Доклады Академии наук СССР
1970. Том 194, № 3

УДК 512.83+62.50 *МАТЕМАТИКА*

В. А. ЯКУБОВИЧ

ФАКТОРИЗАЦИЯ СИММЕТРИЧНЫХ МАТРИЧНЫХ МНОГОЧЛЕНОВ

(Представлено академиком В. И. Смирновым 9 III 1970)

1°. Пусть $A(\lambda)=A_0+\lambda A_1+\ldots+\lambda^N A_N$ — матричный многочлен (A_j — комплексные матрицы порядка $m\times m$, λ — комплексное переменное). Через $A(\lambda)^\nabla$ будем обозначать матричный многочлен $A(\lambda)^\nabla=A^*(-\lambda)=A_0^*-\lambda A_1^*+\ldots+(-\lambda)^N A_N^*$. (Здесь звездочка означает эрмитово сопряжение.) Хорошо известная теорема о факторизации (т. е. о возможности представления $A(\lambda)$ в виде $A(\lambda)=X(\lambda)X(\lambda)^\nabla$, где $X(\lambda)$ — матричный многочлен — см. ниже теорему 4) в последние годы в работах Калмана, Попова и др. неожиданно применена в теории абсолютной устойчивости и в теории оптимального управления (см., например, ([1])).

Можно показать, что многие задачи теории дифференциальных игр также приводят к необходимости факторизации матричного многочлена, однако к факторизации более общего вида, именно $A(\lambda)=X(\lambda)CX(\lambda)^\nabla$, где $C=C^*$ — постоянная матрица, $\det C\neq 0$. Ниже приводится решение этой более общей задачи (теоремы 1—3). Из теоремы 2 ниже выводится теорема 4 о факторизации вида $A(\lambda)=X(\lambda)X(\lambda)^\nabla$. Это доказательство теоремы 4 существенно проще, чем прямое доказательство теоремы 4, приведенное, например, в книге ([1]).

Теорема 1. *Для того чтобы заданный матричный многочлен (м.м.) $A(\lambda)$ можно было представить в виде $A(\lambda)=X(\lambda)CX(\lambda)^\nabla$, где $C=C^*$ — некоторая неособая матрица и $X(\lambda)$ — некоторый м.м. такой, что $\det X(\lambda)\neq 0$ при $\operatorname{Re}\lambda=0$, необходимо и достаточно, чтобы было выполнено:* (а) $A(\lambda)^\nabla=A(\lambda)$; (б) $\det A(\lambda)\neq 0$ *при* $\operatorname{Re}\lambda=0$.

Необходимость условий (а), (б) очевидна. Достаточность следует из приведенной ниже теоремы 2.

Будем называть м.м. (или просто многочлен) вещественным, если его коэффициенты являются вещественными матрицами (числами).

Теорема 2. *Предположим, что $A(\lambda)=A(\lambda)^\nabla$, $\det A(\lambda)\not\equiv 0$. Пусть $i_1(\lambda),\ldots,i_m(\lambda)$ — инвариантные многочлены ([2]) м.м. $A(\lambda)$, упорядоченные так, что $i_k(\lambda)$ делится на $i_{k+1}(\lambda)$, $k=1,\ldots,m-1$. Предположим, что каждый из многочленов $i_1(\lambda),\ldots,i_m(\lambda)$ либо не имеет нулей на мнимой оси (что будет, если $\det A(\lambda)\neq 0$ при $\operatorname{Re}\lambda=0$), либо имеет нули на мнимой оси, но кратность каждого такого нуля является четным числом. Тогда многочлены $i_k(\lambda)$ допускают факторизацию $i_k(\lambda)=\pm x_k(\lambda)x_k(\lambda)^\nabla$, $k=1,\ldots,m$, где $x_k(\lambda)$ — многочлены и, в частности, такую (эту факторизацию будем называть допустимой), что $x_k(\lambda)$ делится на $x_{k+1}(\lambda)$, $k=1,\ldots,m-1$. Для любой допустимой факторизации набора $i_1(\lambda),\ldots,i_m(\lambda)$ существует неособая матрица $C=C^*$ и м.м. $X(\lambda)$, такие, что*

$$A(\lambda)=X(\lambda)CX(\lambda)^\nabla,\qquad \det X(\lambda)=\pm x_1(\lambda)\ldots x_n(\lambda). \tag{1}$$

Если $A(\lambda)$ — вещественный м.м., то для любой допустимой факторизации набора $i_1(\lambda),\ldots,i_m(\lambda)$, для которой $x_1(\lambda),\ldots,x_m(\lambda)$ — веществен-

ные многочлены, существуют вещественная матрица $C = C^*$ и вещественный м.м. $X(\lambda)$, удовлетворяющие соотношениям (1).

Теорема 3. *Пусть $A(\lambda) = A(\lambda)^\nabla$, $\det A(\lambda) \not\equiv 0$ и, если $\det A(\lambda)$ имеет нули на мнимой оси, что все инвариантные многочлены $A(\lambda)$ имеют на мнимой оси нули лишь четной кратности. Пусть $\det A(\lambda) = \pm\delta(\lambda)\delta(\lambda)^\nabla$, причем многочлены $\delta(\lambda)$ и $\delta(\lambda)^\nabla$ взаимно просты. Существует м.м. $X(\lambda)$ и постоянная матрица $C = C^*$ (вещественные при вещественном м.м. $A(\lambda)$) такие, что $A(\lambda) = X(\lambda)CX(\lambda)^\nabla$ и $\det X(\lambda) = \pm\delta(\lambda)$.*

Из теоремы 2 весьма просто выводится следующая известная (см., например, [1])

Теорема 4. *Пусть $A(\lambda) = A(\lambda)^\nabla$. Для существования м.м. $X(\lambda)$ такого, что $A(\lambda) = X(\lambda)X(\lambda)^\nabla$, необходимо и достаточно, чтобы при $\operatorname{Re}\lambda = 0$ эрмитова матрица $A(\lambda)$ была неотрицательной*. Если $A(\lambda)$ — вещественный м.м. и $A(\lambda) \geqslant 0$ при $\operatorname{Re}\lambda = 0$, то существует вещественный м.м. $X(\lambda)$, такой, что $A(\lambda) = X(\lambda)X(\lambda)^\nabla$. В обоих случаях многочлен $X(\lambda)$ может быть выбран так, что $\det X(\lambda)$ не имеет нулей в области $\operatorname{Re}\lambda > 0$.*

3°. Вспомогательные предложения. Ниже в доказательствах опускается аргумент у матричных многочленов. Через ст A обозначается степень м.м. $A = A(\lambda)$ при условии, что A — ненулевой м.м. Для нулевого м.м. $A(\lambda) \equiv 0$ полагаем ст $A(\lambda) = -1$. Через $\operatorname{diag}(a_1, \ldots, a_n)$ обозначается диагональная матрица с диагональными элементами $a_1, \ldots, a_n$. Ниже используются очевидные соотношения $(A+B)^\nabla = A^\nabla + B^\nabla$, $(AB)^\nabla = B^\nabla A^\nabla$, $\det(A^\nabla) = (\det A)^\nabla$.

Лемма 1. *Пусть $A(\lambda)^\nabla = A(\lambda)$ и $i_1(\lambda), \ldots, i_r(\lambda)$ — инвариантные многочлены матрицы $A(\lambda)$. Тогда $i_k(\lambda)^\nabla = \pm i_k(\lambda)$, $k = 1, \ldots, r$.*

Согласно [2] имеем $A = PDQ$, где P, Q — м.м., $\det P \equiv 1$, $\det Q \equiv \text{const} \neq 0$, $D(\lambda) = \operatorname{diag}(i_1, \ldots, i_r, 0, \ldots, 0)$. Так как $A = A^\nabla = Q^\nabla, D^\nabla, P^\nabla$, $\det P^\nabla = 1$, $\det Q^\nabla \equiv \text{const} \neq 0$, то инвариантные многочлены у м.м. A и D^∇ совпадают. Инвариантными многочленами (со старшим коэффициентом 1) м.м. D^∇ являются $\pm i_k^\nabla$. Считая (без ограничения общности), что i_k занумерованы так, что i_{k+1} делится на i_k $(k = 1, \ldots, r-1)$, получим $i_k = \pm i_k^\nabla$.

Лемма 2. Пусть $A(\lambda)^\nabla = A(\lambda) = \|\alpha_{jk}(\lambda)\|$, $\det A(\lambda) \equiv \text{const} \neq 0$ *и $\alpha_{11}(\lambda) \equiv 0$. Существует м.м. $Y(\lambda)$ (вещественный при вещественном м.м. $A(\lambda)$) такой, что $\det Y(\lambda) \equiv 1$ и такой, что у матрицы $B(\lambda) = Y(\lambda)A(\lambda)Y(\lambda)^\nabla$ элементами верхней строки (слева направо) и левого столбца (сверху вниз) являются числа* $1, 0, \ldots, 0$.

Выделяя верхнюю строку и левый столбец, представим матрицы A, A^{-1}, Y, $B = YAY^\nabla$ в виде

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

Здесь μ, l, n, N — многочлены, так как $\det A = \text{const} \neq 0$ и $A_1^\nabla = A_1$, $N^\nabla = N$. Многочлены ξ, x, y, Y_1 нужно подобрать так, чтобы было выполнено: (I) $\det Y = (\xi - x^\nabla Y_1^{-1} y)\det Y_1 \equiv 1$; (II) $b = ya^\nabla x + Y_1(a\xi^\nabla + A_1 x) \equiv 0$; (III) $\beta = \xi a^\nabla x + x^\nabla a \xi^\nabla - x^\nabla A_1 x \equiv 1$. Легко проверить, что эти соотношения удовлетворены для $x = n$, $Y_1 = I_{m-1}$, $\xi = (1 - n^\nabla A_1 n)/2$, $y = -a\xi - A_1 n$. При этом следует учесть соотношение $a^\nabla n = 1$, которое имеет место, поскольку $AA^{-1} = I_m$.

Важную роль в дальнейшем доказательстве играет лемма 3. Ниже приведено доказательство этой леммы для матриц порядка $m = 2$. Для матриц порядка $m > 2$ доказательство леммы 3 получено Б. Д. Любачевским **.

* Это означает, что $z^*A(\lambda)z \geqslant 0$ для $\operatorname{Re}\lambda = 0$ и для любого m-вектора z. Пусть G — матрица порядка $m \times m$. Ниже будем писать $G > 0$ или $G \geqslant 0$, если $G = G^*$ и, соответственно, $z^*Gz > 0$ или $z^*Gz \geqslant 0$ для $\forall\ z \neq 0$. Через I_m ниже обозначается единичная $m \times m$-матрица.

** Автор выражает признательность И. А. Лебедеву за полезные замечания по лемме 3.

Лемма 3. *Пусть* $A(\lambda)=A(\lambda)^\nabla=\|\alpha_{jh}\|$, $B(\lambda)=Y(\lambda)A(\lambda)Y(\lambda)^\nabla=\|\beta_{jh}\|$. *Либо* (а) *существует м.м.* $Y(\lambda)$, $\det Y(\lambda)\equiv\pm1$, *такой, что* $\beta_{11}=0$, *либо* (б) *существует м.м.* $Y(\lambda)$, $\det Y(\lambda)\equiv\pm1$, *такой, что* $\text{ст}\,\beta_{jh}<\text{ст}\,\beta_{jj}$ *при* $j\neq h$, $0\leqslant\text{ст}\,\beta_{11}\leqslant\text{ст}\,\beta_{22}\leqslant\ldots\leqslant\text{ст}\,\beta_{nn}$.

Пусть $m=2$, $Y=\|\eta_{jh}\|$. При $\eta_{11}=\eta_{22}=0$, $\eta_{12}=\eta_{21}=1$ (Y_1-преобразование) имеем $\beta_{11}=\alpha_{22}$, $\beta_{22}=\alpha_{11}$. При $\eta_{11}=\eta_{22}=1$, $\eta_{12}=0$, $\eta_{21}=-\eta$ (Y_2-преобразование) имеем $\beta_{21}=\alpha_{21}-\eta\alpha_{11}$, $\beta_{12}=\beta_{21}^\nabla$. Если $\text{ст}\,\alpha_{11}>\text{ст}\,\alpha_{22}$, то после Y_1-преобразования получим $\text{ст}\,\beta_{11}<\text{ст}\,\beta_{22}$. Если $\text{ст}\,\alpha_{11}\geqslant\text{ст}\,\alpha_{22}$, то полагаем $Y_1=I$. Если окажется, что $\text{ст}\,\beta_{21}\geqslant\text{ст}\,\beta_{11}$, то, деля β_{21} на β_{11}, получим $\beta_{21}=\eta\beta_{11}+\beta_{21}'$, где $\text{ст}\,\beta_{21}'<\text{ст}\,\beta_{21}$. Проводя Y_2-преобразование с этим η, получим м.м. $B'=\|\beta_{jh}'\|$, где $\beta_{11}'=\beta_{11}$, $\text{ст}\,\beta_{21}'<\text{ст}\,\beta_{11}'$. Возможны случаи: (I) $\text{ст}\,\beta_{22}'\geqslant\text{ст}\,\beta_{11}'$; (II) $\beta_{22}'=0$; (III) $\text{ст}\,\beta_{22}<\text{ст}\,\beta_{11}'$, $\beta_{22}'\neq0$. В случае (I) лемма доказана для $Y=Y_2Y_1$. В случае (II) лемма доказана для $Y=Y_1Y_2Y_1$. В случае (III), применяя Y_1-преобразование, придем к исходной ситуации, но степень элемента с индексами 11 будет меньше, чем вначале. Повторим все операции. Если $\text{ст}\,\alpha_{11}=k$, то процесс закончится через $k'\leqslant k$ повторений.

Лемма 4. *Пусть* $A(\lambda)=A(\lambda)^\nabla$, $\det A(\lambda)\equiv\text{const}\neq0$, $B(\lambda)=Y(\lambda)A(\lambda)Y(\lambda)^\nabla=\|\beta_{jh}\|$. *Существует м.м.* $Y(\lambda)$, $\det(\lambda)\equiv\pm1$ (*вещественный при вещественном* $A(\lambda)$) *такой, что* $\text{ст}\,\beta_{jh}<\text{ст}\,\beta_{jj}$ *при* $j\neq h$.

Применим лемму 3. В случае (а) утверждение доказано. В случае (б) по лемме 2 существует м.м Y_1 такой, что м.м. $B=Y_1AY_1$ имеет указанный в лемме 2 вид. Пусть B_1 — м.м., полученный из B отбрасыванием первой строки и первого столбца. Очевидно, $B_1^\nabla=B_1$ и $\det B_1=\text{const}\neq0$. Если для B_1 имеет место случай (б) леммы 3, то имеем $B_2=Y_2B_1Y_2^\nabla=\|\beta_{jh}\|$, $\text{ст}\,\beta_{jh}<\text{ст}\,\beta_{jj}$ при $j\neq h$, Y_2 — м.м. В этом случае утверждение доказано для $Y=Y_1\,\text{diag}(1,Y_2)$. Если для B_2 имеет место случай (а) леммы 3, то, применяя лемму 2, снова понизим порядок м.м. Через конечное число получим м.м. требуемого вида.

Лемма 5. *Если* $A(\lambda)=A(\lambda)^\nabla$, $\det A(\lambda)=\text{const}\neq0$, *то существуют постоянная матрица* $C=C^*$ *и м.м.* $U(\lambda)$ (*вещественная при вещественном м.м.* $A(\lambda)$) *такие, что* $A(\lambda)=U(\lambda)CU(\lambda)^\nabla$, $\det U(\lambda)\equiv\pm1$.

По лемме 4 $A=UBU^\nabla$, где $U=Y^{-1}$ — м.м., $\det U=\pm1$. Так как $\text{ст}\,\beta_{jh}<\text{ст}\,\beta_{jj}$ при $j\neq h$, то старший член многочлена $\det B$ содержится в слагаемом $\beta_{11}\ldots\beta_{nn}$. Поскольку $\det B=\det A=\text{const}$, то $\beta_{11}=\text{const},\ldots,\beta_{nn}=\text{const}$, а значит, $\beta_{jh}=0$ при $j\neq h$, $B=C=\text{const}$. Для $\text{Re}\,\lambda=0$ имеем $A=UCU^*$, $A=A^*$. Поэтому $C=C^*$.

4°. Доказательство теоремы 2. Согласно ([2]) имеем $A=PDQ$, где PQ — м.м., $D=\text{diag}(i_1,\ldots,i_m)$. По лемме 1 не лежащие на мнимой оси нули многочленов i_h расположены (с учетом их кратности) симметрично относительно мнимой оси. Поскольку чисто мнимые нули имеют четную кратность, то $i_h=\varepsilon_hx_hx_h^\nabla$, где x_h — многочлены, $\varepsilon_h=\pm1$. Пусть указанная факторизация допустима. Положим $D_0=\text{diag}(x_1,\ldots,x_m)$, $E=\text{diag}(\varepsilon_1,\ldots,\varepsilon_m)$. Тогда $D=D_0ED_0^\nabla$, $A=RKR^\nabla$, где $R=PD_0K=D_0^\nabla SD_0^{\nabla-1}$, $S=EQP^{\nabla-1}$. Очевидно, $S=\|s_{jh}\|$ — м.м. Покажем, что K — м.м. Так как $A^\nabla=A$ и $\det R\not\equiv0$, то $K=K^\nabla$. Пусть $K=\|k_{jh}\|$. Имеем $k_{jh}=x_j^\nabla s_{jh}x_h^{\nabla-1}$. Поскольку факторизация допустима, то k_{jh} — многочлены при $j<h$. Так как $K=K^\nabla$, $k_{hj}=k_{jh}^\nabla$, то k_{jh} — многочлены и при $j<h$. Так как $\det S=\text{const}\neq0$, то и $\det K=\text{const}\neq0$. По лемме 5 $K=UCU^\nabla$, где U — м.м. Поэтому $A=XCX^\nabla$, где $X=RU$ — м.м. Поскольку $\det U\equiv\pm1$, то $\det X=\pm\det R=\pm\det D_0=\pm x_1\ldots x_m$. Таким образом установлены соотношения (1). Если A — вещественный м.м., то P, Q, i_h и x_h (при указанной факторизации) — вещественные многочлены. Тогда R, S, K — вещественные м.м. По лемме 5 C и U вещественны, а значит, и X — вещественный м.м.

Доказательство теоремы 3. Предположим вначале, что $\delta(\lambda)$ не имеет корней в области $\text{Re}\,\lambda>0$. Пусть $i_h(\lambda)=\Pi(\lambda-\lambda_j)r_j$. Взяв в качестве $x_h(\lambda)$ произведение множителей $(\lambda-\lambda_j)^{r_j}$ с $\text{Re}\,\lambda_j<0$ и

$(\lambda - \lambda_j)^{r_j/2}$ с $\operatorname{Re}\lambda_j = 0$, получим, очевидно, допустимую факторизацию такую, что $\pm x_1, \ldots, x_m = \delta(\lambda)$. По доказанному выше имеем $\det X \equiv \pm\delta$. В общем случае доказательство аналогично — нужно в качестве $x_h(\lambda)$ взять общий наибольший делитель $i_h(\lambda)$ и $\delta(\lambda)$.

Доказательство теоремы 4. Необходимость очевидна.

Достаточность. Если $A > 0$ при $\operatorname{Re}\lambda = 0$, то, по теореме 2, $A = ZCZ^\nabla$, где Z — м.м. C, Z вещественны при вещественном м.м. A. Очевидно, $C = C^* > 0$. Полагая $X = ZC^{1/2}$, получим $A = XX^\nabla$. При этом $\text{ст}\, X = (\text{ст}\, A)/2$. Действительно, пусть $X(\lambda) = \lambda^M X_0 + \ldots + X_M$, $X_0 \neq 0$, $A(\lambda) = \lambda^N A_0 + \ldots + A_N$, $A_0 \neq 0$. Очевидно, $N \leqslant 2M$. При $N < 2M$ имеем $X_0X_0^* = 0$, $\operatorname{Sp} X_0X_0^* = |X_0|^2 = 0$, $X_0 = 0$. Поэтому $N = 2M$, что и утверждалось. Пусть $A(\lambda) \geqslant 0$ при $\operatorname{Re}\lambda = 0$. Используем прием из ([1]). По доказанному, для любого числа $\varepsilon > 0$ имеем $A(\lambda) + \varepsilon I_m = X_\varepsilon(\lambda) X_\varepsilon(\lambda)^\nabla$, где $X_\varepsilon(\lambda)$ — м.м., $\text{ст}\, X_\varepsilon(\lambda) = N/2$. При $\operatorname{Re}\lambda = 0$ имеем $\operatorname{Sp}(A + \varepsilon I) = |X_\varepsilon(\lambda)|^2$. Поэтому элементы $X_\varepsilon(\lambda)$ ограничены при $0 < \varepsilon < \varepsilon_0$, $\operatorname{Re}\lambda = 0$, $|\operatorname{Re}\lambda| \leqslant r$. Пусть $\varepsilon_n \to 0$. Выбирая из последовательности $X_{\varepsilon_n}(\lambda)$ подпоследовательность, сходящуюся к некоторому $X(\lambda)$ равномерно при $\operatorname{Re}\lambda = 0$, $|\operatorname{Re}\lambda| \leqslant r$, получим $A(\lambda) = X(\lambda)X(\lambda)^\nabla$. Поскольку степени $X_\varepsilon(\lambda)$ равномерно ограничены, то $X(\lambda)$ — м.м.

Ленинградский государственный университет
им. А. А. Жданова

Поступило
13 X 1970

ЦИТИРОВАННАЯ ЛИТЕРАТУРА

[1] V. M. Popov, Hiperstabilitatea sistemelor automate, România, 1966. [2] Ф. Р. Гантмахер, Теория матриц, М., гл. 6, 1954.