Numerische Mathematik 10, 153—161 (1967)

Remarks on Band Matrices

L. ELSNER and R. M. REDHEFFER

Received August 15, 1966

Abstract. In this note we consider band- or tridiagonal-matrices of order k, whose elements above, on, and below the diagonal are denoted by b_i, a_i, c_i. In the periodic case, i.e. $b_{i+m}=b_i$ etc., we derive for $k=nm$ and $k=nm-1$ formulas for the characteristic polynomial and the eigenvectors under the assumption that $\prod_{i=1}^{m} c_i b_i > 0$. In the latter case it is shown that the characteristic polynomial is divisible by the $m-1$-th minor, as was already observed by RÓZSA. We also give estimations for the number of real roots and an application to Fibonacci numbers.

1. Introduction

Throughout this note a_i, b_i and c_i for $i=1, 2, \ldots, m$ are complex numbers with

$$b=b_1 b_2 \ldots b_m, \qquad c=c_1 c_2 \ldots c_m, \qquad \varkappa^2 = bc \neq 0.$$

The letter $B=B^m$ denotes a matrix of m rows and $m+2$ columns,

$$B=\begin{pmatrix} c_1 & a_1 & b_1 & 0 & 0 \ldots 0 & 0 & 0 & 0 \\ 0 & c_2 & a_2 & b_2 & 0 \ldots 0 & 0 & 0 & 0 \\ 0 & 0 & c_3 & a_3 & b_3 \ldots 0 & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & 0 \ldots & a_{m-1} & b_{m-1} & 0 \\ 0 & 0 & 0 & 0 & 0 \ldots & c_m & a_m & b_m \end{pmatrix}.$$

The square m by m matrix obtained from B when the first and last columns are deleted is called a band matrix and is denoted by B'. The square $m-1$ by $m-1$ matrix obtained from B' when the bottom row and right-hand column are deleted, for $m>1$, is denoted by B''.

Using a square m by m zero matrix 0_m with B^m we can form a matrix $B_n = B_n^m$ of nm rows and $nm+2$ columns,

$$B_n=\begin{pmatrix} B^m & 0_m & 0_m & \ldots & 0_m \\ 0_m & B^m & 0_m & \ldots & 0_m \\ 0_m & 0_m & B^m & \ldots & 0_m \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0_m & 0_m & 0_m & \ldots & B^m \end{pmatrix}.$$

The square mn by mn matrix obtained from B_n when the first and last columns are deleted is called a periodic band matrix and is denoted by B'_n. If the bottom row and last column of B'_n are deleted the resulting $mn-1$ by $mn-1$ matrix is also called a periodic band matrix and is denoted by B''_n. In justification of this terminology it should be observed that B_n has the same form as B, except that m has been replaced by mn and the elements are periodic. That is, if the elements of B_n are denoted by a_i, b_i and c_i with $1 \leqq i \leqq mn$ then

$$a_{i+m} = a_i, \qquad b_{i+m} = b_i, \qquad c_{i+m} = c_i.$$

As far as we know RÓZSA and LOVASS-NAGY of Budapest were the first to make a systematic study of periodic band matrices with $m>1$, and in particular [2] and [3] contain the complete solution for $m=2$. Our interest in the subject stems from a lecture which RÓZSA held at the University of Hamburg in the summer semester, 1966. This lecture was devoted to the proof of the following theorem, which generalizes the result [3] to the case of arbitrary m:

Theorem A (RÓZSA). *Let determinants $D_m = D_m(\lambda)$ be defined by*

$$D_m(\lambda) = \begin{vmatrix} a_1-\lambda & b_1 & 0 & \dots 0 & 0 \\ c_2 & a_2-\lambda & b_2 & \dots 0 & 0 \\ 0 & c_3 & a_3-\lambda & \dots 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \dots c_m & a_m-\lambda \end{vmatrix},$$

and let $D^*_m(\lambda)$ denote the determinant obtained from $D_m(\lambda)$ when the first row and column are removed. Then, if B''_n is symmetric, the characteristic values of B''_n are the values λ for which $D_{m-1}(\lambda)=0$ or for which

$$D_m(\lambda) - b_m^2 D^*_{m-1}(\lambda) = 2 b_1 b_2 \dots b_m \cos\frac{k\pi}{n}, \qquad k = 1, 2, \dots, n-1.$$

Our reasons for writing are threefold. First, our method is wholly elementary, while RÓZSA used a number of advanced tools from the algebra of matrices and determinants. Second, we obtain results for B'_n as well as B''_n, while RÓZSA was able, he said, to get results in the case of arbitrary m for B''_n only. Third, we treat the general case while RÓZSA's analysis applies only to the symmetric case

$$c_1 = b_m, \; c_2 = b_1, \; \dots, \; c_m = b_{m-1}.$$

In particular, we shall find that RÓZSA's result continues to hold in the asymmetric case, if the factor b_m^2 on the left is replaced by $b_m c_1$ and the factor $b_1 b_2 \dots b_m$ on the right is replaced by $\varkappa$.

2. The Recursive Solution

Let x' be an m-dimensional column vector with components $(x_1, x_2, \dots, x_m)$ and let x be an $(m+2)$-dimensional column vector with components

$$(x_0, x_1, \dots, x_m, x_{m+1}).$$

Evidently λ is a characteristic value for B' if and only if the system $B'x' = \lambda x'$ has a nontrivial solution x'. Since $x_1 = 0$ implies $x' = 0$ we can suppose that $x_1 = 1$.

The system $B'x' = \lambda x'$ has an unsymmetric structure because two columns were deleted in passing from B to B'. But the system $Bx = \lambda x'$ has a very symmetric structure; it is

$$c_k x_{k-1} + (a_k - \lambda) x_k + b_k x_{k+1} = 0 \qquad (k = 1, 2, \ldots, m). \tag{1}$$

If, now, we require $x_0 = x_{m+1} = 0$ the resulting system is identical with $B'x' = \lambda x'$. In other words λ is a characteristic value for B' if and only if the system (1) has a solution with $x_0 = x_{m+1} = 0$ and $x_1 = 1$. Replacing m by $m-1$ in this observation, we see that λ is a characteristic value for B'' if and only if (1) has a solution with $x_0 = x_m = 0$ and $x_1 = 1$.

If the k-th equation (1) is solved for x_{k+1} and the result is used together with $x_k = x_k$ we get a recurrence formula,

$$\begin{pmatrix} x_k \\ x_{k+1} \end{pmatrix} = \frac{1}{b_k} \begin{pmatrix} 0 & b_k \\ -c_k & \lambda - a_k \end{pmatrix} \begin{pmatrix} x_{k-1} \\ x_k \end{pmatrix}.$$

This is solved by means of the products

$$\begin{pmatrix} p_k & q_k \\ r_k & s_k \end{pmatrix} = \begin{pmatrix} 0 & b_k \\ -c_k & \lambda - a_k \end{pmatrix} \cdots \begin{pmatrix} 0 & b_2 \\ -c_2 & \lambda - a_2 \end{pmatrix} \begin{pmatrix} 0 & b_1 \\ -c_1 & \lambda - a_1 \end{pmatrix}$$

where the polynomials $p_k = p_k(\lambda)$, and so on, are defined by the equation. For brevity we denote this product by $\prod_k (\lambda)$, $1 \leq k \leq m$. Since

$$\begin{pmatrix} x_k \\ x_{k+1} \end{pmatrix} = \frac{1}{b_k b_{k-1} \cdots b_2 b_1} \begin{pmatrix} p_k & q_k \\ r_k & s_k \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

the values of x_k and the desired condition $x_m = 0$ or $x_{m+1} = 0$ can be expressed with ease.

We write the characteristic polynomial of B'_n or B''_n as

$$\lambda^{mn} + \cdots \quad \text{or} \quad \lambda^{mn-1} + \cdots$$

so that the leading coefficient is 1. The leading term of $p_m(\lambda)$, $q_m(\lambda)$, $r_m(\lambda)$ or $s_m(\lambda)$ is

$$-b_m c_1 \lambda^{m-2}, \quad b_m \lambda^{m-1}, \quad -c_1 \lambda^{m-1}, \quad \lambda^m$$

respectively, as is easily proved by induction. Since the characteristic polynomial is wholly determined by the characteristic roots, we can summarize our analysis as follows:

Remark 1. *Let polynomials p_k, q_k, r_k, s_k be defined as above. Then the characteristic polynomial of B' or B'' is $s_m(\lambda)$ or $b_m^{-1} q_m(\lambda)$, respectively. The characteristic vector for B' belonging to λ is*

$$\left(1, \frac{s_1(\lambda)}{b_1}, \frac{s_2(\lambda)}{b_1 b_2}, \ldots, \frac{s_{m-1}(\lambda)}{b_1 b_2 \cdots b_{m-1}}\right),$$

where $s_m(\lambda) = 0$, and that for B'' is the same, with $m-1$ in place of m, and with $q_m(\lambda) = 0$ instead of $s_m(\lambda) = 0$.

We define now D_m, D_m^* as in Theorem A and analogously D_{m-1}, D_{m-1}^* with $D_0 = 1$, $D_0^* = 0$, $D_1^* = 1$ by convention. Then we obtain easily

Remark 2. *The following relations hold for $m \geqq 1$:*

$$p_m(\lambda) = (-)^{m-1} b_m c_1 D^*_{m-1}(\lambda), \qquad q_m(\lambda) = (-)^{m-1} b_m D_{m-1}(\lambda),$$
$$r_m(\lambda) = (-)^m c_1 D^*_m(\lambda), \qquad s_m(\lambda) = (-)^m D_m(\lambda).$$

For proof, form B^* by deleting the first row and column of B. Apply Remark 1 to the corresponding product $\prod_m^*$ as well as to $\prod_m$ and note that $\prod_m^*$ and $\prod_m$ are simply related.

3. The Roots in the Periodic Case

We introduce

$$A = A(\lambda) = \varkappa^{-1} \prod_m (\lambda).$$

Since $\det \prod_m (\lambda) = \varkappa^2$, we have $\det A = 1$. The trace of A is $2t$, where $t = t(\lambda)$ is defined by

$$2\varkappa\, t(\lambda) = p_m(\lambda) + s_m(\lambda).$$

We have by the Cayley-Hamilton theorem

$$A^2 - 2tA + I = 0 \qquad I = \text{identity-matrix} \tag{2}$$

and hence with $M_n = A^n$

$$M_{n+1} + M_{n-1} = 2t M_n. \tag{3}$$

Setting $t = \cos\theta$, where θ is a real or complex angle, we see that the solution of the difference equation (3) is given by

$$M_n = M_1 \frac{\sin n\theta}{\sin\theta} - M_0 \frac{\sin(n-1)\theta}{\sin\theta}.$$

(Here and elsewhere, $\sin k\theta/\sin\theta$ is to be replaced by its limit, $k\cos k\theta/\cos\theta$, if $\sin\theta = 0$.) It follows, in particular, that

$$A^n = A \frac{\sin n\theta}{\sin\theta} - I \frac{\sin(n-1)\theta}{\sin\theta}. \tag{4}$$

Upon applying Remark 1 to B_n instead of B we obtain

Remark 3. *The characteristic roots of B'_n are the values λ for which simultaneously*

$$2\varkappa\cos\theta = p_m(\lambda) + s_m(\lambda) \quad \text{and} \quad s_m(\lambda)\frac{\sin n\theta}{\sin\theta} = \varkappa \frac{\sin(n-1)\theta}{\sin\theta}.$$

The characteristic roots of B''_n are the values λ for which

$$q_m(\lambda) = 0 \quad \text{or} \quad p_m(\lambda) + s_m(\lambda) = 2\varkappa\cos\frac{k\pi}{n}, \qquad k = 1, 2, \ldots, n-1.$$

Remark 2 gives a corresponding version of Remark 3 with p_m, q_m and s_m replaced by appropriate subdeterminants of D_m. The part of Remark 3 pertaining to B''_n thus gives the generalized version of RÓZSA's theorem mentioned in the introduction. Conversely, our result for B''_n can be deduced from RÓZSA's by an affine transformation.

4. Characteristic Vectors

To discuss the characteristic vectors x' for B'_n and x'' for B''_n it is convenient to denote the elements of B_n by a_i, b_i and c_i with $1 \leqq i \leqq mn$ and with the periodicity condition. The products $\prod_i = \prod_i(\lambda)$ for $i > m$ are interpreted accordingly. Any index $i \geqq 1$ can be written $i = mj + l$ with $1 \leqq l \leqq m$. Then

$$b_1 b_2 \dots b_i = b^j b_1 b_2 \dots b_l \quad \text{and} \quad \prod_i = \prod_l \left(\prod_m\right)^j.$$

Since $\prod_m = \varkappa A$ the formula (4) for A^n gives a similar formula for $\left(\prod_m\right)^j$. Upon applying Remark 1 to B_n instead of B we get:

Remark 4. *Let $\theta = \theta(\lambda)$ be defined by*

$$2\varkappa \cos\theta = p_m(\lambda) + s_m(\lambda)$$

where λ is a characteristic value for B'_n or B''_n, as the case may be. Let the corresponding characteristic vector be (x_i) with $1 \leqq i \leqq mn$ or $1 \leqq i \leqq mn - 1$, respectively. Then for $j = 0, 1, 2, \dots$ and $1 \leqq l \leqq m$ we have

$$x_{mj+l} = \varrho^{j-1} \frac{\sin j\theta}{\sin\theta} x_{m+l} - \varrho^j \frac{\sin(j-1)\theta}{\sin\theta} x_l$$

where $\varrho = \sqrt{c/b}$.

The interest of the result is that it gives x_i for all i as soon as λ and the initial values

$$x_1, x_2, \dots, x_{2m}$$

are known. These initial values can be found by the formula of Remark 1 or by recursive solution of (1). Of course the recursive solution gives x_i for all i, but if i is large, the formula of Remark 4 is simpler and more accurate.

5. The Characteristic Polynomial

The polynomial of degree n defined by

$$U_n(\cos\theta) = \frac{\sin(n+1)\theta}{\sin\theta} \tag{5}$$

is called the Chebychev polynomial of the second kind [*1*]. In terms of U_n the result (4) reads

$$A^n = A\, U_{n-1}(t) - I\, U_{n-2}(t). \tag{6}$$

Applying Remark 1 to B'_n and B''_n, using (6) and the relation

$$U_{n+1}(t) + U_{n-1}(t) = 2t U_n(t) \tag{7}$$

we obtain

Remark 5. *Let $t = t(\lambda)$ be defined by $2\varkappa t = p_m(\lambda) + s_m(\lambda)$. Then the characteristic polynomial of B'_n is*

$$\varkappa^n U_n(t) - p_m(\lambda)\, \varkappa^{n-1} U_{n-1}(t)$$

and the characteristic polynomial of B_n'' *is*

$$\varkappa^{n-1} s_{m-1}(\lambda)\, U_{n-1}(t)$$

under the convention that $s_0(\lambda)=1$.

Remark 5 gives identities similar to those of Remark 2 for B_n' and B_n''. In particular, setting $\lambda=0$ we get the constant term of the characteristic equation,

$$(-)^{mn}\det B_n' \quad \text{or} \quad (-)^{mn-1}\det B_n''.$$

Since the choice $\theta=\pi/2$ in (5) gives $U_n(0)$ we are led to the following as a special case: Suppose

$$p_m(0)+s_m(0)=0.$$

Then

$$(-)^{mn}\det B_n'=\varkappa^n, \quad -p_m(0)\,\varkappa^{n-1}, \quad -\varkappa^n, \quad p_m(0)\,\varkappa^{n-1}$$

according as $n\equiv 0, 1, 2$ or $3 \mod (4)$, respectively. In the same circumstances the respective values for $\det B_n''$ are given by

$$(-)^{mn}\det B_n''=0, \quad -s_{m-1}(0)\,\varkappa^{n-1}, \quad 0, \quad s_{m-1}(0)\,\varkappa^{n-1}.$$

6. Conditions for Real Roots

In this section we give conditions under which B_n' and B_n'' have at least mn-const distinct real roots as $n\to\infty$. It is convenient to assume that $\varkappa$ is real and positive, and of course, that a_i, b_i and c_i are real.

The desired results can be read off from a plot of $y=t(\lambda)$ together with the horizontal lines

$$y=\cos\frac{k\pi}{n}, \qquad k=1,2,\ldots,n-1. \tag{8}$$

Such a plot might have the appearance suggested by the figure when $m=5$ and $n=10$. Each intersection point gives a real root of B_n''.

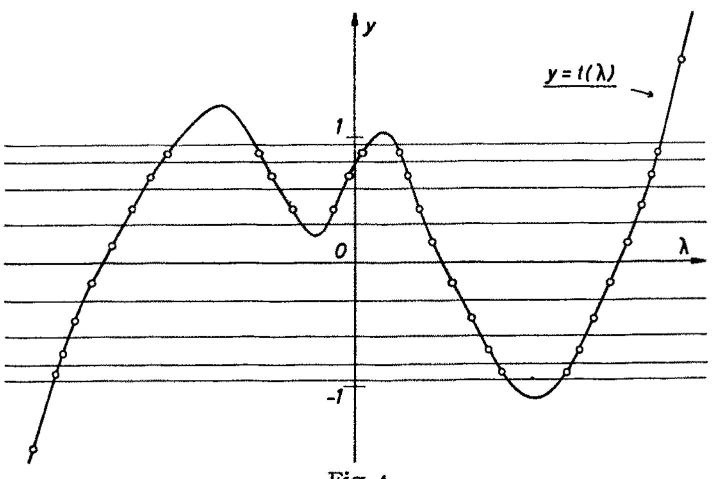

Fig. 1

To discuss the corresponding graphical interpretation for B'_n we note that a plot of

$$y = \frac{U_n(t)}{U_{n-1}(t)}$$

vs. t consists of n continuous curves, in each of which y ranges from $-\infty$ to ∞. The equation

$$p_m(\lambda) = \varkappa \frac{U_n(t)}{U_{n-1}(t)}$$

given by Remark 5 therefore has at least one real root in each interval on which the graph of $t(\lambda)$ joins one line (8) to an adjacent one, and also in each of the two unbounded intervals. These portions of the graph are indicated by the round dots in the figure.

We now introduce the following definition:

Definition. *A real polynomial $P(x)$ of degree n is of Chebychev type if there are $n+1$ values x_i such that*

$$x_0 > x_1 > x_2 \ldots > x_n \quad \textit{and} \quad (-)^k P(x_k) \geqq 1 \quad (k = 0, 1, \ldots, n).$$

Evidently, a Chebychev polynomial is an extremal polynomial of this class, in several senses of the word "extremal".

By the Lagrange interpolation formula we see that a polynomial is of Chebychev type if and only if it admits the representation

$$P(x) = \sum_{k=0}^{n} \frac{Q(x)}{x - x_k} \frac{(-)^k A_k}{Q'(x_k)}$$

where

$$Q(x) = (x - x_0)(x - x_1) \ldots (x - x_n), \qquad A_k \geqq 1,$$

and x_k are as in the definition. If $P(x)$ admits such a representation then it admits one in which the values $x_1, x_2, \ldots, x_{n-1}$ are the roots of $P'(x) = 0$, and in which furthermore

$$P(x_0) = (-)^n P(x_n) = 1.$$

The latter representation is unique. Other characterizations can be obtained by writing

$$P'(x) = (\text{const})(x - x_1)(x - x_2) \ldots (x - x_{n-1})$$

and integrating to get $P(x)$.

The graphical interpretation of the foregoing discussion leads to the following:

Remark 6. *Let a_i, b_i, c_i and $\varkappa = \sqrt{bc}$ be real, with $\varkappa > 0$. Let a polynomial $t(\lambda)$ be defined by*

$$2\varkappa\, t(\lambda) = p_m(\lambda) + s_m(\lambda).$$

Then if $t(\lambda)$ is a polynomial of Chebychev type the matrices B'_n and B''_n have at least

$$mn - 2m + 2 \quad \textit{and} \quad mn - m$$

distinct real characteristic values, respectively. But if $t(\lambda)$ is not a polynomial of Chebychev type there is a constant $m_1 < m$ such that the number of real characteristic roots of B'_n or B''_n, counting multiplicity, does not exceed $m_1 n$ for sufficiently large n.

It should be observed that the derivation of Remark 6 not only gives somewhat more information about the number of real roots than is there stated, but also gives quite detailed information about their location.

So far we have assumed $bc>0$. If $bc<0$ Remark 3 shows that the only real roots of B_n'' are among the roots of $q_m(\lambda)$ or $p_m(\lambda)+s_m(\lambda)$ and hence, there are at most $2m-1$ real roots. The corresponding question for B_n' is left as an open problem.

An evident consequence of Remark 6 is:

Remark 7. *If the products $c_1 b_m, c_2 b_1, c_3 b_2, \ldots, c_m b_{m-1}$ are all positive, then $t(\lambda)$ is a polynomial of Chebychev type.*

7. An Application to Number Theory

The Fibonacci numbers 0, 1, 1, 2, 3, 5, ... are defined by

$$f_0=0, \qquad f_1=1, \qquad f_{n+1}=f_n+f_{n-1}, \qquad n \geqq 1.$$

It was pointed out by COLLATZ that Remark 5 gives an extremely simple proof of the following well-known theorem:

Remark 8 (e.g. [4], p. 148). *If k is divisible by m, then f_k is divisible by f_m.*

COLLATZ'S proof, presented here by permission, is as follows. By induction, we obtain the well-known formula

$$f_{j+1}=D_j \quad \text{with} \quad D_j=\left.\begin{vmatrix} 1 & -1 & & & 0 \\ 1 & 1 & \ddots & & \\ & 1 & \ddots & \ddots & \\ & & \ddots & \ddots & -1 \\ 0 & & & 1 & 1 \end{vmatrix}\right\} j.$$

The assumption that m divides k gives

$$k-1=(n-1)m+m-1$$

for an integer n.

Now, if we regard D_{k-1} as the determinant of a periodic band-matrix with period m, Remark 5 with $\lambda=0$ gives

$$D_{k-1}=K^{n-1} s_{m-1}(0)\, U_{n-1}(t)$$

where

$$K=(i)^{m-1} \quad \text{and} \quad 2Kt=p_m(0)+q_m(0).$$

It is evident that t is a Gaussian integer (that is, $t=\alpha+\beta i$ with α and β integers) and the same is true of $U_{n-1}(t)$. Since Remark 2 gives $s_{m-1}(0)=(-)^{m-1} D_{m-1}$, we conclude that

$$D_{k-1}=D_{m-1} \cdot (\text{Gaussian integer}).$$

The Gaussian integer in the equation is necessarily real, as the D_j's are real, and thus, the proof is complete.

Remark 8 can be generalized to sequences of the following types:

$$f_0=0, \quad f_1=1, \quad f_{n+1}=a\,f_n+b^2 f_{n-1}$$

respectively

$$f_0=0, \quad f_1=1, \quad f_{n+1}=a\,f_n-b^2 f_{n-1}$$

with integers a, b. Observe that $K^{n-1}\,U_{n-1}(r/K)$ is an integer for r and K integral.

We get analogous results for sequences of polynomials which are recursively defined by

$$p_0=1, \quad p_1(\lambda)=a-\lambda, \quad p_{i+1}(\lambda)=(a-\lambda)\,p_i(\lambda)+b\,p_{i-1}(\lambda)$$

for arbitrary a, b. Here we have: If $i|m$, then $p_{i-1}(\lambda)\,|\,p_{m-1}(\lambda)$ $\big($e.g. $p_n(\lambda)=U_n(\lambda)\big)$.

References

[1] Abramowitz, M., and I. Stegun: Handbook of mathematical functions, p. 776. New York: Dover Pub. Inc. 1965.

[2] Lovass-Nagy, V., u. P. Rózsa: Die Berechnung von Ausgleichvorgängen auf längskompensierten Fernleitungen. Arch. Elektrotech. **49**, 260–270 (1964).

[3] — — Matrix analysis of transient voltage distributions in alternating ladder networks. Proc. I.E.E., vol. 110, No. 9, September 1963.

[4] Hardy, G. H., and E. M. Wright: An introduction to the theory of numbers. London: Oxford University Press 1960.

Department of Mathematics
University of California
Los Angeles, California 90024, USA

Institut für Angewandte Mathematik
der Universität Hamburg
2000 Hamburg 13
Rothenbaumchaussee 67/69