# A Spectral Characterization of Stochastic Matrices*

RICHARD A. BRUALDI†
*University of Wisconsin*
*Madison, Wisconsin*

HELMUT W. WIELANDT
*University of Wisconsin, Madison, Wisconsin, and*
*University of Tübingen, Tübingen, Germany*

Communicated by Hans Schneider

## 1. INTRODUCTION

Let $F$ be a field with identity element 1, and $A$ an $n \times n$ matrix over $F$. If $A$ is either row or column stochastic (that is, all row sums or all column sums equal 1), then 1 is a characteristic root of $A$. Moreover, whenever $A$ is row or column stochastic so is $PA$ where $P$ is a $n \times n$ permutation matrix; thus 1 is a characteristic root of $PA$ for every $n \times n$ permutation matrix $P$. It is shown here that this property actually characterizes stochastic matrices over $F$ (Theorem 1). Matrices $A$ over $F$ for which the spectrum of $A$ is the same as the spectrum of $PA$ for every $n \times n$ permutation matrix $P$ are also characterized (Theorem 2).

## 2. MAIN RESULT

The vector space of $n$-tuples over $F$, written as columns, is denoted by $F^n$. If $y \in F^n$, then $y^i$ denotes the $i$th coordinate of $y$ ($i = 1, 2, \ldots, n$).

* Sponsored by the Mathematics Research Center, United States Army, Madison, Wisconsin, under Contract No. DA-31-124-ARO-D-462.

† The research of this author was partially supported by NSF Grant No. GP-3993.

THEOREM 1. *Let $A$ be an $n \times n$ matrix over a field $F$ with identity element* 1. *If* 1 *is a characteristic root of $PA$ for all $n \times n$ permutation matrices $P$, then $A$ is either row stochastic or column stochastic.*

*Proof.* If $P$ is any $n \times n$ permutation matrix, then there exists a nonzero vector $x_P \in F^n$ such that $PAx_P = x_P$. For $P$ fixed let

$$X_P = \{x : x \in F^n, PAx = x\}.$$

Each vector space $X_P$ has dimension at least one. If $I$ denotes the $n \times n$ identity matrix over $F$, then we may assume without loss of generality that $X_I$ has the least dimension of all the $X_P$ (replace $A$ by a suitable $QA$, $Q$ an $n \times n$ permutation matrix).

Let $x_I$ be a nonzero vector in $X_I$. If $x_I^1 = x_I^2 = \cdots = x_I^n$, then $A$ is row stochastic. Hence we may assume that not all the coordinates of $x_I$ are equal. After possibly replacing $A$ by a suitable $RAR^T$, $R$ a permutation matrix, there exists an integer $k$ with $1 \leqslant k < n$ such that

$$x_I^i \neq x_I^j, \qquad 1 \leqslant i \leqslant k < j \leqslant n.$$

(The $X_I$ corresponding to $RAR^T$ will still have minimal dimension.) If $k > 1$, then for $2 \leqslant i \leqslant k$ let $P_i$ be the $n \times n$ permutation matrix associated with the transposition $(i, n)$. For $k + 1 \leqslant j \leqslant n$, let $Q_j$ be the $n \times n$ permutation matrix corresponding to the transposition $(1, j)$.

For $i = 2, \ldots, k$ pick $x_i \in X_{P_i}$ with $x_i^i \neq x_i^n$. This is possible, for otherwise $X_{P_i}$ is a proper subspace of $X_I$ which contradicts the minimality of the dimension of $X_I$. For $j = k + 1, \ldots, n$ pick $y_j \in X_{Q_j}$ with $y_j^1 \neq y_j^j$. Assume

$$x_I b + \sum_{i=2}^{k} x_i c_i + \sum_{j=k+1}^{n} y_j d_j = 0, \tag{1}$$

where the first summation is vacuous if $k = 1$. Apply $A$ to Eq. (1) to obtain, because of $P_i^{-1} = P_i$, $Q_j^{-1} = Q_j$,

$$x_I b + \sum_{i=2}^{k} P_i x_i c_i + \sum_{j=k+1}^{n} Q_j y_j d_j = 0. \tag{2}$$

Subtract Eq. (2) from Eq. (1) to get

$$\sum_{i=2}^{k} (x_i - P_i x_i) c_i + \sum_{j=k+1}^{n} (y_j - Q_j y_j)\, d_j = 0. \tag{3}$$

Compare the $\kappa$th component on both sides, $2 \leqslant \kappa \leqslant k$. Since $P_i$ $(i \neq \kappa)$ and $Q_j$ do not change the $\kappa$th coordinate of their corresponding vectors, (3) implies

$$(x_\kappa^{\ \kappa} - x_\kappa^{\ n})c_\kappa = 0$$

and therefore

$$c_\kappa = 0, \qquad \kappa = 2, \ldots, k. \tag{4}$$

If $k + 1 \leqslant \lambda \leqslant n - 1$, then $P_i$ and $Q_j$ $(j \neq \lambda)$ do not change the $\lambda$th coordinate of their corresponding vectors. Therefore

$$d_\lambda = 0, \qquad \lambda = k + 1, \ldots, n - 1. \tag{5}$$

The use of (4) and (5) in (3) gives

$$(y_n^{\ n} - y_n^{\ 1})\, d_n = 0$$

and hence

$$d_n = 0. \tag{6}$$

Now (4), (5), and (6) when used in Eq. (1) give $b = 0$. Hence the vectors $x_1, x_2, \ldots, x_k, y_{k+1}, \ldots, y_n$ form a set of $n$ linearly independent vectors.

Simplifying the notation we may state that there are $n$ permutation matrices $R_1, R_2, \ldots, R_n$ and $n$ linearly independent vectors $z_1, z_2, \ldots, z_n$ such that

$$R_i A z_i = z_i, \qquad i = 1, 2, \ldots, n.$$

Let $e = (1, 1, \ldots, 1)$. Then

$$e R_i A z_i = e z_i, \qquad i = 1, 2, \ldots, n$$

and therefore

$$e(A - I)z_i = 0, \qquad i = 1, 2, \ldots, n.$$

Since $z_1, z_2, \ldots, z_n$ are linearly independent,

$$e(A - I) = 0$$

or

$$eA = e.$$

That is, $A$ is column stochastic. This completes the proof of the theorem.

*Remark* 1. The proof shows a little more than the theorem states:

(i) It is sufficient to assume that $F$ is a division ring if the statement "1 is a characteristic root of $PA$" is interpreted to mean that there exists $0 \neq x \in F^n$ with $PAx = x$.

(ii) If $A$ is an $n \times n$ row stochastic matrix over $F$ and a vector space $X_P$ of smallest dimension contains a vector with two distinct coordinates, then $A$ is doubly stochastic (both row and column stochastic).

*Remark* 2. It seems difficult to find sets of less than $n!$ permutation matrices which can be used to characterize the stochastic $n \times n$ matrices in the manner of Theorem 1. For instance, in case $n = 4$ neither (i) all transpositions nor (ii) all even permutations nor (iii) all powers of a cyclic permutation will do. Counterexamples are given by

$$\text{(i, ii)} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}; \qquad \text{(iii)} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}.$$

The following corollary is an immediate consequence of Theorem 1.

COROLLARY 1. *Let $F_1$ be an extension field of $F$ and $0 \neq \alpha \in F_1$. Suppose $A$ is an $n \times n$ matrix over $F$ with $\alpha$ a characteristic root of $PA$ for all $n \times n$ permutation matrices $P$. Then actually $\alpha \in F$ and either all row sums of $A$ equal $\alpha$ or all column sums of $A$ equal $\alpha$.*

COROLLARY 2. *Let $A$ be an $n \times n$ matrix over the field $F$ and assume the characteristic of $F$ is either $0$ or relatively prime to $n$. Then two distinct nonzero elements of $F$ cannot both be characteristic roots of $PA$ for all $n \times n$ permutation matrices $P$.*

*Proof.* Suppose $\alpha \neq 0$ and $\beta \neq 0$, $\alpha \neq \beta$, are elements of $F$ both of which are characteristic roots of $PA$ for all $n \times n$ permutation matrices $P$. Then by Corollary 1 all row sums of $A$ equal $\alpha$ and all column sums of $A$ equal $\beta$ (or vice versa). The sum of the entries of $A$ via the row sums is $n\alpha$, while the sum of the entries of $A$ via the column sums is $n\beta$. Hence $n\alpha = n\beta$ and the assumptions on the characteristic of $F$ imply $\alpha = \beta$, a contradiction.

## 3. INVARIANCE OF THE SPECTRUM

We vary the problem by imposing a stronger condition on the matrix $A$. Let $A$ be an $n \times n$ matrix over an algebraically closed field $F$. If $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the $n$ characteristic roots of $A$, then the spectrum of $A$ is denoted by $\operatorname{sp} A = \{\lambda_1, \lambda_2, \ldots, \lambda_n\}$. Our question is: When is the spectrum of $A$ the same (counting multiplicities) as the spectrum of $PA$ for all $n \times n$ permutation matrices $P$? Obviously sufficient is the condition that $A$ or $A^T$ (the transpose of $A$) has the form

$$\begin{bmatrix} x_1 & x_2 & \cdots & x_n \\ x_1 & x_2 & \cdots & x_n \\ & & \vdots & \\ x_1 & x_2 & \cdots & x_n \end{bmatrix} \qquad (x_i \in F). \tag{7}$$

The following theorem gives the complete answer.

THEOREM 2. *Let $A$ be an $n \times n$ matrix ($n \geqslant 2$) over an algebraically closed field $F$.*

(i) *If the characteristic of $F$ is either 0 or relatively prime to $n$, then* $\operatorname{sp} A = \operatorname{sp} PA$ *for all $n \times n$ permutation matrices $P$ if and only if $A$ or $A^T$ has the form given in* (7).

(ii) *If the characteristic of $F$ divides $n$, then* $\operatorname{sp} A = \operatorname{sp} PA$ *for all $n \times n$ permutation matrices $P$ if and only if $A$ has the form*

$$\begin{bmatrix} \alpha_1 + \beta_1 & \alpha_1 + \beta_2 & \cdots & \alpha_1 + \beta_n \\ \alpha_2 + \beta_1 & \alpha_2 + \beta_2 & \cdots & \alpha_2 + \beta_n \\ & & \vdots & \\ \alpha_n + \beta_1 & \alpha_n + \beta_2 & \cdots & \alpha_n + \beta_n \end{bmatrix} \qquad (\alpha_i, \beta_j \in F). \tag{8}$$

*Proof.* Suppose $\operatorname{sp} A = \operatorname{sp} PA$ for all $n \times n$ permutation matrices $P$. Since the trace of a matrix equals the sum of its characteristic roots, the trace of $PA$ equals the trace of $A$ for all $P$. It then follows, for every $2 \times 2$ submatrix of $A$

$$\begin{bmatrix} a_{ij} & a_{ik} \\ a_{lj} & a_{lk} \end{bmatrix} \qquad (1 \leqslant i < l \leqslant n,\ 1 \leqslant j < k \leqslant n),$$

that $a_{ij} + a_{lk} = a_{ik} + a_{lj}$ or

$$a_{ij} - a_{ik} = a_{lj} - a_{lk} \qquad (1 \leqslant i < l \leqslant n,\ 1 \leqslant j < k \leqslant n).$$

This implies that the matrix $A$ has the form (8), and, in particular, has rank at most 2. For $X$ an $n \times n$ matrix, let $e_2(X)$ denote the second elementary symmetric function of the characteristic roots of $X$. Since for $A$ as in (8) the trace of $A$ equals the trace of $PA$ for all $n \times n$ permutation matrices $P$, and since at least $n - 2$ characteristic roots of $A$ are zero, it follows that $\operatorname{sp} A = \operatorname{sp} PA$ for all $P$ if and only if $A$ is as in (8) with $e_2(A) = e_2(PA)$ for all $n \times n$ permutation matrices $P$. From (8), we conclude that

$$\begin{aligned} e_2(A) &= \sum_{n \geqslant i > j \geqslant 1} [(\alpha_i + \beta_i)(\alpha_j + \beta_j) - (\alpha_i + \beta_j)(\alpha_j + \beta_i)] \\ &= \sum_{n \geqslant i > j \geqslant 1} (\alpha_i - \alpha_j)(\beta_j - \beta_i). \end{aligned}$$

Let $P_{k,l}$ denote the $n \times n$ permutation matrix corresponding to the transposition $(k, l)$ $(1 \leqslant k < l \leqslant n)$. A straightforward calculation shows that

$$e_2(A) - e_2(P_{k,l}A) = n(\alpha_k - \alpha_l)(\beta_l - \beta_k). \tag{9}$$

If the characteristic of $F$ is zero or relatively prime to $n$, then from $e_2(A) = e_2(PA)$ for all $n \times n$ permutation matrices $P$ we conclude by (9) that $\alpha_k = \alpha_l$ or $\beta_k = \beta_l$ $(1 \leqslant k < l \leqslant n)$. If $\alpha_1 = \alpha_2 = \cdots = \alpha_n$, then $A$ is of the form (7). Otherwise the $\alpha$'s may be partitioned into two nonempty subsets $S$ and $T$ such that $\alpha_i \in S$, $\alpha_j \in T$ imply $\alpha_i \neq \alpha_j$. This then implies $\beta_1 = \beta_2 = \cdots = \beta_n$ and $A^T$ is of the form (7).

If the characteristic of $F$ divides $n$, then from (9) we conclude that $e_2(A) = e_2(P_{k,l}A)$ $(1 \leqslant k < l \leqslant n)$. Moreover if $A$ has the form (8) then so does $QA$, for every permutation matrix $Q$; hence

$$e_2(QA) = e_2(P_{kl}QA). \tag{10}$$

Now let $P$ be any $n \times n$ permutation matrix. Then $P$ has a representation of the form

$$P = P_{k_t,l_t} \cdots P_{k_2,l_2} P_{k_1,l_1}.$$

Hence from (10) we conclude

$$e_2(A) = e_2(P_{k_1,l_1}A) = e_2(P_{k_2,l_2}P_{k_1,l_1}A) = \cdots = e_2(PA).$$

Hence, in case the characteristic of $F$ divides $n$, all matrices of the form (8) have the desired property. This completes the proof of Theorem 2.

*Received January 19, 1967*