

FACTORIZATION OF SYMMETRIC MATRIX POLYNOMIALS

V. A. Yakubovich

1970

SovietRxiv

View the original and related papers at <https://sovietrxiv.org/items/ru-197001.25870>

Source: Math-Net.Ru and CyberLeninka. Machine translation. Verify with the original.

Abstract

Full Text

UDC 512.83+62.50

MATHEMATICS

V. A. YAKUBOVICH

FACTORIZATION OF SYMMETRIC MATRIX POLYNOMIALS

(Presented by Academician V. I. Smirnov on 9 III 1970)

1°. Let

$$A(\lambda) = A_0 + \lambda A_1 + \dots + \lambda^N A_N$$

be a matrix polynomial (A_j are complex matrices of order $m \times m$, λ is a complex variable). By $A(\lambda)^\nabla$ we shall denote the matrix polynomial

$$A(\lambda)^\nabla = A^*(-\bar{\lambda}) = A_0^* - \lambda A_1^* + \dots + (-\lambda)^N A_N^*.$$

(Here the asterisk denotes Hermitian conjugation.) The well-known theorem on factorization (i.e., on the possibility of representing $A(\lambda)$ in the form $A(\lambda) = X(\lambda)X(\lambda)^\nabla$, where $X(\lambda)$ is a matrix polynomial—see Theorem 4 below) has in recent years, in the works of Kalman, Popov, and others, been unexpectedly applied in the theory of absolute stability and in the theory of optimal control (see, for example, ⁽¹⁾).

It can be shown that many problems in the theory of differential games also lead to the need to factor a matrix polynomial, but to a factorization of a more general form, namely

$$A(\lambda) = X(\lambda)CX(\lambda)^\nabla,$$

where $C = C^*$ is a constant matrix, $\det C \neq 0$. Below a solution of this more general problem is given (Theorems 1-3). From Theorem 2 below, Theorem 4 on factorization of the form $A(\lambda) = X(\lambda)X(\lambda)^\nabla$ is derived. This proof of Theorem 4 is substantially simpler than the direct proof of Theorem 4 given, for example, in the book ⁽¹⁾.

Theorem 1. *In order that a given matrix polynomial (m.p.) $A(\lambda)$ be representable in the form $A(\lambda) = X(\lambda)CX(\lambda)^\nabla$, where $C = C^*$ is some nonsingular matrix and $X(\lambda)$ is some m.p. such that $\det X(\lambda) \neq 0$ for $\operatorname{Re} \lambda = 0$, it is necessary and sufficient that the following hold: (a) $A(\lambda)^\nabla = A(\lambda)$; (b) $\det A(\lambda) \neq 0$ for $\operatorname{Re} \lambda = 0$.*

The necessity of conditions (a), (b) is obvious. Sufficiency follows from Theorem 2 given below.

We shall call an m.p. (or simply a polynomial) real if its coefficients are real matrices (numbers).

Theorem 2. Suppose that $A(\lambda) = A(\lambda)^\nabla$, $\det A(\lambda) \neq 0$. Let $i_1(\lambda), \dots, i_m(\lambda)$ be the invariant polynomials ⁽²⁾ of the m.p. $A(\lambda)$, ordered so that $i_k(\lambda)$ divides $i_{k+1}(\lambda)$, $k = 1, \dots, m-1$. Suppose that each of the polynomials $i_1(\lambda), \dots, i_m(\lambda)$ either has no zeros on the imaginary axis (which will be the case if $\det A(\lambda) \neq 0$ for $\operatorname{Re} \lambda = 0$), or has zeros on the imaginary axis, but the multiplicity of each such zero is an even number. Then the polynomials $i_k(\lambda)$ admit a factorization

$$i_k(\lambda) = \pm x_k(\lambda) x_k(\lambda)^\nabla, \quad k = 1, \dots, m,$$

where $x_k(\lambda)$ are polynomials and, in particular, such (this factorization will be called admissible) that $x_k(\lambda)$ divides $x_{k+1}(\lambda)$, $k = 1, \dots, m-1$. For any admissible factorization of the set $i_1(\lambda), \dots, i_m(\lambda)$ there exist a nonsingular matrix $C = C^*$ and an m.p. $X(\lambda)$ such that

$$A(\lambda) = X(\lambda) C X(\lambda)^\nabla, \quad \det X(\lambda) = \pm x_1(\lambda) \cdots x_m(\lambda). \quad (1)$$

If $A(\lambda)$ is a real m.p., then for any admissible factorization of the set $i_1(\lambda), \dots, i_m(\lambda)$ for which $x_1(\lambda), \dots, x_m(\lambda)$ are real-

polynomials, there exist a Hermitian matrix $C = C^*$ and a Hermitian m.p. $X(\lambda)$ satisfying relations (1).

Theorem 3. Let $A(\lambda) = A(\lambda)^\nabla$, $\det A(\lambda) \neq 0$, and, if $\det A(\lambda)$ has zeros on the imaginary axis, let all invariant polynomials of $A(\lambda)$ have zeros on the imaginary axis only of even multiplicity. Let

$$\det A(\lambda) = \pm \delta(\lambda) \delta(\lambda)^\nabla,$$

where the polynomials $\delta(\lambda)$ and $\delta(\lambda)^\nabla$ are relatively prime. There exists an m.p. $X(\lambda)$ and a constant matrix $C = C^*$ (Hermitian when the m.p. $A(\lambda)$ is Hermitian) such that

$$A(\lambda) = X(\lambda) C X(\lambda)^\nabla$$

and

$$\det X(\lambda) = \pm \delta(\lambda).$$

From Theorem 2 the following well-known result is derived quite simply (see, for example, ⁽¹⁾).

Theorem 4. Let $A(\lambda) = A(\lambda)^\nabla$. For the existence of an m.p. $X(\lambda)$ such that

$$A(\lambda) = X(\lambda) X(\lambda)^\nabla,$$

it is necessary and sufficient that for $\operatorname{Re} \lambda = 0$ the Hermitian matrix $A(\lambda)$ be nonnegative.* If $A(\lambda)$ is a Hermitian m.p. and $A(\lambda) \geq 0$ for $\operatorname{Re} \lambda = 0$, then there exists a Hermitian m.p. $X(\lambda)$ such that

$$A(\lambda) = X(\lambda) X(\lambda)^\nabla.$$

In both cases the polynomial $X(\lambda)$ can be chosen so that $\det X(\lambda)$ has no zeros in the domain $\operatorname{Re} \lambda > 0$.

3. Auxiliary propositions. Below, in the proofs, the argument of matrix polynomials is omitted. By $\operatorname{st} A$ is denoted the degree of the m.p. $A = A(\lambda)$ under the condition that A is a nonzero m.p. For the zero m.p. $A(\lambda) \equiv 0$ we set $\operatorname{st} A(\lambda) = -1$. By $\operatorname{diag}(a_1, \dots, a_n)$ is denoted the diagonal matrix with diagonal entries $a_1, \dots, a_n$. Below we use the obvious relations

$$(A + B)^\nabla = A^\nabla + B^\nabla, \quad (AB)^\nabla = B^\nabla A^\nabla, \quad \det(A^\nabla) = (\det A)^\nabla.$$

Lemma 1. Let $A(\lambda)^\nabla = A(\lambda)$, and let $i_1(\lambda), \dots, i_r(\lambda)$ be the invariant polynomials of the matrix $A(\lambda)$. Then

$$i_k(\lambda)^\nabla = \pm i_k(\lambda), \quad k = 1, \dots, r.$$

According to (2) we have $A = PDQ$, where P, Q are m.p.'s, $\det P \equiv 1$, $\det Q \equiv \operatorname{const} \neq 0$,

$$D(\lambda) = \operatorname{diag}(i_1, \dots, i_r, 0, \dots, 0).$$

Since

$$A = A^\nabla = Q^\nabla D^\nabla P^\nabla, \quad \det P^\nabla = 1, \quad \det Q^\nabla \equiv \operatorname{const} \neq 0,$$

the invariant polynomials of the m.p. A and D^∇ coincide. The invariant polynomials (with leading coefficient 1) of the m.p. D^∇ are $\pm i_k^\nabla$. Assuming (without loss of generality) that the i_k are numbered so that i_{k+1} is divisible by i_k ($k = 1, \dots, r-1$), we obtain

$$i_k = \pm i_k^\nabla.$$

Lemma 2. Let

$$A(\lambda)^\nabla = A(\lambda) = \|a_{jh}(\lambda)\|, \quad \det A(\lambda) \equiv \operatorname{const} \neq 0$$

and $a_{11}(\lambda) \equiv 0$. There exists an m.p. $Y(\lambda)$ (Hermitian when the m.p. $A(\lambda)$ is Hermitian) such that $\det Y(\lambda) \equiv 1$ and such that, for the matrix

$$B(\lambda) = Y(\lambda)A(\lambda)Y(\lambda)^\nabla,$$

the entries of the upper row (from left to right) and of the left column (from top to bottom) are the numbers $1, 0, \dots, 0$.

Separating the upper row and the left column, represent the matrices $A, A^{-1}, Y, B = YAY^\nabla$ in the form

$$A = \begin{Bmatrix} 0 & a^\nabla \\ a & A_1 \end{Bmatrix}, \quad Y = \begin{Bmatrix} \xi & x^\nabla \\ y & Y_1 \end{Bmatrix}, \quad A^{-1} = \begin{Bmatrix} \mu & l^\nabla \\ n & N \end{Bmatrix}, \quad B = \begin{Bmatrix} \beta & b^\nabla \\ b & B_1 \end{Bmatrix}.$$

Here μ, l, n, N are polynomials, since $\det A = \text{const} \neq 0$ and $A_1^\nabla = A_1$, $N^\nabla = N$. The polynomials ξ, x, y, Y_1 must be chosen so that the following be fulfilled:

$$(I) \quad \det Y = (\xi - x^\nabla Y_1^{-1} y) \det Y_1 \equiv 1;$$

$$(II) \quad b = ya^\nabla x + Y_1(a\xi^\nabla + A_1 x) \equiv 0;$$

$$(III) \quad \beta = \xi a^\nabla x + x^\nabla a\xi^\nabla + x^\nabla A_1 x \equiv 1.$$

It is easy to verify that these relations are satisfied for

$$x = n, \quad Y_1 = I_{m-1}, \quad \xi = (1 - n^\nabla A_1 n)/2, \quad y = -a\xi - A_1 n.$$

In doing so one should take into account the relation

$$a^\nabla n = 1,$$

which holds since $AA^{-1} = I_m$.

An important role in the further proof is played by Lemma 3. Below the proof of this lemma is given for matrices of order $m = 2$. For matrices of order $m > 2$ the proof of Lemma 3 was obtained by B. D. Lobachevskii.**

* This means that $z^* A(\lambda) z \geq 0$ for $\text{Re } \lambda = 0$ and for any m -vector z . Let G be a matrix of order $m \times m$. Below we shall write $G > 0$ or $G \geq 0$ if $G = G^*$ and, respectively, $z^* G z > 0$ or $z^* G z \geq 0$ for all $z \neq 0$. By I_m below is denoted the identity $m \times m$ matrix.

** The author expresses gratitude to I. A. Lebedev for useful remarks on Lemma 3.

Lemma 3. Let $A(\lambda) = A(\lambda)^\nabla = \|a_{jh}\|$, $B(\lambda) = Y(\lambda)A(\lambda)Y(\lambda)^\nabla = \|\beta_{jh}\|$. Either (a) there exists an m.m. $Y(\lambda)$, $\det Y(\lambda) \equiv \pm 1$, such that $\beta_{11} = 0$, or (b) there exists an m.m. $Y(\lambda)$, $\det Y(\lambda) = \pm 1$, such that $\text{st } \beta_{jh} < \text{st } \beta_{jj}$ for $j \neq h$, $0 \leq \text{st } \beta_{11} \leq \text{st } \beta_{22} \leq \dots \leq \text{st } \beta_{nn}$.

Let $m = 2$, $Y = \|\eta_{jh}\|$. For $\eta_{11} = \eta_{22} = 0$, $\eta_{12} = \eta_{21} = 1$ (the Y_1 -transformation) we have $\beta_{11} = a_{22}$, $\beta_{22} = a_{11}$. For $\eta_{11} = \eta_{22} = 1$, $\eta_{12} = 0$, $\eta_{21} = -\eta$ (the Y_2 -transformation) we have $\beta_{21} = a_{21} - \eta a_{11}$, $\beta_{12} = \beta_{21}^\nabla$. If $\text{st } a_{11} > \text{st } a_{22}$, then after the Y_1 -transformation we obtain $\text{st } \beta_{11} < \text{st } \beta_{22}$. If $\text{st } a_{11} \leq \text{st } a_{22}$, then we put $Y_1 = I$. If it turns out that $\text{st } \beta_{21} \geq \text{st } \beta_{11}$, then, dividing β_{21} by β_{11} , we obtain $\beta_{21} = \eta \beta_{11} + \beta'_{21}$, where $\text{st } \beta'_{21} < \text{st } \beta_{21}$. Carrying out the Y_2 -transformation with this η , we obtain an m.m. $B' = \|\beta'_{jh}\|$, where $\beta'_{11} = \beta_{11}$, $\text{st } \beta'_{21} < \text{st } \beta'_{11}$. The possible cases are: (I) $\text{st } \beta'_{22} \geq \text{st } \beta'_{11}$; (II) $\beta'_{22} = 0$; (III) $\text{st } \beta'_{22} < \text{st } \beta'_{11}$, $\beta'_{22} \neq 0$. In case (I) the lemma is proved for $Y = Y_2 Y_1$. In case (II) the lemma is proved for $Y = Y_1 Y_2 Y_1$. In case (III), applying the Y_1 -transformation, we arrive at the initial situation, but the degree of the element with indices 11 will be less than at the beginning. We repeat all operations. If $\text{st } a_{11} = k$, then the process terminates after $k' \leq k$ repetitions.

Lemma 4. Let $A(\lambda) = A(\lambda)^\nabla$, $\det A(\lambda) = \text{const} \neq 0$, $B(\lambda) = Y(\lambda)A(\lambda)Y(\lambda)^\nabla = \|\beta_{jh}\|$. There exists an m.m. $Y(\lambda)$, $\det(\lambda) \equiv \pm 1$ (real for real $A(\lambda)$), such that $\text{st } \beta_{jh} < \text{st } \beta_{jj}$ for $j \neq h$.

Apply Lemma 3. In case (a) the assertion is proved. In case (b), by Lemma 2 there exists an m.m. Y_1 such that the m.m. $B = Y_1 A Y_1$ has the form indicated in Lemma 2. Let B_1 be the m.m. obtained from B by deleting the first row and first column. Obviously, $B_1^\nabla = B_1$ and $\det B_1 = \text{const} \neq 0$. If for B_1 case (b) of Lemma 3 holds, then we have $B_2 = Y_2 B_1 Y_2^\nabla = \|\beta_{jh}\|$, $\text{st } \beta_{jh} < \text{st } \beta_{jj}$ for $j \neq h$, where Y_2 is an m.m. In this case the assertion is proved for $Y = Y_1 \text{diag}(1, Y_2)$. If for B_2 case (a) of Lemma 3 holds, then, applying Lemma 2, we again lower the order of the m.m. After a finite number of times we obtain an m.m. of the required form.

Lemma 5. If $A(\lambda) = A(\lambda)^\nabla$, $\det A(\lambda) = \text{const} \neq 0$, then there exist a constant matrix $C = C^*$ and an m.m. $U(\lambda)$ (real for real m.m. $A(\lambda)$) such that $A(\lambda) = U(\lambda)CU(\lambda)^\nabla$, $\det U(\lambda) \equiv \pm 1$.

By Lemma 4, $A = UBU^\nabla$, where $U = Y^{-1}$ is an m.m., $\det U = \pm 1$. Since $\text{st } \beta_{jh} < \text{st } \beta_{jj}$ for $j \neq h$, the leading term of the polynomial $\det B$ is contained in the term $\beta_{11} \dots \beta_{nn}$. Since $\det B = \det A = \text{const}$, it follows that $\beta_{11} = \text{const}, \dots, \beta_{nn} = \text{const}$, and hence $\beta_{jh} = 0$ for $j \neq h$, $B = C = \text{const}$. For $\text{Re } \lambda = 0$ we have $A = UCU^*$, $A = A^*$. Therefore $C = C^*$.

4°. Proof of Theorem 2. According to (2) we have $A = PDQ$, where PQ are m.m., $D = \text{diag}(i_1, \dots, i_m)$. By Lemma 1, the zeros of the polynomials i_h not lying on the imaginary axis are situated (with account of their multiplicities) symmetrically with respect to the imaginary axis. Since purely imaginary zeros have even multiplicity, $i_h = \varepsilon_h x_h x_h^\nabla$, where x_h are polynomials, $\varepsilon_h = \pm 1$. Let the indicated factorization be admissible. Put $D_0 = \text{diag}(x_1, \dots, x_m)$, $E = \text{diag}(\varepsilon_1, \dots, \varepsilon_m)$. Then $D = D_0 E D_0^\nabla$, $A = R K R^\nabla$, where $R = P D_0 K = D_0^\nabla S D_0^{\nabla-1}$, $S = E Q P^{\nabla-1}$. Obviously, $S = \|s_{jh}\|$ is an m.m. We shall show that K is an m.m. Since $A^\nabla = A$ and $\det R \neq 0$, it follows that $K = K^\nabla$. Let $K = \|k_{jh}\|$. We have $k_{jh} = x_j^\nabla s_{jh} x_h^{\nabla-1}$. Since the factorization is admissible, k_{jh} are polynomials for $j < h$. Since $K = K^\nabla$, $k_{hj} = k_{jh}^\nabla$, hence k_{jh} are polynomials also for $j < h$. Since $\det S = \text{const} \neq 0$, we also have $\det K = \text{const} \neq 0$. By Lemma 5, $K = UCU^\nabla$, where U is an m.m. Therefore $A = X C X^\nabla$, where $X = R U$ is an m.m. Since $\det U = \pm 1$, $\det X = \pm \det R = \pm \det D_0 = \pm x_1 \dots x_m$. Thus the relations (1) have been established. If A is a real m.m., then P, Q, i_h , and x_h (under the indicated factorization) are real polynomials. Then R, S, K are real m.m. By Lemma 5, C and U are real, and hence X is a real m.m.

Proof of Theorem 3. Suppose first that $\delta(\lambda)$ has no roots in the domain $\text{Re } \lambda > 0$. Let $i_h(\lambda) = \prod (\lambda - \lambda_j)^{r_j}$. Taking as $x_h(\lambda)$ the product of the factors $(\lambda - \lambda_j)^{r_j}$ with $\text{Re } \lambda_j < 0$ and

$(\lambda - \lambda_j)^{r_j/2}$ with $\text{Re } \lambda_j = 0$, we obtain, obviously, an admissible factorization such that $\pm x_1, \dots, x_m = \delta(\lambda)$. By what was proved above, we have $\det X \equiv \pm \delta$.

In the general case the proof is analogous—it is necessary to take as $x_n(\lambda)$ the greatest common divisor $i_h(\lambda)$ and $\delta(\lambda)$.

Proof of Theorem 4. Necessity is obvious.

Sufficiency. If $A > 0$ for $\operatorname{Re} \lambda = 0$, then, by Theorem 2, $A = ZCZ^\nabla$, where Z is a m.m., C, Z are real for real m.m. A . Clearly, $C = C^* > 0$. Putting $X = ZC^{1/2}$, we obtain $A = XX^\nabla$. In this case $\deg X = (\deg A)/2$. Indeed, let $X(\lambda) = \lambda^M X_0 + \dots + X_M$, $X_0 \neq 0$, $A(\lambda) = \lambda^N A_0 + \dots + A_N$, $A_0 \neq 0$. Clearly, $N \leq 2M$. If $N < 2M$, then $X_0 X_0^* = 0$, $\operatorname{Sp} X_0 X_0^* = |X_0|^2 = 0$, $X_0 = 0$. Therefore $N = 2M$, as asserted. Let $A(\lambda) \geq 0$ for $\operatorname{Re} \lambda = 0$. We use the device from ¹. By what has been proved, for any number $\varepsilon > 0$ we have $A(\lambda) + \varepsilon I_m = X_\varepsilon(\lambda) X_\varepsilon(\lambda)^\nabla$, where $X_\varepsilon(\lambda)$ is a m.m., $\deg X_\varepsilon(\lambda) = N/2$. For $\operatorname{Re} \lambda = 0$ we have $\operatorname{Sp}(A + \varepsilon I) = |X_\varepsilon(\lambda)|^2$. Therefore the elements $X_\varepsilon(\lambda)$ are bounded for $0 < \varepsilon < \varepsilon_0$, $\operatorname{Re} \lambda = 0$, $|\operatorname{Re} \lambda| \leq r$. Let $\varepsilon_n \rightarrow 0$. Choosing from the sequence $X_{\varepsilon_n}(\lambda)$ a subsequence converging to some $X(\lambda)$ uniformly for $\operatorname{Re} \lambda = 0$, $|\operatorname{Re} \lambda| \leq r$, we obtain $A(\lambda) = X(\lambda) X(\lambda)^\nabla$. Since the degrees $X_\varepsilon(\lambda)$ are uniformly bounded, $X(\lambda)$ is a m.m.

Leningrad State University
named after A. A. Zhdanov

Received
13 X 1970

REFERENCES

¹ V. M. Popov. *Hiperstabilitatea sistemelor automate*, România, 1966. ² F. R. Gantmakher, *Theory of Matrices*, Moscow, ch. 6, 1954.

Note: Figure translations are in progress. See original paper for figures.

Source: Math-Net.Ru and CyberLeninka. Machine translation. Verify with the original.