Linear and Multilinear Algebra, 1987, Vol. 21, pp. 289–296
Photocopying permitted by license only

Characteristic Polynomials of Symmetric Rationally Generated Toeplitz Matrices

WILLIAM F. TRENCH
Department of Mathematics, Trinity University, San Antonio, TX 78284

(Received January 24, 1986; in final form December 20, 1986)

Formulas are given for the characteristic polynomials $\{p_n(\lambda)\}$ and eigenvectors of the family $\{T_n\}$ of real symmetric Toeplitz matrices generated by a rational function $R(z)$ with real coefficients such that $R(z) = R(1/z)$. The formulas are in terms of the zeros of a fixed polynomial $P(w; \lambda)$ with coefficients which are simple functions of λ and the coefficients of $R(z)$. The representation for $p_n(\lambda)$ exhibits two factors such that the zeros of one have associated symmetric eigenvectors and the zeros of the other have associated skew-symmetric eigenvectors. In all of these formulas, n is a parameter; that is, the formula does not become more complicated as n increases.

1. INTRODUCTION

Let

$$A(z) = \sum_{j=0}^{q} a_j z^j$$

and

$$C(z) = \sum_{j=-p}^{p} c_j z^j,$$

where $a_0, a_1, \ldots, a_q$ and $c_{-p}, \ldots, c_p$ are real, $c_{-j} = c_j (1 \leqslant j \leqslant p)$, $(p + q)a_0 a_q c_p \neq 0$, and no two of the polynomials $A(z)$, $z^q A(1/z)$, and $z^p C(z)$ have a zero in common. We consider the real symmetric Toeplitz matrices

$$T_n = (t_{j-i})_{i,j=1}^{n} \qquad (n = 1, 2, \ldots), \tag{1}$$

associated with the coefficients $\{t_j\}$ in the formal Laurent series

$$\frac{C(z)}{A(z)A(1/z)} \sim \sum_{j=-\infty}^{\infty} t_j z^j, \tag{2}$$

defined as follows. If $q > 0$, write

$$\frac{1}{A(z)} = \sum_{j=0}^{\infty} \alpha_j z^j, \qquad |z| < R, \tag{3}$$

where

$$R = \sup\{\rho \mid A(z) \neq 0 \text{ if } |z| < \rho\},$$

and define the formal Laurent series on the right of (2) by

$$\sum_{j=-\infty}^{\infty} t_j z^j = C(z)\left[z^q f(z)\left(\sum_{j=0}^{\infty} \alpha_j z^j\right) + g(z) \sum_{j=0}^{\infty} \alpha_j z^{-j}\right], \tag{4}$$

where f and g are the unique polynomials of degree $<q$ such that

$$z^q f(z)A(1/z) + g(z)A(z) = 1. \tag{5}$$

(Recall that $A(z)$ and $z^q A(1/z)$ are relatively prime.) This series formally represents the rational function on the left of (2) in the sense that formal multiplication yields

$$A(z)A(1/z) \sum_{j=-\infty}^{\infty} t_j z^j = C(z),$$

because of (3), (4), and (5). If $q = 0$, then $t_j = c_j$ if $-p \leqslant j \leqslant p$ and $t_j = 0$ if $|j| > p$; thus, (1) is banded if $n > 2p$.

If $R > 1$, then we can replace "$\sim$" in (2) with "$=$" for all z in the annulus $1/R < |z| < R$. The covariance matrices of real-valued autoregressive moving average time series are Toeplitz matrices generated in this way by rational functions of the form (2), where A has no zeros in $|z| \leqslant 1$, $C(z) = B(z)B(1/z)$, and

$$B(z) = \sum_{j=0}^{p} b_j z^j,$$

with $b_0, \ldots, b_p$ real.

In [2] we obtained formulas for the characteristic polynomials of Toeplitz matrices for which the $\{t_j\}$ in (2) are the coefficients of a formal Laurent series of an arbitrary rational function, so that T_n need not be symmetric. Here we start with the results of [2] specialized to the symmetric case, and deduce from them new formulas which give additional insight into the symmetric case. Numerical experiments (discussed in [3]) seem to indicate that these formulas may provide an efficient method for computing the eigenvalues of these matrices.

We let

$$\max(p, q) = m \geqslant 1, \tag{6}$$

and define $\theta_{-p}, \ldots, \theta_p$ by

$$[A(z)A(1/z)] = \sum_{j=-q}^{q} \theta_j z^j. \tag{7}$$

It is convenient to define $c_j = 0$ if $|j| > p$ and $\theta_j = 0$ if $|j| > q$.

We start from the following lemma, which can be obtained by applying Theorem 2 and Corollary 1 of [2] under our present assumptions.

LEMMA *Suppose that λ is such that*

$$c_m - \lambda\theta_m \neq 0, \tag{8}$$

and the Laurent polynomial

$$Q(z;\lambda) = c_0 - \lambda\theta_0 + \sum_{j=1}^{m} (c_j - \lambda\theta_j)(z^j + z^{-j}), \tag{9}$$

has 2m distinct zeros

$$z_1, \ldots, z_m, \frac{1}{z_1}, \ldots, \frac{1}{z_m}. \tag{10}$$

For $n \geqslant 1$, let D_n be the $2m \times 2m$ determinant given in block form by

$$D_n = \begin{vmatrix} [z_s^{r-1}A(z_s)] & [z_s^{-r+1}A(1/z_s)] \\ [z_s^{n+m+r-1}A(1/z_s)] & [z_s^{-n-m-r+1}A(z_s)] \end{vmatrix}, \tag{11}$$

where $1 \leqslant r, s \leqslant m$ in each of the four $m \times m$ matrices on the right. Let V be the $2m \times 2m$ determinant obtained by letting $n = 0$ and $A(z) = 1$ in (11); i.e.,

$$V = \begin{vmatrix} [z_s^{r-1}] & [z_s^{-r+1}] \\ [z_s^{m+r-1}] & [z_s^{-m-r+1}] \end{vmatrix},$$

is the Vandermonde determinant of the zeros (10) of $Q(\ ;\lambda)$. Then the value of the characteristic polynomial

$$p_n(\lambda) = \det(\lambda I_n - T_n), \tag{12}$$

is given by

$$p_n(\lambda) = K_n(c_m - \lambda\theta_m)^n D_n/V, \tag{13}$$

where K_n is a constant. Moreover, if λ is an eigenvalue of T_n and $\{G_1, \dots, G_m, H_1, \dots, H_m\}$ is a nontrivial solution of the $2m \times 2m$ system

$$\text{(a)} \quad \sum_{s=1}^{m} [z_s^{r-1} A(z_s) G_s + z_s^{-r+1} A(1/z_s) H_s] = 0, \qquad 1 \leqslant r \leqslant m,$$
$$\text{(b)} \sum_{s=1}^{m} [z_s^{n+m+r-1} A(1/z_s) G_s + z_s^{-n-m-r+1} A(z_s) H_s] = 0, \qquad 1 \leqslant r \leqslant m, \tag{14}$$

(which has the determinant D_n), then the vector

$$U = [u_1, \dots, u_n]^t, \tag{15}$$

with components

$$u_r = \sum_{s=1}^{m} A(z_s) A(1/z_s) [G_s z_s^{m+r-1} + H_s z_s^{-m-r+1}], \qquad 1 \leqslant r \leqslant m, \tag{16}$$

is a λ-eigenvector of T_n.

The value of K_n is given explicitly in [2], but it is not important here.

2. THE MAIN THEOREM

Following Cantoni and Butler [1], we say that an n-vector (15) is *symmetric* if $u_r = u_{n-r+1}$ ($1 \leqslant r \leqslant n$), or *skew-symmetric* if $u_r = -u_{n-r+1}$ ($1 \leqslant r \leqslant n$). Cantoni and Butler have shown that if T_n is a real symmetric Toeplitz matrix of order n, then R^n has an orthonormal basis consisting of $[n/2]$ skew symmetric and $n - [n/2]$ symmetric eigenvectors of T_n. (Here $[x]$ is the integer part of x.)

The following is our main result.

THEOREM 1 *With $\theta_{-p}, \dots, \theta_p$ as in (7), let*

$$P(w; \lambda) = c_0 - \lambda\theta_0 + 2 \sum_{j=1}^{m} (c_j - \lambda\theta_j) \tau_j(w), \tag{17}$$

where $\tau_1, \dots, \tau_m$ are the Chebychev polynomials; i.e., $\tau_n(\cos t) = \cos nt$. Let λ be such that (8) holds and $P(\ ;\lambda)$ has m distinct zeros $w_1, \dots, w_m$ such that

$$w_j \neq 1, -1, \qquad 1 \leqslant j \leqslant m. \tag{18}$$

Now let

$$\gamma_j = \tfrac{1}{2}\cos^{-1} w_j, \qquad 0 \leqslant \mathrm{Re}(\gamma_j) \leqslant \frac{\pi}{2}, \tag{19}$$

and define

$$C_{rn}(\gamma_j) = \sum_{j=0}^{q} a_j \cos(n + 2r - 2j - 1)\gamma_j, \tag{20}$$

and

$$S_{rn}(\gamma_j) = \sum_{j=0}^{q} a_j \sin(n + 2r - 2j - 1)\gamma_j. \tag{21}$$

Then the value of the characteristic polynomial (12) *is given by*

$$p_n(\lambda) = K_n(c_m - \lambda\theta_m)^n F_{0n}(\lambda) F_{1n}(\lambda), \tag{22}$$

where

$$F_{0n}(\lambda) = \frac{\det[C_{rn}(\gamma_s)]_{r,s=1}^{m}}{\det[\cos(2r-1)\gamma_s]_{r,s=1}^{m}}, \tag{23}$$

and

$$F_{1n}(\lambda) = \frac{\det[S_{rn}(\gamma_s)]_{r,s=1}^{m}}{\det[\sin(2r-1)\gamma_s]_{r,s=1}^{m}}. \tag{24}$$

Moreover, if $F_{ln}(\lambda) = 0$, *then* T_n *has a* λ-*eigenvector which is symmetric if* $l = 0$, *or skew-symmetric if* $l = 1$.

Proof Define

$$z_r = w_r + \sqrt{w_r^2 - 1}, \qquad 1 \leqslant r \leqslant m, \tag{25}$$

so that

$$w_r = \tfrac{1}{2}(z_r + 1/z_r).$$

From the defining property of the Chebychev polynomials,

$$\tau_j(w_r) = \tfrac{1}{2}(z_r^j + z_r^{-j});$$

therefore, z_r and $1/z_r$ are zeros of $Q(z;\lambda)$ (cf. (9) and (17)). Moreover, since $w_1, \ldots, w_m$ are distinct and satisfy (18), the quantities (10) are distinct, and the hypotheses of Lemma 1 are verified.

We now perform manipulations on D_n, as defined in (11). For $1 \leqslant s \leqslant m$, divide column s and multiply column $m + s$ by $z_s^{(n+2m-1)/2}$ to obtain

$$D_n = \begin{vmatrix} [z_s^{-(n+2m-2r+1)/2} A(z_s)] & [z_s^{(n+2m-2r+1)/2} A(1/z_s)] \\ [z_s^{(n+2r-1)/2} A(1/z_s)] & [z_s^{-(n+2r-1)/2} A(z_s)] \end{vmatrix}.$$

Now it is convenient to reverse the order of the first m rows to obtain

$$D_n = \pm \begin{vmatrix} [z_s^{-(n+2r-1)/2} A(z_s)] & [z_s^{(n+2r-1)/2} A(1/z_s)] \\ [z_s^{(n+2r-1)/2} A(1/z_s)] & [z_s^{-(n+2r-1)/2} A(z_s)] \end{vmatrix}. \tag{26}$$

(For typographical reasons we are not specific about the "$\pm$", which will turn out to be irrelevant to our final result.)

Subtracting row r from row $m + r$ $(1 \leq r \leq m)$ in (26) yields

$$D_n = \pm \begin{vmatrix} [z_s^{-(n+2r-1)/2} A(z_s)] & [z_s^{(n+2r-1)/2} A(1/z_s)] \\ E_n & -E_n \end{vmatrix}, \tag{27}$$

with

$$E_n = [z_s^{(n+2r-1)/2} A(1/z_s) - z_s^{-(n+2r-1)/2} A(z_s)]_{r,s=1}^{m}. \tag{28}$$

Now adding column s to column $m + s$ $(1 \leq s \leq n)$ in (27) yields

$$D_n = \pm \begin{vmatrix} [z_s^{-(n+2r-1)/2} A(z_s)] & F_n \\ E_n & O_n \end{vmatrix}, \tag{29}$$

where O_n is the $n \times n$ zero matrix and

$$F_n = [z_s^{(n+2r-1)/2} A(1/z_s) + z_s^{-(n+2r-1)/2} A(z_s)]_{r,s=1}^{m}. \tag{30}$$

Now Laplace's expansion of (29) yields

$$D_n = \pm \det(E_n) \det(F_n). \tag{31}$$

By taking $A(z) = 1$ and $n = 0$, we infer from this result that

$$V = \pm \det(E) \det(F), \tag{32}$$

where

$$E = [z_s^{(2r-1)/2} - z_s^{-(2r-1)/2}]_{r,s=1}^{m}, \tag{33}$$

and

$$F = [z_s^{(2r-1)/2} + z_s^{-(2r-1)/2}]_{r,s=1}^{m}, \tag{34}$$

and the $\pm$ in (32) is the same as in (31). Since V is the Vandermonde determinant of the distinct points (10), (32) implies that $\det(E) \neq 0$ and $\det(F) \neq 0$.

Now (13), (31), and (32) imply (22), with

$$F_{0n}(\lambda) = \frac{\det(F_n)}{\det(F)}, \tag{35}$$

and

$$F_{1n}(\lambda) = \frac{\det(E_n)}{\det(E)}. \tag{36}$$

With γ_j as in (19), $w_j = \cos 2\gamma_j$, and (25) implies that $z_j = e^{2i\gamma_j}$. Substituting this into (28), (30), (33), and (34) shows that (35) and (36) can be rewritten as (23) and (24), respectively. (Recall (20) and (21).)

If $F_{ln}(\lambda) = 0$ ($l = 0$ or 1), then the system

$$\sum_{s=1}^{m} \left[A(z_s)z_s^{-(n+2r-1)/2} + (-1)^l A(1/z_s)z_s^{(n+2r-1)/2}\right]P_s = 0, \qquad 1 \leq r \leq m,$$

has a nontrivial solution $P_1, \dots, P_m$. From this it is straightforward to verify that the system (14) has the nontrivial solution

$$H_s = z_s^{(n+2m-1)/2}P_s, \quad G_s = (-1)^l z_s^{-(n+2m-1)/2}P_s, \quad 1 \leq s \leq m. \tag{37}$$

(To see this, it is convenient to replace r by $m - r + 1$ in (14a).) Substituting (37) into (16) yields the eigenvector (15), with

$$u_r = \sum_{s=1}^{m} A(z_s)A(1/z_s)P_s\left[z_s^{(n-2r+1)/2} + (-1)^l z_s^{-(n-2r+1)/2}\right]. \tag{38}$$

It is easy to verify that $u_{n-r+1} = (-1)^l u_r$, which completes the proof of Theorem 1.

3. A REMARK

It is generally agreed that attempting to compute the eigenvalues of a high order matrix by any method involving the application of root finding techniques to its characteristic polynomial is wildly impractical. However, Theorem 1 essentially reduces the solution of the eigenvalue problem for T_n to finding those values of λ for which either

$$\det[C_{rn}(\gamma_s)]_{r,s=1}^{m} = 0 \qquad \text{or} \qquad \det[S_{rn}(\gamma_s)]_{r,s=1}^{m} = 0. \tag{39}$$

The computations to determine $\gamma_1, \dots, \gamma_m$ for a given λ are independent of n, and n enters into these determinants only as a parameter. Moreover, elementary manipulations make it possible to delete complex terms and factor out exponential factors which occur in the determinants in (39) if some of the quantities $\gamma_1, \dots, \lambda_m$ are not real-valued. Therefore, the determinants in (39) can be replaced by "well-scaled" functions which have the same zeros as $p_n(\lambda)$, but do not vary wildly, as $p_n(\lambda)$ does if n is large; in fact, they are bounded for all n. Numerical experiments already performed (and reported in [3]) show that this approach can be used to obtain *all* eigenvalues of high order rationally generated symmetric Toeplitz matrices such that $m = 1, 2,$ or 3 in (6), and at a cost per eigenvalue which is essentially independent of the order n. There seems to be no theoretical reason to preclude the extension of the method so as to deal with larger values of m.

References

[1] A. Cantoni and P. Butler, Eigenvalues and eigenvectors of symmetric centrosymmetric matrices, *Lin. Alg. Appl.* **13** (1976), 275–288.
[2] W. F. Trench, On the eigenvalue problem for Toeplitz matrices generated by rational functions, *Lin. Multilin. Alg.* **17** (1985), 337–353; Addendum (to appear).
[3] W. F. Trench, Numerical solution of the eigenvalue problem for symmetric rationally generated Toeplitz matrices, in *Linear Algebra in Signals, Systems, and Control* (to appear).