Štefan Schwarz

A Note on the Structure of the Semigroup of Doubly-Stochastic Matrices

Matematický časopis, Vol. 17 (1967), No. 4, 308--316

Persistent URL: `http://dml.cz/dmlcz/127001`

Matematický časopis 17 (1967), No. 4

A NOTE ON THE STRUCTURE OF THE SEMIGROUP OF DOUBLY-STOCHASTIC MATRICES

ŠTEFAN SCHWARZ, Bratislava

An $n \times n$ matrix $P = (p_{ik})$ is called stochastic if $p_{ik} \geqq 0$ and $\sum_{k=1}^{n} p_{ik} = 1$ (for $i = 1,2, \ldots, n$). If moreover $\sum_{i=1}^{n} p_{ik} = 1$ (for $k = 1, 2, \ldots, n$), the matrix is called doubly-stochastic.

Since the product of two stochastic [doubly-stochastic] matrices is again a stochastic [doubly-stochastic] matrix, the set $\mathfrak{S}_n$ of all stochastic and the set $\mathfrak{D}_n$ of all doubly-stochastic matrices are semigroups. Clearly $\mathfrak{D}_n \subset \mathfrak{S}_n$, for $n > 1$ $\mathfrak{D}_n \neq \mathfrak{S}_n$.

Introduce in $\mathfrak{S}_n$ [and $\mathfrak{D}_n$ respectively] a natural topology by the requirement $P^{(n)} = (p_{ik}^{(n)}) \to P = (p_{ik})$ if and only if $p_{ik}^{(n)} \to p_{ik}$. The sets $\mathfrak{S}_n$ and $\mathfrak{D}_n$ become compact Hausdorff semigroups.

In paper [1] we have studied the structure of $\mathfrak{S}_n$ and, in particular, we have shown that the fundamental results concerning Markov chains follow from the general theory of compact semigroups.

The present paper contains some notes concerning the structure of $\mathfrak{D}_n$ $(n > 1)$. First: In contradistinction to $\mathfrak{S}_n$ $(n > 1)$ the semigroup $\mathfrak{D}_n$ contains only a finite number of idempotents. Secondly: If I is an idempotent matrix $\in \mathfrak{S}_n$ of the rank s it has been shown in [1] that the maximal group $G_0(I)$ belonging to I is isomorphic to the symmetric group of s letters. This is not true in $\mathfrak{D}_n$. The maximal groups belonging to two different idempotents of the same rank s need not be isomorphic.

Some further comments on the structure of $\mathfrak{D}_n$ are given.

1. THE IDEMPOTENTS $\in \mathfrak{D}_n$

Lemma 1. *A doubly-stochastic matrix is either irreducible or completely reducible into irreducible doubly-stochastic matrices.*

Proof. Suppose that $P = (p_{ik})$ is a reducible doubly-stochastic $n \times n$ matrix, i. e. there is a permutation matrix W such that

$$W^{-1}PW = \begin{pmatrix} A_1 & 0 \\ B & A_2 \end{pmatrix},$$

where A_1 and A_2 are square matrices of orders $s > 0$ and $n - s > 0$ respectively and B is a rectangular $(n - s) \times s$ matrix. We shall show that all elements of B are zeros.

Write $W^{-1}PW = (x_{ik})$. By supposition we have for $1 \leqq k \leqq n$

$$1 = \sum_{i=1}^{s} x_{ik} + \sum_{i=s+1}^{n} x_{ik} .$$

By summing the first s equations we get

$$s = \sum_{k=1}^{s} \sum_{i=1}^{s} x_{ik} + \sum_{k=1}^{s} \sum_{i=s+1}^{n} x_{ik} .$$

Now for any i with $1 \leqq i \leqq s$ we have by supposition $\sum_{k=1}^{s} x_{ik} = 1$, so that $\sum_{i=1}^{s} \sum_{k=1}^{s} x_{ik} = s$. Hence $\sum_{k=1}^{s} \sum_{i=s+1}^{n} x_{ik} = 0$. Since $x_{ik} \geqq 0$, we conclude $x_{ik} = 0$ for $i = s + 1, \ldots, n$ and $k = 1,2, \ldots, s$.

In the matrix $W^{-1}PW = \text{diag}\,(A_1, A_2)$ both matrices A_1, A_2 are doubly-stochastic. If for instance A_1 is reducible, we may apply the same argument, which shows that A_1 is completely reducible. Repeating this process we obtain Lemma 1.

Lemma 2. *There exists a unique irreducible idempotent $r \times r$ doubly-stochastic matrix, namely the matrix $A = (a_{ik})$ with all a_{ik} equal to the number $\frac{1}{r}$.*

Proof. It is well-known that a non-negative $r \times r$ matrix A is irreducible if and only if $A + A^2 + \ldots + A^r$ is positive. If A is an idempotent, then $A = A^2$, hence an irreducible idempotent matrix is necessarily positive.

For $i = 1,2, \ldots, r$ denote by $\varrho(i)$ the least integer j such that $a_{ji} = \min\,(a_{1i}, a_{2i}, \ldots, a_{ri})$. Since A is an idempotent,

$$a_{\varrho(i),i} = \sum_{k=1}^{r} a_{\varrho(i),k} a_{k,i}.$$

With respect to $1 = \sum_{k=1}^{r} a_{\varrho(i),k}$ this can be written in the form

$$\sum_{k=1}^{r} a_{\varrho(i),k} \, [a_{ki} - a_{\varrho(i),i}] = 0 .$$

Since $a_{\varrho(i),k} > 0$ and $a_{ki} - a_{\varrho(i),i} \geqq 0$, we have $a_{k,i} = a_{\varrho(i),i}$ for $k = 1,2, \ldots, r$. Further $\sum_{k=1}^{r} a_{ki} = 1$ (for every i) implies $r \,.\, a_{\varrho(i),i} = 1$. Hence $a_{ik} = a_{\varrho(i),i} = \frac{1}{r}$ for any i and any k. This proves our statement.

Let now I be any idempotent $\in \mathfrak{D}_n$. By Lemma 1 the matrix I is either irreducible or completely reducible into irreducible doubly-stochastic matrices, i. e. there is a permutation matrix W such that $W^{-1}IW = \mathrm{diag}(Q_1, Q_2, \ldots, Q_s)$, where Q_i are irreducible matrices. This implies the following result:

Theorem 1. *Any idempotent $I \in \mathfrak{D}_n$ is of the form $I = W^{-1}UW$, where W is a permutation matrix and U is a matrix of the form*

$$U = \begin{pmatrix} Q_1 & 0 & \ldots & 0 \\ 0 & Q_2 & \ldots & 0 \\ \vdots & & & \\ 0 & 0 & \ldots & Q_s \end{pmatrix}.$$

Here Q_i is a $r_i \times r_i$ square matrix with all elements equal to $\frac{1}{r_i}$ and $r_1 + r_2 + \ldots + r_s = n$. Conversely: Every matrix of this form is an idempotent $\in \mathfrak{D}_n$ and it is of the rank s.

Corollary. *$\mathfrak{D}_n$ contains only a finite number of idempotents.*

By choosing suitably the permutation matrix W we can obtain that in the expression for U we have $r_1 \geqq r_2 \geqq \ldots \geqq r_s$.

If U contains α_1 matrices of order ϱ_1, α_2 matrices of order $\varrho_2, \ldots, \alpha_\sigma$ matrices of order ϱ_σ, we shall say that I is of the type $(\varrho_1^{\alpha_1}, \varrho_2^{\alpha_2}, \ldots, \varrho_\sigma^{\alpha_\sigma})$. Hereby we may suppose $\varrho_1 > \varrho_2 > \ldots > \varrho_\sigma$ and we have $\alpha_1 + \alpha_2 + \ldots + \alpha_\sigma = s$, $\alpha_1\varrho_1 + \alpha_2\varrho_2 + \ldots + \alpha_\sigma\varrho_\sigma = n$.

To find all idempotents $\in \mathfrak{D}_n$ it is sufficient to find all partitions of n into non necessarily different summands, and after constructing the matrix U to apply all permutation matrices W (which, of course, need not necessarily lead to different idempotents $\in \mathfrak{D}_n$).

Example. To find all idempotents $\in \mathfrak{D}_3$ we consider the partitions $3 = 2 + 1 = 1 + 1 + 1$. There is one idempotent of the type (3^1), namely the matrix

$$I_0 = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix},$$

which is the zero element of $\mathfrak{D}_3$. There is a unique idempotent of the type (1^3), namely

$$I_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

which is the unit element of $\mathfrak{D}_3$. Finally there are three different idempotents of the type $(2^1, 1^1)$. These are the matrices

$$I_2' = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad I_2'' = \begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix}, \quad I_2''' = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

Hence $\mathfrak{D}_3$ contains exactly 5 different idempotents.

2. MAXIMAL GROUPS

We shall now study the maximal group $G(I)$ belonging to a given idempotent $I \in \mathfrak{D}_n$.

We retain the notations from Theorem 1. If $I = W^{-1}UW$, then it is easy to see that $G(I) = W^{-1}G(U)W$. (Cf. [1], Lemma 8.) Hence to get informations concerning the structure of $G(I)$ it is sufficient to study the maximal group $G(U)$ belonging to an idempotent of the form

$$U = \text{diag}\,(Q_1, Q_2, \ldots, Q_s).$$

Recall that an element $P \in \mathfrak{D}_n$ is contained in the group $G(U)$ if and only if: 1. We have $PU = UP = P$. 2. There is an element $P' \in G(U)$ such that $PP' = P'P = U$ and $P'U = UP' = P'$.

A) We shall first find the form of an element $P \in \mathfrak{D}_n$ for which

$$PU = UP = P \tag{1}$$

holds.

Write

$$P = \begin{pmatrix} P_{11}, & \ldots, & P_{1s} \\ \vdots & & \\ P_{s1}, & \ldots, & P_{ss} \end{pmatrix},$$

where P_{ik} is a rectangular $r_i \times r_k$ matrix. The relation (1) implies $P_{ik} = Q_iP_{ik} = P_{ik}Q_k$. Now Q_iP_{ik} and $P_{ik}Q_k$ are $r_i \times r_k$ matrices of the forms

$$\begin{pmatrix} u_1, & u_2, & \ldots, & u_{r_k} \\ \vdots & & & \\ u_1, & u_2, & \ldots, & u_{r_k} \end{pmatrix}, \quad \begin{pmatrix} v_1 & \ldots & v_1 \\ v_2 & \ldots & v_2 \\ \vdots & & \\ v_{r_i} & \ldots & v_{r_i} \end{pmatrix}$$

respectively. Hence $u_1 = \ldots = u_{r_k} = v_1 = \ldots = v_{r_i}$ and P_{ik} is a scalar multiple of the matrix E_{ik}, where E_{ik} is the $r_i \times r_k$ matrix with all entries equal to 1.

For convenience we shall write P_{ik} in the following in both forms:

$$P_{ik} = \frac{c_{ik}}{r_k} E_{ik} = \frac{d_{ik}}{r_i} E_{ik}.$$

We have proved: If P satisfies (1), it is of the form

$$(2) \qquad P = \begin{bmatrix} \frac{c_{11}}{r_1} E_{11}, & \dots, & \frac{c_{1s}}{r_s} E_{1s} \\ \vdots & & \\ \frac{c_{s1}}{r_1} E_{s1}, & \dots, & \frac{c_{ss}}{r_s} E_{ss} \end{bmatrix} = \begin{bmatrix} \frac{d_{11}}{r_1} E_{11}, & \dots, & \frac{d_{1s}}{r_1} E_{1s} \\ \vdots & & \\ \frac{d_{s1}}{r_s} E_{s1}, & \dots, & \frac{d_{ss}}{r_s} E_{ss} \end{bmatrix}.$$

Hereby (since P is doubly-stochastic)

$$(3) \qquad \sum_{k=1}^{s} c_{ik} = \sum_{i=1}^{s} d_{ik} = 1.$$

Conversely: Direct computation shows that if P is of the form (2), and (3) holds, then $PU = UP = P$. For

$$PU = \begin{bmatrix} \frac{c_{11}}{r_1} E_{11}Q_1, & \dots, & \frac{c_{1s}}{r_s} E_{1s}Q_s \\ \vdots & & \\ \frac{c_{s1}}{r_1} E_{s1}Q_1, & \dots, & \frac{c_{ss}}{r_s} E_{ss}Q_s \end{bmatrix}$$

and with respect to

$$\frac{c_{ik}}{r_k} E_{ki}Q_i = \frac{c_{ik}}{r_k} E_{ki} \cdot \frac{1}{r_i} E_{ii} = \frac{c_{ik}}{r_i r_k} (E_{ki}E_{ii}) = \frac{c_{ik}}{r_i r_k} \cdot r_i \cdot E_{ki} = \frac{c_{ik}}{r_k} E_{ki}$$

we get $PU = P$. Analogously $UP = P$.

B) Suppose now that P is contained in $G(U)$. Then there is a matrix $P' \in G(U)$ such that $PP' = P'P = U$. The matrix P' is of the same form as P with coefficients c'_{ik}, d'_{ik} satisfying $\sum_{k=1}^{s} c'_{ik} = 1$, $\sum_{i=1}^{s} d'_{ik} = 1$.

The relation $PP' = \operatorname{diag}(Q_1, Q_2, \dots, Q_s)$ implies

$$\sum_{k=1}^{s} \frac{c_{ik}}{r_k} E_{ik} \frac{c'_{kl}}{r_l} E_{kl} = \begin{cases} \frac{1}{r_i} E_{il} & \text{for } l = i, \\ 0 \text{ (zero matrix)} & \text{for } l \neq i. \end{cases}$$

Since $E_{ik}E_{kl} = r_k E_{il}$, we have

$$\sum_{k=1}^{s} c_{ik}c'_{k} = \begin{cases} 1 & \text{for } l = i, \\ 0 & \text{for } l \neq i. \end{cases}$$

Analogously $P'P = \operatorname{diag}(Q_1, Q_2, \ldots, Q_s)$ implies

$$\sum_{k=1}^{s} c'_{ik}c_{kl} = \begin{cases} 1 & \text{for } l = i, \\ 0 & \text{for } l \neq i. \end{cases}$$

Hence the product of the matrices

$$C = \begin{pmatrix} c_{11} & \ldots & c_{1s} \\ \cdot\cdot & & \\ c_{s1} & \ldots & c_{ss} \end{pmatrix}, \qquad C' = \begin{pmatrix} c'_{11} & \ldots & c'_{1s} \\ \cdot\cdot & & \\ c'_{s1} & \ldots & c'_{ss} \end{pmatrix}$$

is the unit matrix of order s and both matrices are non-singular (of order s).

With respect to the relations $\sum_{k=1}^{s} c_{ik} = \sum_{k=1}^{s} c'_{ik} = 1$ we get

$$\sum_{k=1}^{s} c'_{ik}(1 - c_{ki}) = 0, \quad \sum_{k=1}^{s} c_{ik}(1 - c'_{ki}) = 0.$$

Since each summand is non-negative, we have

$$c'_{ki}(1 - c_{ki}) = 0, \qquad c_{ik}(1 - c'_{ki}) = 0$$

for $i, k = 1, 2, \ldots, s$. If (for some l) $c_{il} = 1$, then for all $k \neq l$ we have $c_{ik} = 0$. On the other hand, if for some i, l, we have $c_{li} < 1$, then $c'_{il}(1 - c_{li}) = 0$ implies $c'_{il} = 0$ and with respect to $c_{li}(1 - c'_{il}) = 0$ we get $c_{li} = 0$. This means: If $c_{li} < 1$, then $c_{li} = 0$. This proves that both matrices C, C' are permutation matrices of order s. By the same method it follows that the matrix $D = (d_{ik})$ is a permutation matrix of order s.

We have proved: If $P \in G(U)$, then (c_{ik}) and (d_{ik}) are permutation matrices. Now both matrices explicitly described in (2) are identical. This implies: If $c_{ik} \neq 0$ (and hence $c_{ik} = 1$), then $d_{ik} \neq 0$ (hence $d_{ik} = 1$) and we necessarily have $r_i = r_k$. Summarily:

The necessary condition in order that

$$P = \begin{bmatrix} \frac{c_{11}}{r_1} E_{11}, & \dots, & \frac{c_{1s}}{r_s} E_{1s} \\ \\ \frac{c_{1s}}{r_1} E_{s1}, & \dots, & \frac{c_{ss}}{r_s} E_{ss} \end{bmatrix}$$

belongs to $G(U)$ is that (c_{ik}) is a permutation matrix and if $c_{ik} \neq 0$, then $r_i = r_k$.

Conversely: If these conditions are satisfied, direct computation shows that $PU = UP = P$ and there is a matrix $P' \in G(U)$ such that $P'U = UP' = P'$ and $PP' = P'P = U$. Clearly if (c'_{ik}) is the inverse matrix to C it is sufficient to take for P' the matrix

$$P' = \begin{bmatrix} \frac{c'_{11}}{r_1} E_{11}, & \dots & \frac{c'_{1s}}{r_s} E_{1s} \\ \vdots & & \\ \frac{c'_{s1}}{r_1} E_{s1}, & \dots & \frac{c'_{ss}}{r_s} E_{ss} \end{bmatrix}.$$

If the numbers $r_1, r_2, \dots, r_s$ all differ from one another and $P \in G(U)$, then $c_{ik} = 0$ for all $i \neq k$ and $G(U)$ contains a unique matrix, namely U itself.

In the second „extreme case" if $r_1 = r_2 = \dots = r_s = r$, the matrix

$$\frac{1}{r} \begin{pmatrix} c_{11}E_{11}, & \dots, & c_{1s}E_{1s} \\ \vdots & & \\ c_{s1}E_{s1}, & \dots, & c_{ss}E_{ss} \end{pmatrix}$$

is contained in $G(U)$ for any permutation matrix (c_{ik}) so that the number of elements of the group $G(U)$ is $s!$.

In general the following theorem follows immediately from our considerations:

Theorem 2. *If U is an idempotent of the type $(\varrho_1^{\alpha_1}, \varrho_2^{\alpha_2}, \dots, \varrho_\sigma^{\alpha_\sigma})$, then $G(U)$ is a finite group of order $\alpha_1! \, \alpha_2! \dots \alpha_\sigma!$.*

Example. Consider the case $n = 4$. The semigroup $\mathfrak{D}_4$ contains (among others) the following two idempotents, both of rank 2:

$$I' = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}, \quad I'' = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}.$$

Here $G(I')$ is a group of order 2 which contains besides I' the matrix

$$\begin{pmatrix} 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \end{pmatrix},$$

while $G(I'')$ is a one point group containing only I'' itself.

Theorem 2 shows a striking „loss of symmetry" of $G(U)$ in comparison with $G_0(U)$ [the maximal group belonging to U in $\mathfrak{S}_n$]. In [1] we have proved that if U is of rank s, then $G_0(U)$ is isomorphic to the symmetric group of s letters. But in $\mathfrak{D}_n$ even the order of $G(U)$ depends on the partition of n into s positive summands. (See our example.) This result is rather unexpected since the set of all doubly-stochastic matrices seems to be at first glance a „much more symmetric entity" than the set of all merely stochastic matrices.

To explain the situation call — for a while — a matrix C-stochastic if it is non-negative and all the column sums are equal to 1. Denote by $\mathfrak{S}_n^*$ the semigroup of all C-stochastic matrices and by $G_0^*(U)$ the maximal group in $\mathfrak{S}_n^*$ belonging to a doubly-stochastic idempotent matrix U. Clearly $\mathfrak{D}_n = \mathfrak{S}_n \cap \mathfrak{S}_n^*$ and $U \in \mathfrak{D}_n$. The groups $G_0(U)$ and $G_0^*(U)$ considered as subgroups of the semigroup of all non-negative matrices are isomorphic. But they are not identical. The intersection $G_0(U) \cap G_0^*(U)$ is a subgroup of $\mathfrak{D}_n$ and we clearly have $G_0(U) \cap G_0^*(U) = G(U)$.

This can be illustrated by our example. Consider the idempotent I''. Then $G_0(I'')$ is a group of order 2 containing I'' and the stochastic (but not doubly-stochastic) matrix

$$\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 \end{pmatrix}.$$

Analogously $G_0^*(I'')$ contains I'' and the C-stochastic (but not doubly-stochastic) matrix

$$\begin{pmatrix} 0 & 0 & 0 & \frac{1}{3} \\ 0 & 0 & 0 & \frac{1}{3} \\ 0 & 0 & 0 & \frac{1}{3} \\ 1 & 1 & 1 & 0 \end{pmatrix}.$$

We have $G_0(I'') \cap G_0^*(I'') = I''$.

Remark (added in October 1966). After this paper had been sent to print the paper [2] appeared. It contains (in essential) the results of our paper. The proofs are, however, different.

REFERENCES

[1] Schwarz Št., *On the structure of the semigroup of stochastic matrices*, Magyar Tud. Akad. Mat. Kutató Int. Közl. *9* (1964), 297—311.

[2] Farahat H. K., *The semigroup of doubly-stochastic matrices*, Proc. Glasgow Math. Ass. *7* (1966), 178—183.

Received August 5, 1966.

Matematický ústav
Slovenskej akadémie vied,
Bratislava

ERRATUM

B. Zelinka, A CONTRIBUTION TO MY ARTICLE „INTRODUCING AN ORIENTATION INTO A GIVEN NON-DIRECTED GRAPH“, Mat. časop. *17* (1967), 142 — 145.

In Theorem 1a — 2a instead of „tree with a finite diameter“ there should be „tree without infinite peths“.