A GENERALIZATION OF THE FROBENIUS COMPANION MATRIX

By

LAJOS LÁSZLÓ

National Planning Office Budapest, V., Roosevelt square 7–8.

Introduction

Present paper generalizes the results of [2] concerning a set of matrices having a prescribed characteristic polynomial, and presents another special case giving Frobenius' companion matrix.

Theorem. Let $P(z) = \sum\limits_{i=0}^{n} c_i\, z^i$ be a monic complex polynomial. For an arbitrary set $\{A_i\}_{i=1}^{n-1}$ of complex numbers, there exists an $n \times n$ matrix B containing these numbers and having P as its characteristic polynomial.

In detail, the following is true. Let us define the first $n-1$ rows of B via (a)–(c):

(a) For $1 \le i \le n-1$ let

$$b_{ii} = A_i.$$

b) Let us fix $1 \le i \le n-1$, then there are two cases:

(b_1) If $A_j \neq A_i$ for every $i < j < n$, then let

$$b_{in} = 1.$$

(b_2) If $A_j = A_i$ for some $i < j < n$ and $A_k \neq A_i$ if $i < k < j$, then let

$$b_{ij} = 1.$$

(c) For all the remaining elements of the first $n-1$ rows let

$$b_{ij} = 0.$$

Then there exist complex numbers $\{b_{nk}\}_{k=1}^{n}$ in a unique way with the equality

(1) $$det(z\,I - B) = P(z)$$

holding for the matrix $B = (b_{ij})_{i,\,j=1}^{n}$. $\boxed{\times}$

Proof. Comparing the coefficients of z^{n-1} in (1), we have the requirement $c_{n-1} = -\sum\limits_{i=1}^{n} b_{ii}$. Using this equality, b_{nn} can be determined:

$$(2) \qquad\qquad b_{nn} = -c_{n-1} - \sum_{i=1}^{n-1} b_{ii} \,.$$

For uniformity — see (a) — let $A_n = b_{nn}$. Expanding the determinant in (1) by the last row we get

$$(3) \qquad det(z\,I - B) = \prod_{i=1}^{n} (z - A_i) - \sum_{k=1}^{n-1} b_{nk} P_k(z) \,.$$

If (2) is valid, then the degree of the polynomial

$$(4) \qquad\qquad \Delta(z) = \prod_{i=1}^{n} (z - A_i) - det(zI - B)$$

is not greater than $n-2$. (1) is equivalent to

$$\sum_{k=1}^{n-1} b_{nk} P_k(z) = \Delta(z) = \prod_{i=1}^{n} (z - A_i) - P(z)$$

hence it is evident that there exist numbers $\{b_{nk}\}_{k=1}^{n-1}$ with the desired property if and only if the polynomials P_k, $1 \le k \le n-1$ constitute a basis in H_{n-2}. (H_{n-2} stands for the space of the polynomials with a degree less than or equal to $n-2$.)

Now we introduce the following notations. Let $\{a_i\}_{i=1}^{m}$ be the set of the distinct numbers in $\{A_i\}_{i=1}^{n-1}$ and suppose that a_i occurs ν_i times among the A_i's. Obviously: $m \le n-1$ and $\sum\limits_{i=1}^{m} \nu_i = n-1$. Let

$$h_i(z) = \prod_{\substack{j=1 \\ j \ne i}}^{m} (z - a_j)^{\nu_j}, \quad 1 \le i \le m\,.$$

As a consequence of the next lemma, the set P_k, $1 \le k \le n-1$ turns out to be identical with the set of the polynomials

$$(5) \qquad h_{ij}^{*}(z) = h_i(z)\,(z - a_i)^j, \quad 0 \le j \le \nu_i - 1, \quad 1 \le i \le m\,.$$

These are similar to Hermite's interpolatory polynomials

$$(6) \qquad h_{ij}(z) = h_i(z)\,(z - a_i)^j\,Q_{ij}(z), \quad 0 \le j \le \nu_i - 1, \quad 1 \le i \le m\,,$$

belonging to the same nodes ($Q_{ij} \in H_{\nu_i - j - 1}$).

Since every h_{ij} can be expressed in the form

$$(7) \qquad\qquad h_{ij} = \sum_{s=j}^{\nu_i - 1} \gamma_{ijs}\, h_{ls}^{*},$$

furthermore, the number of the polynomials (5) is $n-1$, and finally, since the polynomials (6) form a basis in H_{n-2}, the same is true for the h_{ij}^*-s, i.e., for the polynomials P_k, $1 \leqq k \leqq n-1$. The proof is complete. $\boxtimes$

In the proof we used the following

Lemma. Let us use the notations of the Theorem, let $1 \leqq k \leqq n-1$, and $A_k = a_i$. Let us further suppose that the number of A_l's, $l = 1, 2, \ldots,$ $k-1$ being equal to a_i is j. Then

$$P_k = h_{ij}^*.$$

Proof. Denote the submatrix arising from $zI-B$ by deleting the last row and the k-th column by $B_k(z)$. Then, by definition

(8) $$P_k(z) = (-1)^{n+k} \det B_k(z).$$

Let $B_k'(z)$ be the left upper $(k-1) \times (k-1)$ submatrix and $B_k''(z)$ the right lower $(n-k) \times (n-k)$ submatrix of $B_k(z)$, then the partitioning

$$B_k(z) = \begin{pmatrix} B_k'(z) & V \\ 0 & B_k''(z) \end{pmatrix}$$

shows that

(9) $$\det B_k(z) = \det B_k'(z) \det B_k''(z).$$

Now $\det B_k''(z)$ will be decomposed further, into $v_i - j$ factors. In $B_k''(z)$ there are exactly $v_i - j$ -1's with $z-a_i$ in the same row or column. For each of these -1's, let us denote the square matrix being symmetric with respect to the main diagonal and having the -1 at its upper right corner, by $B_{ks}''(z)$, $s = j+1, \ldots, v_i$. In this way we get:

Above the blocks of the $B_{ks}''(z)$-s there are zeros or, possibly, -1's, while below them — except for the $z-a_i$'s located at the junctions of these blocks — all the rest is zero. It is easy to see that, at calculating $\det B_k''(z)$, each term containing $z-a_i$ vanishes. Thus

(10) $$\det B_k''(z) = \prod_{s=j+1}^{v_i} \det B_{ks}''(z).$$

The matrices $B_{ks}''(z)$ are of the form

$$
B_{ks}''(z) = \begin{bmatrix} 0 \ . \ . \ . \ 0 & -1 \\ & 0 \\ U_{ks}(z) & \vdots \\ & 0 \end{bmatrix}.
$$

Here $U_{ks}(z)$ is an upper triangular matrix having $z - a_l$ ($l \neq i$) in the main diagonal. If we denote the order of $B_{ks}''(z)$ by n_{ks}, then

(11) $$ det \, B_{ks}''(z) = (-1)^{n_{ks}} det \, U_{ks}(z). $$

From (7) − (10) we get

$$
P_k(z) = (-1)^{n+k} det \, B_k'(z) \prod_{s=j+1}^{r_j} (-1)^{n_{ks}} det \, U_{ks}(z).
$$

Consequently, $P_k(z)$ contains all the factors $z - a_l$ with $l \neq i$. The factor $z - a_i$ is included only in $B_k'(z)$, and − by definition − exactly j times. As for the sign, the exponent of − 1 is equal to

$$
n + k + \sum_{s=j+1}^{r_j} n_{ks} = n + k + n - k = 2n,
$$

therefore,

$$
P_k(z) = \prod_{\substack{s=1 \\ s \neq i}}^{m} (z - a_s)^{r_s} \cdot (z - a_i)^j = h_{ij}^*(z)
$$

which was to be proved. ⊠

Remarks.

1. If $A_i = 0$ for $1 \leq i \leq n-1$, then we have the system $h_{1j}(z) = z^j$, $0 \leq j \leq n-2$, and B is nothing else than Frobenius' companion matrix (see [1]).

2. If $\{A_i\}_{i=1}^{n-1}$ are distinct then the interpolatory polynomials of Hermite are reduced to those of Lagrange (i.e. to h_{i0}, $1 \leq i \leq n-1$), and B is the matrix in [2], or [3].

3. More can be said about the numbers $\{b_{nk}\}_{k=1}^{n-1}$. They are linear combinations of the derivatives of P taken at the given points. Indeed, by the basic formula of the Hermite interpolation and by (7),

$$
\Delta = \sum_{i=1}^{m} \sum_{j=0}^{r_i-1} \frac{\Delta^{(j)}(a_i)}{j!} h_{ij} = \sum_{i=1}^{m} \sum_{j=0}^{r_i-1} \frac{\Delta^{(j)}(a_i)}{j!} \sum_{s=j}^{r_i-1} \gamma_{ijs} h_{is}^* =
$$

$$
= \sum_{i=1}^{m} \sum_{s=0}^{r_i-1} \left\{ \sum_{j=0}^{s} \gamma_{ijs} \frac{\Delta^{(j)}(a_i)}{j!} \right\} h_{is}^*.
$$

Since

$$\prod_{i=1}^{n-1} (z - A_i) = \prod_{i=1}^{m} (z - a_i)^{r_i},$$

on the ground of (1) and (4) we have

$$\Lambda^{(j)}(a_i) = -P^{(j)}(a_i), \quad 0 \le j \le v_i - 1, \quad 1 \le i \le m.$$

Thus, the coefficient of h_{ij}^* (i.e., one of the b_{nk}'s) is the linear combination of the numbers $P^{(s)}(a_i)$, $s = 0, 1, \ldots, j$.

4. Let $P(z) = \prod\limits_{i=1}^{m} (z - a_i)^{r_i}$ (i.e. we choose the A_i's to be the roots of P with the respective multiplicities). Then, by the previous remark, we have $b_{nk} = 0$, $1 \le k \le n-1$, i.e. B becomes to an upper triangular matrix (comp. with the Jordan' canonical form of matrices!).

5. Let P be a real polynomial with distinct real roots α_i, $1 \le i \le n$, and suppose that they are separated by the numbers A_i, $1 \le i \le n-1$:

$$\alpha_1 < A_1 < \alpha_2 < \ldots \alpha_{n-1} < A_{n-1} < \alpha_n.$$

Then the matrix B obtained from the Theorem can be symmetrized. Indeed

$$b_{ni} = -\frac{P(A_i)}{\prod\limits_{\substack{j=1 \\ j \ne i}}^{n-1} (A_i - A_j)}, \quad 1 \le i \le n-1.$$

(see Remark 2.).
Since sign $P(A_i) = (-1)^i$, $1 \le i \le n-1$ (P is monic!) obviously $b_{ni} > 0$ for every i. Thus with

$$d_i = \sqrt{b_{ni}}, \quad 1 \le i \le n-1, \quad d_n = 1, \quad D = diag(d_i)$$

the matrix DBD^{-1} will be symmetric.

6. Since the h_{ij}^*'s are relative primes, similarly to the Frobenius-matrix, the characteristic and minimal polynomials of B coincide (see [1]).

REFERENCES

[1] P. Lancaster, Theory of Matrices, Academic Press, New York, 1969.
[2] L. Laszlo, Matrix Methods for Finding Roots of Polynomials, Annales Univ. Sci. Budapest, Sectio Comp., II. (1978), pp. 29−37.
[3] L. Mirsky, Matrices with Prescribed Characteristic Roots and Diagonal Elements, J. London Math. Soc., 33 (1958), pp. 14−21.