CONVEX MATRIX EQUATIONS

BY J. S. MONTAGUE AND R. J. PLEMMONS

Communicated by Eugene Isaacson, May 15, 1972

1. Introduction. Let $\Delta_n$ denote the set of all $n \times n$ complex matrices $A$ whose spectral norm $\|A\|$ is at most one. Then $\Delta_n$ forms a convex topological semigroup under matrix multiplication ([6], [7]). The subsemigroup $\Sigma_n$ of $\Delta_n$, consisting of all real nonnegative matrices in $\Delta_n$, is the set of all $n \times n$ doubly substochastic matrices; that is, real nonnegative matrices whose row and column sums are at most one. The subsemigroup of $\Sigma_n$ consisting of all $n \times n$ doubly stochastic matrices will be denoted by $\Omega_n$.

Geometrically, $\Omega_n$ is the convex hull of the group of all $n \times n$ permutation matrices ([1], [8]), while $\Sigma_n$ is the convex hull of the subsemigroup of all $n \times n$ subpermutation matrices [9]. The following theorem establishes a similar result for $\Delta_n$.

THEOREM 1. $\Delta_n$ is the convex hull of the set of all $n \times n$ unitary matrices.

The proof of the theorem can be obtained by establishing that the unitary matrices form the set of extreme points of $\Delta_n$. The result then follows by the Krein-Milman theorem. The complete proof will appear elsewhere [10]. Another proof of this result is given in [15].

Several authors have considered matrix equations involving doubly stochastic matrices. In particular, S. Sherman [14] and S. Schreiber [13] have considered the solvability of the equation $AX = B$ and D. J. Hartfiel [5] has considered the solvability of the equation $AXB = X$, where $A$, $B$, and $X$ are doubly stochastic. The main purpose of this note is to consider the system of matrix equations

\begin{equation}
AX = B \quad \text{and} \quad BY = A,
\end{equation}


Key words and phrases. Convex hull, convex topological semigroup, doubly stochastic matrix, Green's relations, Moore-Penrose generalized inverse, regularity.

1 Research supported in part by the National Science Foundation grant GP-15943.
where $A$ and $B$ are arbitrary complex or real $m \times n$ matrices and where $X$ and $Y$ are in $\Delta_n$, $\Sigma_n$ or $\Omega_n$. These ideas are then used in §§3 and 4 to investigate the Green's relations and regularity in those semigroups.

2. The equations $AX = B$, $BY = A$. The following theorems characterize the solvability of the equations (1.1) over $\Delta_n$, $\Sigma_n$ and $\Omega_n$ in terms of solvability over certain matrix groups.

**Theorem 2.** For arbitrary $m \times n$ complex matrices $A$ and $B$, the equations (1.1) are solvable for $X, Y \in \Delta_n$ if and only if $A = BU$ for some unitary matrix $U$.

**Theorem 3.** For arbitrary $m \times n$ real matrices $A$ and $B$, the equations (1.1) are solvable for $X, Y \in \Sigma_n$ [$\Omega_n$] if and only if $A = BP$ for some permutation matrix $P$.

The proof of Theorem 2 is based on a theorem of Witt [16] and is fairly straightforward. Although it might be expected that the proof of Theorem 3 would follow from or be similar to the proof of Theorem 2, an entirely different approach is apparently needed. The proof for $\Sigma_n$ and $\Omega_n$ is based in part on a theorem of Hardy, Littlewood and Pólya [4, Theorem 4.6]. These solvability theorems will now be used to investigate the algebraic structures of the convex semigroups $\Delta_n$, $\Sigma_n$ and $\Omega_n$.

3. The Green's relations. The Green's relations $\mathcal{R}$, $\mathcal{L}$, $\mathcal{J}$, $\mathcal{H}$, and $\mathcal{D}$ play a fundamental role in the study of the algebraic structure of semigroups. For an arbitrary semigroup $S$ with $a, b \in S$, the relation $\mathcal{R}$ [$\mathcal{L}$, $\mathcal{J}$] is defined by $a \mathcal{R} b$ [$a \mathcal{L} b, a \mathcal{J} b$] if and only if $a$ and $b$ generate the same principal right [left, two-sided] ideal in $S$. Then $\mathcal{H} = \mathcal{R} \cap \mathcal{L}$ and $\mathcal{D}$ is defined to be the join $\mathcal{R} \lor \mathcal{L}$ [2, Chapter 3]. The problem of characterizing the Green's relations on $\Delta_n$, $\Sigma_n$ and $\Omega_n$ can be solved by characterizing solutions to certain matrix equations. The results in this section follow from Theorems 2 and 3, together with their duals obtained by taking transposes in the equations (1.1). Notice that $\mathcal{D} = \mathcal{J}$ on these semigroups since they are compact [6].

**Theorem 4.** Let $A$ and $B$ belong to the semigroup $\Delta_n$. Then

(i) $A \mathcal{R} B$ if and only if $A = BU$ for some unitary matrix $U$;
(ii) $A \mathcal{L} B$ if and only if $A = VB$ for some unitary matrix $V$;
(iii) $A \mathcal{H} B$ if and only if $A = BU = VB$ for some unitary matrices $U, V$;
(iv) $A \mathcal{D} B$ if and only if $A = VBU$ for some unitary matrices $U$ and $V$.

**Theorem 5.** Let $A$ and $B$ belong to the semigroup $\Sigma_n$ of doubly substochastic matrices. Then
(i) $A \mathcal{R} B$ if and only if $A = BP$ for some permutation matrix $P$;
(ii) $A \mathcal{L} B$ if and only if $A = QB$ for some permutation matrix $Q$;
(iii) $A \mathcal{H} B$ if and only if $A = BP = QB$ for some permutation matrices $P$ and $Q$;
(iv) $A \mathcal{D} B$ if and only if $A = QBP$ for some permutation matrices $P$ and $Q$.

It follows that Theorem 5 also characterizes the Green's relations on the semigroup $\Omega_n$ of doubly stochastic matrices [11], since the permutation matrices are contained in $\Omega_n$. Moreover, these characterizations can be used to show that the maximal subgroups of $\Delta_n$ are isomorphic to full unitary groups, while $\Sigma_n$ and $\Omega_n$ have finitely many maximal subgroups, each of which is isomorphic to a direct product of full symmetric groups [3].

4. Regularity. An element $a$ in a semigroup $S$ is said to be regular if the equation $a = axa$ is solvable for $x \in S$. If in addition $x = xax$, then $a$ and $x$ are said to be semi-inverses. In this section the regular elements in $\Delta_n$, $\Sigma_n$ and $\Omega_n$ are investigated. Clearly not every matrix in $\Delta_n$ is regular. In particular, the only nonsingular regular members of $\Delta_n$ are the unitary matrices. The following concepts will facilitate the characterizations of regularity.

The singular values of an $n \times n$ complex matrix $A$ are the positive square roots of the eigenvalues of $A^*A$, where $A^*$ denotes the conjugate transpose of $A$. Now $A$ is a partial isometry if the linear transformation represented by $A$ preserves distances on the range of $A^*$. The Moore-Penrose generalized inverse [12] of $A$ is the unique matrix $A^+$ defined by $A^+y = x$ if $Ax = y$ and $x$ is in the range of $A^*$, and $A^+y = 0$ if $A^*y = 0$.

Now it can be shown [10] that $E \in \Delta_n$ is idempotent if and only if there is a unitary matrix $U$ such that $UEU^*$ is diagonal with 0's and 1's on the diagonal. This fact, together with the results in §3, lead to the following characterizations of regularity. The details of the proof will appear elsewhere ([10], [11]).

**Theorem 6.** Let $A \in \Delta_n [\Sigma_n, \Omega_n]$. Then the following statements are equivalent.

(i) $A$ is regular in $\Delta_n [\Sigma_n, \Omega_n]$.
(ii) $A^*$ is the unique semi-inverse of $A$ in $\Delta_n [\Sigma_n, \Omega_n]$.
(iii) The singular values of $A$ are 0 and 1.
(iv) $A$ is a partial isometry.
(v) $\|A\| = 1$ if $A$ is nonzero.
(vi) $A^+ = A^*$. 

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use
REFERENCES


DEPARTMENT OF MATHEMATICS AND COMPUTER SCIENCE, UNIVERSITY OF TENNESSEE, KNOXVILLE, TENNESSEE 37916