Linear Algebra and its Applications 532 (2017) 387–396

Contents lists available at ScienceDirect

Linear Algebra and its Applications

www.elsevier.com/locate/laa

Maximal doubly stochastic matrix centralizers ☆

Henrique F. da Cruz [a], Gregor Dolinar [b,c], Rosário Fernandes [d,*], Bojan Kuzma [e,c]

[a] *Departamento de Matemática da Universidade da Beira Interior, Rua Marquês D'Avila e Bolama, 6201-001 Covilhã, Portugal*
[b] *University of Ljubljana, Faculty of Electrical Engineering, Tržaška cesta 25, 1000 Ljubljana, Slovenia*
[c] *IMFM, Jadranska cesta 19, 1000 Ljubljana, Slovenia*
[d] *Departamento de Matemática, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, 2829-516 Caparica, Portugal*
[e] *University of Primorska, Glagoljaška 8, 6000 Koper, Slovenia*

A R T I C L E   I N F O

Article history:
Received 21 September 2016
Accepted 19 June 2017
Available online 3 July 2017
Submitted by M. Tsatsomeros

MSC:
15A27

Keywords:
Doubly stochastic matrices
Matrix algebra
Centralizers

A B S T R A C T

We describe doubly stochastic matrices with maximal centralizers.

☆ This work was partially supported by Portuguese funds of FCT—Foundation for Science and Technology under the projects UID/MAT/00212/2013 and UID/MAT/00297/2013, and by Slovenian Research Agency (research core funding No. P1-0288 and No. P1-0222).

* Corresponding author.
E-mail addresses: hcruz@ubi.pt (H.F. da Cruz), gregor.dolinar@fe.uni-lj.si (G. Dolinar), mrff@fct.unl.pt (R. Fernandes), bojan.kuzma@upr.si (B. Kuzma).

http://dx.doi.org/10.1016/j.laa.2017.06.029
0024-3795/

1. Introduction

Let M_n be the algebra of all n-by-n matrices over the complex field $\mathbb{C}$ and let $I \in M_n$ be its identity. One of the relations which is often used on M_n, both in pure and applied problems, is commutativity [13–16]. In the study of commutativity the notion of *centralizer* (also called commutant) has an important role. For $A \in M_n$ its centralizer, denoted by $\mathcal{C}(A)$, is the set of all matrices commuting with A, that is

$$\mathcal{C}(A) = \{X \in M_n : AX = XA\},$$

and for a set $S \subseteq M_n$ its centralizer, denoted by $\mathcal{C}(S)$, is the intersection of centralizers of all its elements, that is

$$\mathcal{C}(S) = \{X \in M_n : AX = XA, \text{ for every } A \in S\}.$$

The Centralizer Theorem (see [6], p. 113, Corollary 1, or [16], p. 106, Theorem 2) states that $\mathcal{C}(\mathcal{C}(A)) = \mathbb{C}[A]$ where $\mathbb{C}[X] \subseteq M_n$ denotes the unital algebra spanned by $X \in M_n$. It is well known that the central elements of M_n are the scalar matrices, $\mathcal{C}(M_n) = \{\alpha I : \alpha \in \mathbb{C}\}$.

The centralizer induces an equivalence relation, $\sim$, and a preorder relation, $\preceq$, on M_n:

- A and B are $\mathcal{C}$-equivalent, $A \sim B$, if $\mathcal{C}(A) = \mathcal{C}(B)$,
- $A \preceq B$ if $\mathcal{C}(A) \subseteq \mathcal{C}(B)$.

For a preorder $\preceq$ on M_n we say that:

- a non-scalar matrix A is minimal if for every matrix X with $\mathcal{C}(X) \subseteq \mathcal{C}(A)$ it follows $\mathcal{C}(A) = \mathcal{C}(X)$,
- a non-scalar matrix A is maximal if for every non-scalar matrix X with $\mathcal{C}(X) \supseteq \mathcal{C}(A)$ it follows $\mathcal{C}(A) = \mathcal{C}(X)$.

The characterization of minimal and maximal matrices in M_n is known. For minimal matrices the classification consists of several equivalent conditions, we list only a few of them.

Proposition 1.1 *(see [4]). For $A \in M_n$ the following statements are equivalent.*

(1) *A is minimal.*
(2) *A is nonderogatory, i.e., the minimal polynomial of A equals its characteristic polynomial.*
(3) $\mathcal{C}(A) = \mathbb{C}[A]$.

For maximal matrices, the characterization is as follows:

Proposition 1.2 *(see Lemma 3.1 in [9]). A matrix $A \in M_n$ is maximal if and only if it is $\mathcal{C}$-equivalent to a non-scalar idempotent or $\mathcal{C}$-equivalent to a non-scalar square-zero matrix.*

It is our aim to classify doubly stochastic matrices, denoted henceforth by $\Omega_n \subseteq M_n$, with extremal centralizers. Recall that a *doubly stochastic matrix* is a square matrix of nonnegative real numbers with each row and column summing to 1. It is well-known (see, e.g. [8, Theorem A.2]) that Ω_n is a convex closure of the set of permutational matrices. Doubly stochastic matrices are important in the study of Markov chains, combinatorics, statistics and probability [1,7,10,11]. Matrices that commute with doubly stochastic matrices are the subject of several papers, see for example [2] and [12]. In particular, in graph theory a graph G with adjacency matrix A is called *compact* if A commutes with every doubly stochastic matrix.

The main results of our paper is a complete classification of maximal doubly stochastic matrices and matrices which are maximal when their centralizers are restricted to doubly stochastic matrices. Unlike for maximal matrices, the set of minimal doubly stochastic matrices is open (see Remark 3.8) and consists of nonderogatory doubly stochastic matrices. Therefore no further classification of minimal doubly stochastic matrices is given.

2. Matrices with maximal centralizers

We begin this section with a characterization of the maximal doubly stochastic matrices. Recall that in a doubly stochastic matrix P every entry lies in the interval $[0,1] \subseteq \mathbb{R}$. Also, it is a trivial observation that an entry-wise nonnegative matrix M is doubly-stochastic if and only if $M\mathbf{1} = \mathbf{1}$ and $\mathbf{1}^T M = \mathbf{1}^T$, where $\mathbf{1} = [1 \dots 1]^T$ is a column vector of ones and X^T denotes a transposition of (possibly non-square) matrix X.

Theorem 2.1. *A non-scalar matrix M is maximal doubly stochastic if and only if*

$$M = I - \alpha(I - P),$$

where $P = [p_{ij}] \neq I$ is a doubly stochastic idempotent and $\alpha \in \mathbb{R}$ satisfies $0 < \alpha \leq \min\left\{\frac{1}{1-p_{ii}} : \ p_{ii} \neq 1\right\}$.

Proof. If

$$M = I - \alpha(I - P),$$

with P an idempotent doubly stochastic matrix and α as in the statement of the theorem, then M is $\mathcal{C}$-equivalent to a non-scalar idempotent and so by Proposition 1.2, M is maximal. Also, $p_{ij} \geq 0$ implies that the off-diagonal entries of M are nonnegative.

Moreover, P is doubly stochastic so $p_{ii} \leq 1$ and then the assumption on α gives that diagonal entries of M are also nonnegative. Since clearly $M\mathbf{1} = \mathbf{1}$ and $\mathbf{1}^T M = \mathbf{1}^T$ we see that M is doubly stochastic.

Conversely, suppose that M is a maximal doubly stochastic matrix. Then by Proposition 1.2, M is $\mathcal{C}$-equivalent to either a square-zero matrix or an idempotent matrix.

In the first case $M = \lambda I + N$, where N is a square-zero matrix. Since M is doubly stochastic,

$$\mathbf{1} = M\mathbf{1} = (\lambda I + N)\mathbf{1} = \lambda\mathbf{1} + N\mathbf{1}.$$

Then

$$N\mathbf{1} = (1-\lambda)\mathbf{1}$$

and as N is square-zero we conclude that $\lambda = 1$ and $M = I + N$. Also, $\mathrm{Tr}(N) = 0$ and each diagonal entry of $N = M - I$ is non-positive since M is doubly stochastic. Therefore each diagonal entry of N equals zero and each diagonal entry of $M = N + I$ equals one. Since M is doubly stochastic this gives $M = I$, a contradiction.

In the second case $M = \lambda I + \mu Q$ with $\mu \neq 0$ and Q a non-scalar idempotent matrix. Clearly, the spectrum of M equals $\{\lambda, \lambda + \mu\}$. Since M is doubly stochastic, 1 is also its eigenvalue. Hence

$$\lambda = 1 \quad \text{or} \quad \lambda + \mu = 1. \tag{1}$$

In addition, since $M \geq 0$ entry-wise,

$$0 \leq \mathrm{Tr}(M) = \lambda\,\mathrm{Tr}(I) + \mu\,\mathrm{Tr}(Q) = \lambda\, n + \mu\,\mathrm{rank}(Q). \tag{2}$$

Then, (1)–(2) imply $\lambda, \mu \in \mathbb{R}$. By the Perron–Frobenius theorem for nonnegative matrices, the spectral radius $\rho(M)$ is an eigenvalue and the corresponding eigenvector x can be taken entrywise nonnegative. Multiplying $Mx = \rho(M)x$ by $\mathbf{1}^T$ on the left, we get

$$\mathbf{1}^T x = \mathbf{1}^T M x = \rho(M)\mathbf{1}^T x$$

and hence $\rho(M) = 1$. Therefore, as the spectrum of M equals $\{\lambda, \lambda + \mu\} \subseteq \mathbb{R}$, we obtain the following: if $\lambda = 1$ then $\mu < 0$ and if $\lambda + \mu = 1$ then $\mu > 0$. So,

$$M = I - \alpha(I - P),$$

where $P = Q$ and $\alpha = \mu > 0$ if $\lambda + \mu = 1$, and where $P = I - Q$ and $\alpha = -\mu > 0$ if $\lambda = 1$. In both cases P is an idempotent matrix. We are going to prove that P is in both cases a doubly stochastic matrix.

Since $M = I - \alpha(I - P)$ is doubly stochastic and $\alpha > 0$, the off-diagonals entries of $P = [p_{ij}]$ are non-negative. Let $i \in \{1, \ldots, n\}$. Then

$$p_{ii} = p_{ii}^2 + \varepsilon, \qquad \text{where} \qquad \varepsilon = \sum_{\substack{1 \le k \le n \\ k \neq i}} p_{ik} p_{ki} \ge 0$$

and we conclude that $p_{ii}^2 \le p_{ii}$. Consequently, $0 \le p_{ii} \le 1$. Moreover,

$$\mathbf{1} = M\mathbf{1} = (I - \alpha(I - P))\mathbf{1}$$

easily implies $P\mathbf{1} = \mathbf{1}$ and likewise $\mathbf{1}^T M = \mathbf{1}^T$ implies $\mathbf{1}^T P = \mathbf{1}^T$ so P is doubly stochastic. The condition that $0 < \alpha \le \min\left\{\frac{1}{1-p_{ii}} : \; p_{ii} \neq 1\right\}$ follows easily by comparing the diagonal entries in equation $M = I - \alpha(I - P)$ and using that M is doubly stochastic. □

We let $\mathbb{1} = \mathbf{1}\mathbf{1}^T$ be the matrix full of ones. When its size is important, we shall write $\mathbb{1}_k$ to denote the k-by-k matrix full of ones.

Remark 2.2. The characterization of the idempotent doubly stochastic matrices is known (see [3,5] or [8, Theorem I.2.]): If Q is a doubly stochastic idempotent, then Q is permutationally similar to a block-diagonal matrix

$$\left(\tfrac{1}{k_1}\mathbb{1}_{k_1}\right) \bigoplus \left(\tfrac{1}{k_2}\mathbb{1}_{k_2}\right) \bigoplus \cdots \bigoplus \left(\tfrac{1}{k_p}\mathbb{1}_{k_p}\right)$$

for some positive integer p and some positive integers k_i.

Now it is easy to characterize which permutation matrices are maximal. We denote by S_n the symmetric group of degree n. If $\sigma \in S_n$ we let $P(\sigma)$ be the corresponding permutation matrix.

Theorem 2.3. *Let $\sigma \in S_n$. Then $P(\sigma)$ is maximal in M_n if and only if σ is a product of disjoint transpositions.*

Proof. Assume that $\sigma \in S_n$ is the product of disjoint transpositions. Then $P(\sigma)^2 = I$ and so $P(\sigma) = I - 2P$ where $P = \frac{1}{2}(I - P(\sigma))$ is an idempotent. By Proposition 1.2, $P(\sigma)$ is maximal.

Conversely, assume that $P(\sigma)$ is a maximal matrix. Then by Theorem 2.1 there is an idempotent doubly stochastic matrix P such that

$$P(\sigma) = I - \alpha(I - P), \;\; \alpha > 0.$$

Clearly, $P(\sigma)$ is not an idempotent, so $\alpha \neq 1$. Note that each row and column of $P(\sigma)$ has only one entry different from zero. Hence, each row and column of $P = \frac{1}{\alpha}P(\sigma) - \frac{(1-\alpha)}{\alpha}I$ has at most two nonzero entries and one of them is in the main diagonal. Let $i \in \{1, \dots, n\}$. Now, if P has just one nonzero entry in the i-th row, then it is in the main diagonal. Since P is a doubly stochastic matrix this entry is 1. If P has two nonzero entries in the row i, then according to a classification of doubly stochastic idempotents given in Remark 2.2, these two entries are equal to $\frac{1}{2}$. At least one row of P does contain $\frac{1}{2}$ for otherwise, $P = P(\sigma) = I$, a contradiction. This implies that $\alpha = 2$, and hence σ is a product of disjoint transpositions. $\square$

Recall that every doubly stochastic matrix is a convex combination of permutation matrices. However, not every maximal doubly stochastic matrix can be obtained as a convex combination of maximal permutation matrices.

Example 2.4. There exists a maximal doubly stochastic matrix which is not a convex combination of maximal permutation matrices. To see this, consider the doubly stochastic idempotent $P = \frac{1}{3}\mathbb{1}_3$. By Theorem 2.1 the matrix

$$M = I - \tfrac{4}{3}(I - P) = \frac{1}{9}\begin{bmatrix} 1 & 4 & 4 \\ 4 & 1 & 4 \\ 4 & 4 & 1 \end{bmatrix}$$

is maximal doubly stochastic. By Theorem 2.3 the only maximal permutation matrices are $P(12), P(13)$, and $P(23)$. So, if

$$M = \alpha_1 P(12) + \alpha_2 P(13) + \alpha_3 P(23), \qquad \alpha_i \in [0,1],\ \sum \alpha_i = 1,$$

then it easily follows that $\alpha_1 = 1/9$ and $\alpha_1 = 4/9$, a contradiction.

With the next example we show that the polytope of the doubly stochastic matrices has two adjacent vertices that are maximal matrices. However, the edge between them only has minimal matrices.

Example 2.5. There exist two maximal doubly stochastic matrices such that the entire line between them consists of minimal matrices. Namely, as noted in Theorem 2.3, the matrices $P(12)$ and $P(13)$ are maximal. But the doubly stochastic matrices

$$B_t = tP(12) + (1-t)P(13)) = \begin{bmatrix} 0 & t & (1-t) \\ t & (1-t) & 0 \\ (1-t) & 0 & t \end{bmatrix}, \qquad t \in (0,1),$$

are minimal. Namely, the characteristic polynomial of B_t equals $p(\lambda) = (\lambda - 1)(1 - 3t + 3t^2 - \lambda^2)$. The zeros of the second factor equal

$$\lambda_{2,3} = \pm\sqrt{3\left(t - \frac{1}{2}\right)^2 + \frac{1}{4}}$$

and are clearly distinct for each real t, and also distinct from 1 if $t \in (0,1)$. Hence, for $t \in (0,1)$ the matrices B_t have three distinct real eigenvalues (one of them is 1), so they are nonderogatory and hence minimal.

3. Maximal centralizers within doubly stochastic matrices

Within this section we will investigate the following more restricted problem: What are the doubly stochastic matrices with maximal centralizer within the set of doubly stochastic matrices? We will require the following well-known fact (recall that $\mathbf{1}$ denotes the column vector full of ones). We give a hint of the proof for the sake of convenience. We acknowledge that the main idea came from http://math.stackexchange.com/questions/70569/span-of-permutation-matrices.

Lemma 3.1. *The algebra generated by doubly stochastic matrices is the same as their linear span and equals the space of all matrices with $\mathbf{1}$ as a right and $\mathbf{1}^T$ as a left eigenvector.*

Proof. The first part follows since doubly-stochastic matrices are a convex hull of permutation matrices (see [1]), which are closed under multiplication. For the last part, a doubly stochastic matrix has $\mathbf{1}$ as a right and $\mathbf{1}^T$ as a left eigenvector. Same then holds for each matrix from $\mathrm{Lin}_{\mathbb{R}}(\Omega_n)$ (here and throughout, given a set $\Xi \subseteq M_n$, we denote by $\mathrm{Lin}_{\mathbb{R}}(\Xi)$ the smallest real vector subspace of M_n which contains Ξ) which bounds above its dimension to $\dim \mathrm{Lin}_{\mathbb{R}}(\Omega_n) \le 1 + (n-1)^2$. Conversely, one can check that the $1 + (n-1)^2$ permutation matrices coming from the permutations id, $(1, r)$, $(1, r, s)$ for $1 \ne r \ne s \ne 1$, are linearly independent. $\square$

With the help of the previous result we can prove the following lemma, which will be used several times in the sequel.

Lemma 3.2. *There exists orthogonal matrix $U \in M_n(\mathbb{R})$ such that*

$$U^T \Omega_n U \subseteq 1 \oplus M_{n-1}(\mathbb{R}).$$

Moreover,

$$U^T (\mathrm{Lin}_{\mathbb{R}}\, \Omega_n) U = \mathbb{R} \oplus M_{n-1}(\mathbb{R}).$$

Proof. There exists a real orthogonal matrix U whose first column equals $\frac{1}{\sqrt{n}}\mathbf{1}$. Hence, if $e_1, \ldots, e_n$ is a standard basis of column vectors in $\mathbb{R}^n$ then $Ue_1 = \frac{1}{\sqrt{n}}\mathbf{1}$. Since $U^T = U^{-1}$ we see that if A is doubly stochastic, then $U^T A U e_1 = e_1$ and $e_1^T U^T A U = e_1^T$, wherefrom $U^T A U \in 1 \oplus M_{n-1}(\mathbb{R})$.

To prove the last equality of the Lemma, note that every matrix $A \in \mathbb{R} \oplus M_{n-1}(\mathbb{R})$ has (e_1^T, e_1) as a left/right eigenvector, hence UAU^T has $(\frac{1}{\sqrt{n}}\mathbf{1}^T, \frac{1}{\sqrt{n}}\mathbf{1})$ as a left/right eigenvector. The equality then follows from Lemma 3.1. $\square$

Corollary 3.3. *A doubly stochastic matrix $A \in M_n$, $n \geq 2$, commutes with every doubly stochastic matrix if and only if*

$$A = d\mathbb{1}_n + (1 - nd)I, \qquad d \in \left[0, \tfrac{1}{(n-1)}\right]. \tag{3}$$

Proof. Let U be a unitary such that $U^T \Omega_n U \subseteq 1 \oplus M_{n-1}(\mathbb{R})$. If $A \in \Omega_n$ commutes with Ω_n then clearly $U^T AU$ commutes with every matrix from $U^T(\mathrm{Lin}_{\mathbb{R}}\, \Omega_n)U = \mathbb{R} \oplus M_{n-1}(\mathbb{R})$, wherefrom $U^T AU = 1 \oplus (\lambda I_{n-1}) = \lambda I_n + (1 - \lambda)E_{11}$ for some scalar λ (here, E_{ij} denotes the standard matrix unit). Thus, $A = U(\lambda I + (1 - \lambda)E_{11})U^T = \lambda I + \frac{(1-\lambda)}{n}\mathbb{1}_n$. This is entrywise nonnegative if and only if $d := \frac{1-\lambda}{n} \in [0, \frac{1}{n-1}]$. $\square$

Here is a restatement of the previous Corollary.

Corollary 3.4. *For a doubly stochastic matrix A one has $\mathcal{C}(A) \cap \Omega_n = \Omega_n$ if and only if $A = d\mathbb{1} + (1 - nd)I$ for some $d \in \left[0, \frac{1}{(n-1)}\right]$.*

As another immediate corollary, obtained after choosing $d = \frac{1}{n}$ in Corollary 3.3:

Corollary 3.5. *If a doubly stochastic matrix $A \in \Omega_n$ commutes with $B \in \mathrm{Lin}_{\mathbb{R}}\, \Omega_n$, then A also commutes with $\frac{1}{n}\mathbb{1} - \beta B$ for every real β.*

The lemma below is crucial to classify which doubly stochastic matrices have maximal doubly stochastic centralizer.

Lemma 3.6. *Let A, B be doubly stochastic. Then $\mathcal{C}(A) \cap \Omega_n \subseteq \mathcal{C}(B) \cap \Omega_n$ if and only if $\mathcal{C}(A) \cap \mathrm{Lin}_{\mathbb{R}}(\Omega_n) \subseteq \mathcal{C}(B) \cap \mathrm{Lin}_{\mathbb{R}}(\Omega_n)$.*

Proof. Assume $\mathcal{C}(A) \cap \Omega_n \subseteq \mathcal{C}(B) \cap \Omega_n$. Take any $X \in \mathcal{C}(A) \cap \mathrm{Lin}_{\mathbb{R}}\, \Omega_n$. Then, X is a real matrix so $\hat{X} = \frac{1}{n}\mathbb{1} - \lambda X$ has all entries positive if $\lambda \in \mathbb{R}$ is sufficiently small, nonzero. As such, $\hat{X}$ is doubly-stochastic, and hence, by Corollary 3.5, $\hat{X} \in \mathcal{C}(A) \cap \Omega_n$. By the assumptions then also $\hat{X} \in \mathcal{C}(B)$ wherefrom, again by Corollary 3.5, $X \in \mathcal{C}(B) \cap \mathrm{Lin}_{\mathbb{R}}(\Omega_n)$. The converse implication is trivial. $\square$

We say that a doubly stochastic matrix B has maximal centralizer within Ω_n if

$$\mathcal{C}_{\Omega_n}(B) := \mathcal{C}(B) \cap \Omega_n = \Omega_n.$$

Such matrices were classified in Corollary 3.5.

We say that B has strictly maximal centralizer within Ω_n if $\mathcal{C}_{\Omega_n}(B) \subsetneq \Omega_n$ and for every doubly stochastic X we have that $\mathcal{C}_{\Omega_n}(B) \subsetneq \mathcal{C}_{\Omega_n}(X)$ implies $\mathcal{C}_{\Omega_n}(X) = \Omega_n$. Below we classify those matrices.

Theorem 3.7. *A doubly stochastic matrix A has strictly maximal centralizer within Ω_n if and only if $A = d\mathbb{1} + (1 - nd)I + \mu Q$ where $Q \in M_n(\mathbb{R})$ is a nontrivial idempotent with vanishing row and column sums and $d, \mu \in \mathbb{R}$ are chosen so that $A \geq 0$, entrywise.*

Proof. Again choose a real orthogonal matrix U such that $U^T\Omega_nU \subseteq 1 \oplus M_{n-1}(\mathbb{R})$. Assume first $A \in \Omega_n$ has strictly maximal centralizer within Ω_n. Choose any matrix $U^TXU \in \mathbb{R} \oplus M_{n-1}(\mathbb{R})$ such that

$$\mathcal{C}(U^TAU) \cap (\mathbb{R} \oplus M_{n-1}(\mathbb{R})) \subsetneqq \mathcal{C}(U^TXU) \cap (\mathbb{R} \oplus M_{n-1}(\mathbb{R})). \tag{4}$$

By subtracting a suitable scalar matrix we can achieve that $U^TXU \in 0 \oplus M_{n-1}(\mathbb{R})$ while not affecting the property (4). Since $U \in M_n(\mathbb{R})$ we have that $X \in M_n(\mathbb{R})$. Hence, there exists a small enough positive number λ such that $X_1 = \frac{1}{n}\mathbb{1}_n - \lambda X$ has nonnegative entries. Recall that $U^T\mathbf{1}\mathbf{1}^TU = U^T\mathbb{1}_nU = nE_{11}$ so $U^TX_1U \in 1 \oplus M_{n-1}(\mathbb{R})$. Therefore, $X_1\mathbf{1} = \mathbf{1}$ and $\mathbf{1}^TX_1 = \mathbf{1}^T$ and hence X_1 is doubly stochastic. Clearly, (4) holds for X_1 as well (because $U^TX_1U = E_{11} - \lambda U^TXU$) and is equivalent to

$$\mathcal{C}(A) \cap \mathrm{Lin}_{\mathbb{R}}(\Omega_n) \subsetneqq \mathcal{C}(X_1) \cap \mathrm{Lin}_{\mathbb{R}}(\Omega_n).$$

Lemma 3.6 then implies $\mathcal{C}_{\Omega_n}(A) \subsetneqq \mathcal{C}_{\Omega_n}(X_1)$ which due to strict maximality of A and Corollary 3.4 implies that $X_1 = d\mathbb{1}_n + (1 - nd)I$ for suitable $d \geq 0$. We deduce that U^TAU has maximal centralizer within $\mathbb{R} \oplus M_{n-1}(\mathbb{R})$.

Conversely, if $U^TAU \in \mathbb{R} \oplus M_{n-1}(\mathbb{R})$ has maximal centralizer within $\mathbb{R} \oplus M_{n-1}(\mathbb{R})$ then, after subtracting form A a suitable scalar, multiplying the resulting matrix with small enough $\lambda > 0$ and adding $\frac{1}{n}\mathbb{1}_n$ we obtain a doubly stochastic matrix A_1 with the same relative centralizer as A and by Lemma 3.6 A_1 has strictly maximal centralizer within Ω_n.

Therefore, it suffices to classify $A \in \Omega_n$ such that $\mathcal{C}(U^TAU) \cap (\mathbb{R} \oplus M_{n-1}(\mathbb{R}))$ is maximal within $\big(\mathbb{R} \oplus M_{n-1}(\mathbb{R})\big)$ but not equal to $\big(\mathbb{R} \oplus M_{n-1}(\mathbb{R})\big)$. With the help of Proposition 1.2 it is then easy to see that the solution is $U^TAU = 1 \oplus (\mu P_1 + \nu I_{n-1}) = \nu I + (1 - \nu)E_{11} + \mu P$ for some nontrivial idempotent $P = 0 \oplus P_1 \in 0 \oplus M_{n-1}(\mathbb{R})$ and suitably chosen real scalars μ, ν. Since $Pe_1 = 0$ and $e_1^TP = 0$ we get that $A = d\mathbb{1} + (1 - nd)I + \mu Q$ where $d = \frac{1-\nu}{n}$ and where $Q = UPU^T \in M_n(\mathbb{R})$ is a nontrivial idempotent with $Q\mathbf{1} = 0$ and $\mathbf{1}^TQ = 0$ (i.e., with vanishing row and column sums). $\square$

Remark 3.8. The characterization of minimal doubly stochastic matrices is more involved, since the set of minimal doubly stochastic matrices is open within Ω_n.

To see this it suffices to show that the set of all nonderogatory matrices of M_n is open in M_n, or equivalently that the set of derogatory matrices is closed in M_n. Assume otherwise. Then, there would exist a nonderogatory matrix A which would be a limit of a sequence of derogatory matrices A_m. Passing to a subsequence we could achieve that all their minimal polynomials would have a fixed degree $k < n$. Hence, minimal polynomials of a A_m equal

$$f_m(x) = (x - \lambda_{1m})^{i_{1m}} \dots (x - \lambda_{tm})^{i_{t,m}}$$

for some index $t = t(m) \in \{1, \dots, k\}$ and exponents $i_{1,m}, \dots, i_{tm}$ with sum equal to k. Since $|\lambda_{im}| \leq \|A_m\| \xrightarrow{m\to\infty} \|A\|$ we see that the coefficients of minimal polynomials are bounded. Passing again to a subsequence we can achieve that all coefficients of monic polynomials $f_m(x)$ converge so that minimal polynomials of A_m also converge to some polynomial $f(x)$ of degree k. By continuity, $f(A) = \lim f_m(A_m) = 0$, hence A is not nonderogatory, a contradiction.

So, characterization of minimal doubly stochastic matrices would have to include, for example the matrices of the form given in (5) below. Let $n = 3$ and $P := \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$. Observe that P is cyclic, hence nonderogatory and so minimal. Since the set of minimal matrices within Ω_3 is open, we see that

$$\begin{bmatrix} a & b & 1-a-b \\ 1-c-d & c & d \\ d-a+c & 1-b-c & a+b-d \end{bmatrix} \tag{5}$$

is a minimal doubly stochastic matrix provided that a, b, c, d are sufficiently close to 0 and each entry is nonnegative.

References

[1] R.A. Brualdi, Combinatorial Classes of Matrices, Encyclopedia of Mathematics and its Applications, vol. 108, Cambridge University Press, Cambridge, 2006.

[2] R. Brualdi, Some applications of doubly stochastic matrices, Linear Algebra Appl. 107 (1988) 77–100.

[3] J.L. Doob, Topics in the theory of Markov chains, Trans. Amer. Math. Soc. 52 (1942) 37–64.

[4] G. Dolinar, A. Guterman, B. Kuzma, P. Oblak, Extremal matrix theory, Linear Algebra Appl. 438 (2013) 2904–2910.

[5] S.-G. Hwang, S.-S. Pyo, Doubly stochastic matrices whose powers eventually stop, Linear Algebra Appl. 330 (2001) 25–30.

[6] N. Jacobson, Lectures in Abstract Algebra: Linear Algebra, vol. II, Graduate Texts in Mathematics, vol. 31, Springer Verlag, Berlin, 1953.

[7] D. Luck, Doubly stochastic matrices in quantum mechanics, Found. Phys. 27 (1997) 1085–1104.

[8] A.W. Marshall, I. Olkin, B.C. Arnold, Inequalities: Theory of Majorization and its Applications, second edition, Springer, New York, 2011.

[9] P. Šemrl, Non-linear commutativity preserving maps, Acta Sci. Math. (Szeged) 71 (2005) 781–819.

[10] E. Senata, Non-Negative Matrices and Markov Chains, 2nd ed., Springer Verlag, New York, 1980.

[11] R. Sinkhorn, A relationship between arbitrary positive matrices and doubly stochastic matrices, Ann. Math. Stat. 35 (2) (1964) 876–879.

[12] R. Sinkhorn, Matrices which commute with doubly stochastic matrices, Linear Multilinear Algebra 4 (1976) 201–203.

[13] D.A. Suprunenko, R.A. Tyshkevich, Commutative Matrices, Nauka i Technika, Minsk, 1966.

[14] O. Taussky, Commutativity of finite matrices, Amer. Math. Monthly 64 (1957) 239–255.

[15] Adrian R. Wadsworth, The algebra generated by two commuting matrices, Linear Multilinear Algebra 27 (1990) 159–162.

[16] J.H.M. Wedderburn, Lectures on Matrices, American Mathematical Society Colloquium Publications, 1934.