A Characterization of Tridiagonal Matrices*

MIROSLAV FIEDLER
Czechoslovak Academy of Sciences, Institute of Mathematics
Prague, Czechoslovakia

INTRODUCTION

The purpose of this paper is to prove that symmetric irreducible tridiagonal matrices and their permutations are the only symmetric matrices (of order $n \geqslant 2$) the rank of which cannot be diminished to less than $n - 1$ by any change of diagonal elements.

The main part of the proof was obtained as a byproduct of a minimum problem solution (cf. [1]).

1. NOTATION AND CONVENTIONS

All vectors and matrices considered are real. The transpose of a matrix M is denoted by M^T, the identity matrix by I. The term vector always means column vector. The scalar product of vectors x and y is denoted by (x, y). If x is a vector, we shall denote by $||x||$ the Euclidean norm of x, i.e., $(x, x)^{1/2}$. As is well known, the corresponding operator norm of a square matrix A defined by

$$||A|| = \operatorname*{lub}_{||x|| \leqslant 1} ||Ax||$$

is equal to the square root of the maximum eigenvalue of AA^T. Especially, if A is symmetric, i.e., $A = A^T$, $||A||$ is equal to $\max_i |\lambda_i|$ where λ_i are the eigenvalues of A. For remaining usual definitions we refer to the book [2].

For our purpose it will be convenient to denote by C_n ($n \geqslant 2$ integer) the class of all symmetric matrices A of order n such that for any diagonal matrix D the rank of $A + D$ is greater than or equal to $n - 1$.

* Dedicated to Professor A. M. Ostrowski on his 75th birthday.

Linear Algebra and Its Applications 2(1969), 191–197

2. In this section, we are going to prove the main theorem (2.8). First, let us state three obvious lemmas.

LEMMA 2.1. *If $A \in C_n$ and D is diagonal then $A + D \in C_n$.*

LEMMA 2.2. *If $A \in C_n$ then all eigenvalues of A are simple.*

LEMMA 2.3. *If $A \in C_n$ then A is irreducible.*

The following lemma will be of main importance.

LEMMA 2.4. *If $A \in C_n$ then there exists a permutation matrix P such that PAP^T has the (nontrivially partitioned) form*

$$\begin{pmatrix} D_1 & B_{12} \\ B_{12}^T & D_2 \end{pmatrix}$$

where D_1, D_2 are diagonal matrices.

Proof. If $n = 2$, the assertion is true. Let now n be greater than two and let $A \in C_n$. It is easy to see that there exists a diagonal matrix D_0 such that

$$||A + D_0|| \leqslant ||A + D|| \tag{1}$$

for all diagonal matrices D. Denote $A_0 = A + D_0$. Let λ_0 be the maximum, μ_0 the minimum eigenvalue of A_0. From (1), it follows immediately that

$$\lambda_0 > 0 \tag{2}$$

and

$$\mu_0 = -\lambda_0. \tag{3}$$

The set of all matrices $P = (p_{ik})$ with the same off-diagonal elements as A forms an n-parametric system $\mathscr{P}$ depending on the diagonal elements p_{ii}. With each $P \in \mathscr{P}$ we can associate its maximum eigenvalue $\lambda(P)$ and minimum eigenvalue $\mu(P)$. According to (1), (2), and (3), we have

$$\max_{P \in \mathscr{P}}(\lambda(P) - \mu(P)) = \lambda_0 - \mu_0. \tag{4}$$

By 2.2, λ_0 and μ_0 are simple eigenvalues of A_0. Hence $\lambda(P) - \mu(P)$ is an analytic function of the diagonal elements p_{ii} in the neighborhood of A_0, i.e.,

$$\left[\frac{\partial(\lambda(P) - \mu(P))}{\partial p_{ii}}\right]_{P=A_0} = 0, \qquad i = 1, \ldots, n. \tag{5}$$

To compute these derivatives, denote $\lambda = \lambda(P)$, x the corresponding eigenvector:

$$Px = \lambda x, \qquad \|x\| = 1, \tag{6}$$

and denote by $\dot{}$ the differentiation $\partial/\partial p_{ii}$.

From (6), we get

$$\dot{P}x + P\dot{x} = \dot{\lambda}x + \lambda\dot{x}$$

so that

$$(\dot{P}x, x) + (P\dot{x}, x) = \dot{\lambda}(x, x) + \lambda(\dot{x}, x).$$

Since

$$(P\dot{x}, x) = (Px, \dot{x}) = \lambda(x, \dot{x}),$$

it follows that

$$\dot{\lambda} = (\dot{P}x, x). \tag{7}$$

Let $y = (y_k)$ and $z = (z_k)$ be normed eigenvectors of A_0 corresponding to λ_0 and μ_0:

$$A_0 y = \lambda_0 y, \qquad \|y\| = 1,$$

$$A_0 z = \mu_0 z, \qquad \|z\| = 1.$$

It follows from (7) that

$$\left[\frac{\partial\lambda(P)}{\partial p_{ii}}\right]_{P=A_0} = y_i^2$$

so that the condition (5) yields

$$y_i^2 = z_i^2, \qquad i = 1, \ldots, n. \tag{8}$$

If $N = \{1, 2, \ldots, n\}$, denote by M that subset of N consisting of all indices i for which

$$y_i = z_i;$$

it follows from (8) that if $j \in N - M$ then

$$y_j = -z_j \neq 0.$$

Since $z \neq y \neq -z$, $\emptyset \neq M \neq N$ and $y + z \neq 0$, $y - z \neq 0$.

Let m be the number of elements in M and let P be a permutation which transforms M into the set of the first m indices of N. Then, the matrix

$$B = PA_0P^T$$

has the property that for $Py = \tilde{y}$, $Pz = \tilde{z}$

$$\begin{aligned} B(\tilde{y} - \tilde{z}) &= PA_0(y - z) = P\lambda_0(y + z) \\ &= \lambda_0(\tilde{y} + \tilde{z}), \\ B(\tilde{y} + \tilde{z}) &= \lambda_0(\tilde{y} - \tilde{z}) \end{aligned}$$

according to (3). However, in the partitioned form

$$\tilde{y} + \tilde{z} = \begin{pmatrix} u \\ 0 \end{pmatrix}, \qquad \tilde{y} - \tilde{z} = \begin{pmatrix} 0 \\ v \end{pmatrix}$$

where $u \neq 0$, $v \neq 0$.

Let

$$\begin{pmatrix} B_{11} & B_{12} \\ B_{12}^T & B_{22} \end{pmatrix}$$

be the corresponding partitioning of B.

Hence

$$\begin{aligned} B_{11}u &= 0, \\ B_{12}^T u &= \lambda_0 v, \\ B_{12}v &= \lambda_0 u, \\ B_{22}v &= 0. \end{aligned} \tag{9}$$

We are now able to prove that B_{11} as well as B_{22} are zero matrices. Assume $B_{11} \neq 0$ and denote for $\sigma > 0$ by $B(\sigma)$ the matrix

$$B(\sigma) = \begin{pmatrix} \sigma^{-1}B_{11} & B_{12} \\ B_{12}^T & \sigma B_{22} \end{pmatrix}.$$

By (9), the vectors $\tilde{y}$ and $\tilde{z}$ are eigenvectors of $B(\sigma)$ with eigenvalues λ_0 and $-\lambda_0$. The remaining eigenvalues depend continuously on σ. If σ tends to zero, $B(\sigma)$ has arbitrarily large elements and since $B(\sigma)$ is symmetric, some eigenvalue has arbitrarily large modulus. On the other hand, the remaining eigenvalues of $B(1)$ are contained in the interval $(-\lambda_0, \lambda_0)$. Consequently, λ_0 or $-\lambda_0$ is a multiple eigenvalue of $B(\sigma_0)$ for some $\sigma_0 > 0$. This implies that for $\varepsilon = 1$ or $\varepsilon = -1$ the matrix

$$\begin{pmatrix} \sigma_0{}^{-1}B_{11} - \varepsilon\lambda_0 I_1 & B_{12} \\ B_{12}^T & \sigma_0 B_{22} - \varepsilon\lambda_0 I_2 \end{pmatrix}$$

has rank smaller than $n - 1$ so that

$$P^T \begin{pmatrix} B_{11} - \sigma_0\varepsilon\lambda_0 I_1 & B_{12} \\ B_{12}^T & B_{22} - \sigma_0{}^{-1}\varepsilon\lambda_0 I_2 \end{pmatrix} P$$

with the same off-diagonal elements as A has rank smaller than $n - 1$. This contradiction proves $B_{11} = 0$. Similarly, $B_{22} = 0$. Hence, PAP^T has the form as asserted.

Remark 2.5. The assertion of Lemma 2.4 can be also formulated as follows. If $A \in C_n$ then there exists a subset M of $N = \{1, 2, \ldots, n\}$, $\emptyset \neq M \neq N$, such that $a_{ik} = 0$ whenever $i \neq k$ and both i and k belong either to M or to $N - M$. This property is also called property A due to Young [3].

LEMMA 2.6. *If $A \in C_n$ then no row of A contains more than two off-diagonal elements different from zero.*

Proof. Let $A \in C_n$ be already in the form

$$\begin{pmatrix} D_1 & B_{12} \\ B_{12}^T & D_2 \end{pmatrix}$$

and assume that the matrix $B_{12} = (a_{ip})$, $i \in M = \{1, 2, \ldots, m\}$, $p \in N - M = \{m + 1, \ldots, n\}$ has three entries in the first row a_{1j}, a_{1k}, a_{1l} different from zero.

If D_1 is nonsingular, we can write

$$\begin{pmatrix} D_1 & B_{12} \\ B_{12}^T & D_2 \end{pmatrix} = \begin{pmatrix} I_1 & 0 \\ P_{21} & I_2 \end{pmatrix} \begin{pmatrix} D_1 & B_{12} \\ 0 & Q_2 \end{pmatrix}$$

where $P_{21} = B_{12}^T D_1^{-1}$, $Q_2 = D_2 - B_{12}^T D_1^{-1} B_{12}$.

It follows that for any nonsingular diagonal D_1 the matrices $D_2 - B_{12}^T D_1^{-1} B_{12}$ as well as $B_{12}^T D_1^{-1} B_{12}$ belong to C_{n-m} since otherwise the rank of A could be diminished to less than $n - 1$ by a proper change of diagonal elements. Now, it is possible to choose a nonsingular $D_1 = \operatorname{diag}\{d_i\}$, $i \in M$, in such a manner that the following entries of $B_{12}^T D_1^{-1} B_{12}$ will be different from zero:

$$\sum_{i \in M} a_{ji} a_{ki} d_i^{-1} \neq 0, \qquad \sum_{i \in M} a_{ji} a_{li} d_i^{-1} \neq 0, \qquad \sum_{i \in M} a_{ki} a_{li} d_i^{-1} \neq 0.$$

Since $B_{12}^T D_1^{-1} B_{12} \in C_{n-m}$, there exists by 2.5 a decomposition of $N - M$ into two subsets S_1, S_2, and any two of the indices j, k, l have to belong to different subsets. Since this is impossible, the proof is complete.

LEMMA 2.7. *Let $a_1, \ldots, a_n$ $(n \geqslant 3)$ be all different from zero. Then the matrix*

$$A = \begin{vmatrix} 0 & a_1 & 0 & \ldots & 0 & a_n \\ a_1 & 0 & a_2 & \ldots & 0 & 0 \\ 0 & a_2 & 0 & \ldots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \ldots & 0 & a_{n-1} \\ a_n & 0 & 0 & \ldots & a_{n-1} & 0 \end{vmatrix}$$

does not belong to C_n.

Proof. There exist numbers $y_1, \ldots, y_n$ all different from zero such that

$$\sum_{i=1}^{n-1} (a_i y_i y_{i+1})^{-1} + (a_n y_n y_1)^{-1} = 0.$$

Hence there exist numbers $z_1, \ldots, z_n$ all different from zero such that

$$z_i y_i^{-1} - z_{i+1} y_{i+1}^{-1} = (a_i y_i y_{i+1})^{-1}, \qquad i = 1, \ldots, n-1,$$

$$z_n y_n^{-1} - z_1 y_1^{-1} = (a_n y_n y_1)^{-1}.$$

Define the numbers $d_1, \ldots, d_n$ by

$$a_1 y_2 + a_n y_n = d_1 y_1,$$
$$a_2 y_3 + a_1 y_1 = d_2 y_2,$$
$$\cdots\cdots\cdots\cdots$$
$$a_n y_1 + a_{n-1} y_{n-1} = d_n y_n.$$

An easy computation shows that then also

$$a_1 z_2 + a_n z_n = d_1 z_1,$$
$$a_2 z_3 + a_1 z_1 = d_2 z_2,$$
$$\cdots\cdots\cdots\cdots$$
$$a_n z_1 + a_{n-1} z_{n-1} = d_n z_n.$$

Hence the matrix $A - D$ where D is a diagonal matrix with diagonal elements $d_1, \ldots, d_n$ has rank smaller than $n - 1$. The proof is complete.

THEOREM 2.8. *Let $A \in C_n$. Then there exists a permutation matrix P such that PAP^T is tridiagonal irreducible.*

Proof. According to 2.6, no row of C_n contains more than two off-diagonal entries different from zero, by 2.7 and 2.3 some row contains exactly one such off-diagonal element. Choose this as the first row of the permuted matrix $B = (b_{ik})$ and as the element b_{12}. In the second row, $b_{21} = b_{12} \neq 0$. Let b_{23} be second nonzero element in this row, further $b_{34} \neq 0$ etc. This process can be performed completely since otherwise $B \in C_n$ would not be irreducible. Hence B is tridiagonal and the proof is complete.

REFERENCES

1 M. Fiedler, Metric problems in the space of matrices, *Proc. Symp. C. N. R. S.*, Programmation en mathématiques numériques, Paris, 1968, 93–103.
2 F. R. Gantmacher, *The Theory of Matrices*, Chelsea Publ. Co., New York, 1959.
3 D. M. Young, Iterative methods for solving partial difference equations of elliptic type, *TAMS* **76**(1954), 92–111.

Received March 27, 1968