COMPOSITIO MATHEMATICA

H. S. THURSTON

On factoring a matric polynomial with scalar coefficients

Compositio Mathematica, tome 6 (1939), p. 235-238
<http://www.numdam.org/item?id=CM_1939__6__235_0>

On factoring a matric polynomial with scalar coefficients

by

H. S. Thurston
University of Alabama

1. *Introduction.* In a paper published about four years ago [1]), A. Hermann stated the following theorem: *If μ is the number of characteristic values of a matrix A, then any polynomial of degree t, with simple zeros, breaks up in $t^{\mu-1}$ ways into a product of t linear factors in the ring $R(A)$.* The purpose of this note is to call attention to the fact that in the proof of the theorem all ways of factoring were not taken into consideration, and to show that the total number of ways is $(t!)^{\mu-1}$.

2. *Determination of factors.* Let the minimum equation of a given matrix A be $\psi(\lambda) = \prod_{i=1}^{\mu}(\lambda - \alpha_i)^{\nu_i} = 0$. Corresponding to the μ distinct roots α_i, there exist μ linearly independent, non-zero matrices e_i in the ring $R(A)$, satisfying the relations

$$\Sigma e_i = 1, \quad e_i^2 = e_i, \quad e_i e_j = 0, \quad (i \neq j).$$

These are called the principal idempotent elements [2]) of A.

Now let $F(\lambda)$ be any polynomial with scalar coefficients having simple zeros $s_1, s_2, \ldots, s_t$. Choosing any μ of these zeros (repetitions being permitted), $s_{i_1}, \ldots, s_{i_\mu}$, the function

$$X_0(A) = s_{i_1}e_1 + s_{i_2}e_2 + \ldots + s_{i_\mu}e_\mu$$

is a root of $F(X) = 0$. If each subscript i_k is increased by i $(i = 1, 2, \ldots, t-1)$, and reduced modulo t, we obtain a matrix $X_i(A)$ which is also a root of $F(X) = 0$, and is called by Hermann the i-th conjugate of X_0. The matrices X_i are commutative and it is readily seen that their elementary symmetric functions are

---

1) A. Hermann, Über Matrixgleichungen und die Zerlegung von Polynomen in Linearfactoren [Compositio Mathematica 1 (1934), 284—302].

2) J. H. M. Wedderburn, Lectures on Matrices [Amer. Math. Soc. Publications (1934)], 29.

the corresponding elementary symmetric functions of the roots of $F(\lambda)=0$. Hence

$$F(X) = \Pi_{i=0}^{t-1}(X - X_i).$$

As the total number of matrices $X(A)$ thus formed is t^{μ}, and as each factorization of $F(X)$ requires t of them, Hermann concludes that there are $t^{\mu-1}$ ways of factoring $F(X)$ in the ring $R(A)$.

For convenience we shall denote the matrix X_0 symbolically by $(s_{i_1}s_{i_2}\ldots s_{i_\mu})$, the linear factor $X-X_0$ by $[s_{i_1}s_{i_2}\ldots s_{i_\mu}]$, with similar symbols for X_i and its corresponding linear factor. It is obvious that all the factorizations obtained by Hermann may be obtained by letting the first element in each initial matrix X_0 be s_{i_1} and permuting the other elements in all possible ways. Thus X_0 may be chosen in $t^{\mu-1}$ ways. However it is not necessary to employ Hermann's rule to obtain the remaining conjugates. It is sufficient to choose these matrices so that the elements occupying the k-th position in the t conjugates shall be distinct roots of $F(\lambda)=0$. Keeping X_0 fixed, it is obvious that no distinct factorizations of $F(X)$ will be lost if we let the first element of X_1 be s_{i_2}. The remaining elements of X_1, and hence X_1 itself, may be chosen in $(t-1)^{\mu-1}$ ways. Similarly, for any fixed X_0 and X_1, we may choose X_2 in $(t-2)^{\mu-1}$ ways. Finally, X_{t-1} may be chosen in one way. Hence, we have

THEOREM 1. *If μ is the number of distinct latent roots of a matrix A, then any polynomial $F(\lambda)$ of degree t, with simple zeros, breaks up in $(t!)^{\mu-1}$ ways in the ring $R(A)$.*

Below are shown, by way of example, the 36 ways of factoring when $t=\mu=3$, and $F(\lambda)=(\lambda-1)(\lambda-2)(\lambda-3)$. The first factorization in each line is obtained by Hermann's rule, and the corresponding matrices may be called conjugates in the sense of Hermann. It has been shown [3]) that when A has simple latent roots, the sets (123), (231), (312) and (132), (213), (321) are conjugates in the sense of Taber, while all other sets whose leading matrix is (123) or (132) are conjugates in the sense of Franklin.

[111][222][333], [111][223][332], [111][233][322], [111][232][323],
[112][223][331], [112][231][323], [112][221][333], [112][233][321],
[113][221][332], [113][222][331], [113][231][322], [113][232][321],

[3]) Matric conjugates in a ring $R(A)$ [Bulletin Amer. Math. Soc. **44** (1938), 258—261].

[121][232][313],	[121][233][312],	[121][212][333],	[121][213][332],
[122][233][311],	[122][211][333],	[122][231][313],	[122][213][331],
[123][231][312],	[123][212][331],	[123][232][311],	[123][211][332],
[131][212][323],	[131][213][322],	[131][222][313],	[131][223][312],
[132][213][321],	[132][211][323],	[132][221][313],	[132][223][311],
[133][211][322],	[133][212][321],	[133][222][311],	[133][221][312].

3. *Similar factorizations.* Two different sets of factors will be called *similar* if there exists a non-singular matrix T which transforms each matrix of the first set into a matrix of the second set, the same T serving for every transformation. In general, the factorizations given by Theorem 1 are not all dissimilar. For example, if A has simple latent roots, then among the above 36 sets there are 13 sets which are mutually dissimilar. The scalar set (111)(222)(333) is, of course, not transformable into any other set. The sets (123)(231)(312) and (132)(213)(321) are similar, while the remaining sets are similar in groups of three, as, for example, (112)(221)(333), (121)(212)(333), and (211)(122)(333). Although the number of dissimilar sets in the case of a general matrix A has not been investigated, in the special case to which we have just referred we may pick out similar sets by means of the following theorem.

THEOREM 2. *If A has simple latent roots* [4]), *two sets $\{X_i\}$ and $\{Y_i\}$ are similar if the symbols for corresponding matrices differ only by a cyclic rearrangement of the numbers in the symbols.*

Let the sets $\{X_i\}$ and $\{Y_i\}$ consist of matrices $X_0, X_1, \ldots, X_{t-1}$ and $Y_0, Y_1, \ldots, Y_{t-1}$ respectively. Since the distinct numbers within any matric symbol $(s_{i_1} s_{i_2} \ldots s_{i_n})$ are the roots of the minimum equation of the matrix [5]), the elementary divisors of all matrices under consideration are linear and any such matrix can be transformed to diagonal form [6]) with diagonal elements $s_{i_1}, s_{i_2}, \ldots, s_{i_n}$. Also, since X_i and Y_i have the same elementary divisors, there exists a non-singular matrix T such that $T^{-1}X_iT = Y_i$. We shall prove that the same matrix T will serve for all values of i.

We shall consider the case where Y_i is obtained cyclicly from

[4]) With slight modification, the proof may be carried through if A has a Segre characteristic $[kk \ldots k]$.

[5]) Wedderburn, loc. cit. p. 26.

[6]) C. C. MACDUFFEE, Theory of Matrices [Berlin, Springer (1933)], p. 70.

X_i by replacing i_k by i_{k-1}. The general case may be proved by repetition of the argument.

Let

$$X_0 = (s_{i_1} s_{i_2} \ldots s_{i_n}) = s_{i_1} e_1 + s_{i_1} e_2 + \ldots + s_{i_n} e_n$$

and

$$Y_0 = (s_{i_n} s_{i_1} \ldots s_{i_{n-1}}) = s_{i_n} e_1 + s_{i_1} e_2 + \ldots + s_{i_{n-1}} e_n.$$

Then if P transforms X_0 into the diagonal matrix A_0, we have

$$\begin{aligned} P^{-1} X_0 P &= s_{i_1} P^{-1} e_1 P + \ldots + s_{i_n} P^{-1} e_n P \\ &= s_{i_1} \varepsilon_1 + \ldots + s_{i_n} \varepsilon_n \\ &= A_0, \end{aligned}$$

where the ε_i are the principal idempotent elements of A_0, having unity as the i-th diagonal element and zeros elsewhere.

Similarly, P transforms Y_0 into a diagonal matrix B_0. Let Q transform A_0 into B_0. Then we have

$$Y_0 = P B_0 P^{-1} = P Q^{-1} A_0 Q P^{-1} = P Q^{-1} P^{-1} X_0 P Q P^{-1}.$$

Since $P^{-1} e_i P = \varepsilon_i$, it follows that P will transform each X_1 into its corresponding diagonal matrix A_i, and Y_i into B_i. And since A_i is transformed into B_i by an interchange of rows and of columns (the same for every i), this can always be effected by the same matrix Q. Hence $T = PQP^{-1}$ will transform each X_i into the corresponding Y_i, and the two sets are similar.

(Received March 31st, 1938.)