Alberto Mendoza, Lázaro Recht, Alejandro Varela, Minimal matrices and the corresponding minimal curves on flag manifolds in low dimension, Linear Algebra Appl. 430 (8–9) (2009) 1906–1928.
[2] Rajendra Bhatia, Matrix Analysis, Graduate Texts in Mathematics, vol. 169, Springer-Verlag, New York (1997).
[3] Inderjit S. Dhillon, W. Heath, Jr., Mátyŝ A. Sustik, Joel A. Tropp, Generalized finite algorithms for constructing Hermitian matrices with prescribed diagonal and spectrum, SIAM J. Matrix Anal. Appl. (electronic) 27 (1) (2005) 61–71.
[4] C.E. Durán, L.E. Mata-Lorenzo, L. Recht, Metric geometry in homogeneous spaces of the unitary group of a C-algebra: Part I – minimal curves, Adv. Math. 184 (2) (2004) 342–366.
[5] Alfred Horn, Doubly stochastic matrices and the diagonal of a rotation matrix, Amer. J. Math. 76 (1954) 620–630.
[6] Albert W. Marshall, Ingram Olkin, Inequalities: theory of majorization and its applications, Mathematics in Science and Engineering, vol. 143, Academic Press, Inc. (Harcourt Brace Jovanovich), Publishers, New York–London (1979).
[7] L. Mirsky, Matrices with prescribed characteristic roots and diagonal elements, J. London Math. Soc. 33 (1958) 14–21.
[8] M.A. Rieffel, Leibni tz seminorms and best approximation from C-subalgebras, Preprint arXiv:1008.3733v4 [math.OA].
[9] W. Tadej, K. Zyczkowski, Defect of a unitary matrix, with an appendix by Wojciech Slomczynski, Linear Algebra Appl. 429 (2-3) (2008) 447–481.
[10] K. Zyczkowski, M. Kus, W. Slomczynski, H.J. Sommers, Random unistochastic matrices, J. Phys. A: Math. Gen. 36 (12) (2003) 3425–3450.