G. I. Malashonok, A computation of the characteristic polynomial of an endomorphism of a free module, *Zap. Nauchn. Sem. POMI*, 1999, Volume 258, 101–114

Записки научных
семинаров ПОМИ
Том 258, 1999 г.

G. I. Malaschonok

A COMPUTATION OF THE CHARACTERISTIC POLYNOMIAL OF AN ENDOMORPHISM OF A FREE MODULE

Abstract. Two methods are given for computation of the characteristic polynomial of an endomorphism of a free module over a commutative domain, that require $O(n^3)$ and $O(n^{\log 7})$ ring operations.

1. Introduction

Let $\mathcal{K}$ be a commutative ring, E a n-dimensional free module over $\mathcal{K}$. Let f be an endomorphism of E and A the matrix of f in some basis in E. The characteristic polynomial of the endomorphism f is $p_f(t) = \det(tI_n - A)$. One of the important problems of commutative algebra is to construct an effective method for computation of characteristic polynomials.

As $tI_n - A$ is the matrix over the polynomial ring $\mathcal{K}[t]$, it is possible to use one of the methods for matrix determinant computation over commutative ring [2–5]. However, complexity of the Dodgson's method is $O(n^3)$, and for recurcive method is $O(n^{\log 7})$ operations in the ring $\mathcal{K}[t]$. Therefore the complexity of these methods are $O(n^5)$ and $O(n^{2+\log_2 7})$ operations in the ring $\mathcal{K}$ itself. But it is known, that in the case when $\mathcal{K}$ is a field, there exist the methods with the complexity $O(n^3)$ operations in this field. Rapid algorithms for computation of characteristic polynomials are considered in [1] and [3].

In this paper we give two methods of the computation of the characteristic polynomial with the complexity $O(n^3)$ and $O(n^{\log 7})$ when $\mathcal{K}$ is an commutative domain.

The first method is based on the transferring to a such basis, in which the matrix of the endomorphism f is thridiagonal. It is allowable an arbitrary initial location of zero elements in the matrix.

The second method is based on the transferring to another basis in which the endomorphism matrix has the nonzero second diagonal, the first row and the last column.

In both of these bases the characteristic polynomials are easily computed in $O(n^2)$ operations in the ring $\mathcal{K}$.

1.1. The associated matrices.

Let us introduce the operation of association.

Let $K^L_{p,q}$ denote the subgroup of the group $GL(p+q,\mathcal{K})$, that formed by the matrices

$$\begin{pmatrix} aI_p & 0 \\ C & dI_q \end{pmatrix}, \quad a,d\in\mathcal{K}^*, C\in\mathcal{K}^{q\times p},$$

I_p – the identity matrix of order p, $\mathcal{K}^*=\mathcal{K}\backslash\{0\}$. The transposition transfers this subgroup to isomorphic subgroup of upper triangular matrices. This subgroup will be denoted by $K^R_{p,q}$. The involutory automorphism of $K^L_{p,q}$:

$$G=\begin{pmatrix} aI_p & 0 \\ C & dI_q \end{pmatrix} \to \begin{pmatrix} dI_p & 0 \\ -C & aI_q \end{pmatrix} = \widetilde{G},$$

will be called the *association*. This operation will be denoted by the sign "tilde".

The involutory automorphism of $K^R_{p,q}$:

$$\begin{pmatrix} aI_p & B \\ 0 & dI_q \end{pmatrix} \to \begin{pmatrix} dI_p & -B \\ 0 & aI_q \end{pmatrix}$$

is defined in such way that the operations of the transposition and the association are commutative

$$\widetilde{G^T}=(\widetilde{G})^T.$$

Obviously, $G\widetilde{G}=\widetilde{G}G=adI_{p+q}$ is a scalar matrix, $\widetilde{FG}=\widetilde{G}\widetilde{F}$, $\widetilde{\widetilde{G}}=G$, $\widetilde{I}=I$.

The involutory automorphism of the group $GL(2,\mathcal{K})$:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \to \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

will be also called the association. If G is the matrix of the second order, then its associated matrix is equal to the adjoint one: $\widetilde{G}=G^*$.

2. The thridiagonal method

2.1. The triangular Dodgson decomposition.

Let $A=(a_{ij})$ be the matrix of order n over the ring $\mathcal{K}$,

$$A^k_{i,j}=\begin{pmatrix} a_{1,1} & \dots & a_{1,k-1} & a_{1,j} \\ \dots & \dots & \dots & \dots \\ a_{k-1,1} & \dots & a_{k-1,k-1} & a_{k-1,j} \\ a_{i,1} & \dots & a_{i,k-1} & a_{i,j} \end{pmatrix}$$

is the submatrix of order k, $k = 2, \dots, n$, formed by bordering the submatrix of order $k-1$ standing in the upper left corner by row i and column j. We denote its determinant by a_{ij}^k,

$$a_{ij}^k = \det(A_{i,j}^k)$$

and $a_{ij}^1 = a_{ij}$. Let $\delta_k = a_{kk}^k$ be the corner minor of order k, I_k – the unit matrix of order k,

$$\mathbf{D}_k = \mathrm{diag}\,(I_{k+1}, \delta_k I_{n-k-1})$$

the diagonal matrix of order n. Let

$$A_u^{(k)} = \begin{pmatrix} A_1^k & A_2^k \\ A_3^k & A_4^k \end{pmatrix}.$$

$$(A_1^k, A_2^k) = (a_{i,j}^i), i = 1, \dots, k-1, \;\; j = 1, \dots, n,$$
$$(A_3^k, A_4^k) = (a_{i,j}^k), i = k, \dots, n, \quad j = 1, \dots, n,$$

be the matrix with blocks A_1^k of size $(k-1) \times (k-1)$ and A_4^k of size $(n-k+1) \times (n-k+1)$.

Let us note that $a_{i,j}^k = 0$ for $i < k$ and/or $j < k$ so A_1^k is the upper triangular matrix and $A_3^k = 0$. For $k = n$ we have $A_u^{(n)}$ – the upper triangular matrix. Let

$$v_k = (a_{k+1,k}^k, \dots, a_{n,k}^k)^T$$

be the column with $n-k$ elements,

$$L_k = \begin{pmatrix} \delta_k & 0 \\ v_k & I_{n-k} \end{pmatrix}, \quad \widetilde{L}_k = \begin{pmatrix} 1 & 0 \\ -v_k & \delta_k I_{n-k} \end{pmatrix}, \quad k = 1, 2, \dots, n, \quad (2.1)$$

are matrices of order $n-k+1$ and

$$\mathbf{L}_k = \mathrm{diag}\,(I_{k-1}, L_k) = \begin{pmatrix} 1 & \dots & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \delta_k & 0 & \dots & 0 \\ 0 & \dots & a_{k+1,k}^k & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & a_{n,k}^k & 0 & \dots & 1 \end{pmatrix},$$

$$\widetilde{\mathbf{L}}_k = \mathrm{diag}\,(I_{k-1}, \widetilde{L}_k) = \begin{pmatrix} 1 & \dots & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 1 & 0 & \dots & 0 \\ 0 & \dots & -a_{k+1,k}^k & \delta_k & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & -a_{n,k}^k & 0 & \dots & \delta_k \end{pmatrix},$$

$$\mathbf{L} = \mathbf{L}_1\mathbf{L}_2\cdots\mathbf{L}_{n-1},$$

$$\overline{\mathbf{L}} = \mathbf{D}_{n-2}^{-1}\widetilde{\mathbf{L}}_{n-1}\cdots\mathbf{D}_1^{-1}\widetilde{\mathbf{L}}_2\widetilde{\mathbf{L}}_1$$

are the lower triangular matrices of order n. We have

$$\widetilde{L}_k L_k = \delta_k I_{n-k+1}, \quad \overline{\mathbf{L}}\mathbf{L} = T$$

where

$$T = \operatorname{diag}(\delta_1, \delta_1\delta_2,\ \delta_2\delta_3, \dots, \delta_{n-2}\delta_{n-1},\ \delta_{n-1}).$$

Proposition 1. *If all the diagonal minors δ_k $(k = 1, 2, \dots, n-1)$ of matrix A are nonzero, then we have the identity:*

$$A_u^n = \overline{\mathbf{L}}A, \tag{2.2}$$

Multiplying by the matrices $\mathbf{D}_k^{-1}$ always provides the "exact" division. The proof is based on

Proposition 2. *For any k, $k = 2, \dots, n-1$, we have the identity*

$$\mathbf{D}_{k-1}A_u^{(k+1)} = \widetilde{\mathbf{L}}_k A_u^{(k)} \text{ and } A_u^{(2)} = \widetilde{\mathbf{L}}_1 A. \tag{2.3}$$

The proof of (2.3) is based on the Sylvester identity [2]

$$a_{k-1,k-1}^{k-1}a_{i,j}^{k+1} = a_{k,k}^k a_{i,j}^k - a_{i,k}^k a_{k,j}^k. \tag{2.4}$$

If all the diagonal minors δ_k are nonzero, then we have

$$A_u^{(k+1)} = \mathbf{D}_{k-1}^{-1}\widetilde{\mathbf{L}}_k A_u^{(k)} \text{ and}$$

$$A_u^n = \mathbf{D}_{n-2}^{-1}\widetilde{\mathbf{L}}_{n-1}(\mathbf{D}_{n-3}^{-1}\widetilde{\mathbf{L}}_{n-2}(\dots(\widetilde{\mathbf{L}}_1 A)\dots)) \tag{2.5}$$

We can compute step by step $A_u^{(k)}$, $k = 2, 3, \dots, n$, because all the elements of the matrices $\mathbf{D}_k$ and $\widetilde{\mathbf{L}}_k$ are the elements of the matrix $A_u^{(k)}$.

2.1.1. Decomposition with permutations.

Note that the condition to be nonzero for diagonal minors δ_k $(k = 1, 2, \dots, n-1)$ may be taken off if we use the permutation matrices in the factorization (2.2).

If the diagonal minor δ_k of order k equals zero and $v_k \neq 0$, then it is necessary to interchange rows i and k if $a_{i,k}^k \neq 0$.

So it is necessary to multiply $A_u^{(k)}$ by the permutation matrix $P_k = P_{(i,k)} = I_n + E_{ik} + E_{ki} - E_{kk} - E_{ii}$. Here E_{ik} denotes the matrix, all the elements of which are zeros, except of (i,k), that equals one.

If the diagonal minor δ_k of order k equals zero and $v_k = 0$, then $P_k = \widetilde{\mathbf{L}}_k = \mathbf{D}_{k-1} = I_n$, $\mathbf{D}_k = \mathbf{D}_{k-1}$. The formula (2.3) is true in the general case, but now $\widetilde{\mathbf{L}}_k = \operatorname{diag}(I_{k-1}, \widetilde{L}_k)P_k$.

2.1.2. Symmetric decomposition. There is also true the identity, symmetric to (2.2):

$$A_l^{(n)} = A\overline{\mathbf{U}}, \tag{2.6}$$

where A_l is the lower triangular matrix, $\overline{\mathbf{U}}$ is the upper triangular matrix.

$$\overline{\mathbf{U}} = \widetilde{\mathbf{U}}_1\widetilde{\mathbf{U}}_2\mathbf{C}_1^{-1}\cdots\widetilde{\mathbf{U}}_{n-1}\mathbf{C}_{n-2}^{-1}, \tag{2.7}$$

$\widetilde{\mathbf{U}}_k = Q_k\mathrm{diag}\,(I_{k-1}, \widetilde{U}_k)$, where Q_k is the permutation matrix, $C_k = \delta_k I_{n-k}$, $\mathbf{C}_k = \mathrm{diag}\,(I_k, C_k)$,

$$U_k = \begin{pmatrix} \delta_k & h_k \\ 0 & I_{n-k} \end{pmatrix}, \quad \widetilde{U}_k = \begin{pmatrix} 1 & -h_k \\ 0 & \delta_k I_{n-k} \end{pmatrix} \tag{2.8}$$

$h_k = (a^k_{k,k+1}, \ldots, a^k_{k,n})$, $U_k\widetilde{U}_k = C_k$.

2.2. Similar p-triangular matrix.

Further $L(A), T(A), U(A), \overline{L}(A), \overline{U}(A)$ denote $\mathbf{L}, \mathbf{T}, \mathbf{U}, \overline{\mathbf{L}}, \overline{\mathbf{U}}$ for the matrix A. If A is rectangular matrix, then the computations of the matrices $\mathbf{L}$, $\mathbf{T}$, $\mathbf{U}$, $\widetilde{\mathbf{L}}$, $\widetilde{\mathbf{U}}$ are fulfilled for the maximal left upper square block of A.

Let $A = \begin{pmatrix} \mathbf{a} & \mathbf{b} \\ \mathbf{c} & \mathbf{d} \end{pmatrix}$ be the matrix over R with blocks $\mathbf{a}$ of order p and $\mathbf{d}$ of order n. We shall call the matrix A the upper p-triangular, if the block $(\mathbf{c}, \mathbf{d})$ has the upper triangular form, the lower p-triangular, if the block $(\mathbf{b}, \mathbf{d})^T$ has the lower triangular form.

We denote by the calligraphic letters the block-diagonal matrices of order $(n+p)$ of the form $\mathrm{diag}\,(I_p, G) = \mathcal{G}$, where G is the matrix of order n.

Let G be a some $n \times n$ matrix,

$$\widetilde{\mathbf{L}} = \widetilde{L}((\mathbf{c}, \mathbf{d}G)),$$

$$\mathbf{L} = L((\mathbf{c}, \mathbf{d}G)),$$

$$\mathbf{T} = T((\mathbf{c}, \mathbf{d}G)).$$

If we put $G = \mathbf{L}$, $\overline{\mathcal{L}} = \mathrm{diag}\,(I_p, \overline{\mathbf{L}})$, $\mathcal{L} = \mathrm{diag}\,(I_p, \mathbf{L})$, and

$$A_u = \overline{\mathcal{L}}A\mathcal{L}$$

then

$$A_u = \begin{pmatrix} I_p & 0 \\ 0 & \overline{\mathbf{L}} \end{pmatrix}\begin{pmatrix} \mathbf{a} & \mathbf{b} \\ \mathbf{c} & \mathbf{d} \end{pmatrix}\begin{pmatrix} I_p & 0 \\ 0 & \mathbf{L} \end{pmatrix} = \begin{pmatrix} \mathbf{a} & \mathbf{bL} \\ \overline{\mathbf{L}}\mathbf{c} & \overline{\mathbf{L}}\mathbf{dL} \end{pmatrix}$$

is similar to the matrix $\mathrm{diag}(I_p, \mathbf{T})A$, because $\overline{\mathbf{L}}\mathbf{L} = \mathbf{T}$.

Let us show that the matrix A_u is the upper p-triangular matrix.

Consider the multiplying of matrix A of order $p+n$ from the right by the matrix $\mathcal{L}_k = \mathrm{diag}(I_p, L_k)$. It does not change the first $p+k-1$ columns, the last $n-k$ columns are only multiplied by the number δ_k. The $p+k$ column (the column with the number $p+k$) is changed by the linear combination of the columns $p+k, p+k+1, \dots, p+n$.

The multiplying from the left by the matrix $\widetilde{\mathcal{L}}_k$ does not change the first $p+k$ rows. Each row with the number $i, (i = p+k+1, \dots, n)$ is changed by the linear combination of the rows i and $p+k$ and the (i,k) element becomes zero. So all the elements in the k column, which are bellow the $p+k$ row (below the lower p-diagonal – $\{(p,1), (p+1,2), \dots, (n, n-p+1)\}$), are zero.

The multiplying by the diagonal matrix do not change the form of the matrix. Therefore, after the sequential multiplyings by $\mathcal{L}_1$ and $\widetilde{\mathcal{L}}_1$ the elements of the first column (below the lower p-diagonal) vanish, by $\mathcal{L}_2$, $\widetilde{\mathcal{L}}_2$ and $\mathbf{D}_1^{-1}$ – the second, and so on. As the result the matrix would be the upper p-triangular.

If the original matrix already the lower p-triangular (all the (i,j) elements, $i < p, j > p+1$), are zero), then the multiplyings by $\mathcal{L}_k$ and $\widetilde{\mathcal{L}}_k$ do not change this form. So the result matrix would be the p-diagonal.

The factors of the matrices $\mathbf{L}$ and $\overline{\mathbf{L}}$ may be computed sequentially.

We have $(\mathbf{c}, \mathbf{d}G) = (\mathbf{c}, \mathbf{dL})$. The first p columns of the matrix $(\mathbf{c}, \mathbf{dL})$ form the block $\mathbf{c}$, and do not depend on $\mathbf{L}$. So using these p columns we may compute consequently the first p factors of the matrix $\widetilde{\mathbf{L}}$: $\widetilde{\mathbf{L}}_1, \mathbf{D}_1, \widetilde{\mathbf{L}}_2, \dots, \mathbf{D}_{p-1}, \widetilde{\mathbf{L}}_p$. Then we may write the first p factors of the matrix $\mathbf{L}$, compute p columns of the matrix $\mathbf{dL}$, and then the following p factors of $\overline{\mathbf{L}}$, etc.

An example of the upper 1-triangle matrix computation see in Appendix A.

Analogously the lower p-triangle matrix

$$A_l = \mathcal{U} A \overline{\mathcal{U}},$$

is computed. Here we denote $\overline{\mathcal{U}} = \mathrm{diag}(I_p, \overline{\mathbf{U}})$, $\mathcal{U} = \mathrm{diag}(I_p, \mathbf{U})$, $\overline{\mathbf{U}} = \overline{U}((\mathbf{b}, H\mathbf{d})^T)$, $\mathbf{T} = T((\mathbf{b}, H\mathbf{d})^T)$, $H = \mathbf{U}$. The matrix A_l is similar to $\mathcal{T}A$ where $\mathcal{T} = \mathrm{diag}(I_p, \mathbf{T})$.

2.3. Computation of $2p+1$ diagonal matrix.

To compute the $2p+1$ diagonal matrix D $(p \geq 1)$, we can sequentially compute the upper and the lower p-triangular matrices:

$$A_u = \overline{\mathcal{L}} A \mathcal{L}$$

$$D = \mathcal{U} A_u \overline{\mathcal{U}}$$

Here $\overline{\mathbf{L}} = \overline{L}((\mathbf{c}, \mathbf{d}G))$, $\mathbf{T}_l = T((\mathbf{c}, \mathbf{d}G))$, $G = \mathbf{L}$, $\overline{\mathbf{U}} = \overline{U}(\mathcal{H}\hat{A}_u)$, $\mathbf{T}_u = T((\mathcal{H}\hat{A}_u))$, $H = \mathbf{U}$, and $\hat{A}_u$ is the right $(n+p) \times n$ block of the matrix A_u.

The matrix D is $2p+1$ diagonal but the matrix D does not similar to $\mathcal{T}A$, where $\mathcal{T} = \operatorname{diag}(I_p, \mathbf{T}_u\mathbf{T}_l)$ is a diagonal matrix.

The necessary condition is the absence of any permutation in the computation of lower p-triangular matrix: the permutations may lead to the appearance of nonzero elements under the lower p-diagonal.

It is possible to take off this limitation and in the same time to comput the matrix D that is similar to A if the computation of the upper and the lower p-triangular matrices are fulfilled simultaneously:

$$A^{(2)} = \underset{5}{\mathbf{U}_1} \; \underset{4}{\widetilde{\mathbf{L}}_1} \; \underset{1}{P_1} \; \underset{0}{A} \; \underset{2}{P_1^{-1}} \; \underset{3}{\mathbf{L}_1} \; \underset{6}{\widetilde{\mathbf{U}}_1}$$

$$A^{(k+1)} = \underset{6}{\mathbf{U}_k} \; \underset{5}{\mathbf{D}_{k-1}^{-1}} \; \underset{4}{\widetilde{\mathbf{L}}_k} \; \underset{1}{P_k} \; \underset{0}{A^{(k)}} \; \underset{2}{P_k^{-1}} \; \underset{3}{\mathbf{L}_k} \; \underset{7}{\widetilde{\mathbf{U}}_k} \; \underset{8}{\mathbf{C}_{k-1}^{-1}},$$

$$k \geq 2,$$

$$D = A^{(n)}.$$

where $\mathbf{D}_k$ and $\mathbf{C}_k$ are the matrices of the next kind $\operatorname{diag}(I_{k-1}, \delta_k I_{n-k+1})$ and only lasts $n-k-1$ diagonal elements of the matrices $\mathbf{D}_k$ and $\mathbf{C}_k$ provides the "exact" division.

The numbers indicate the order of actions. The permutations matrix P_k is chosen in such way that the pivot element $a^k_{k,k+p}$ in the upper p-diagonal and the pivot element $a^k_{k+p,k}$ in the lower p-diagonal are nonzero.

The following cases are possible:

1. $a^k_{k,k+p} \neq 0, a^k_{k+p,k} \neq 0$, then $P_k = I_n$;
2. $a^k_{k,k+r} \neq 0, a^k_{k+r,k} \neq 0$, for some $r, (r > p)$, then $P_k = P_{(k,k+r)}$;
3. $a^k_{k,k+p} \neq 0, a^k_{k+p,k} = 0, a^k_{k+p,r} \neq 0, a^k_{r,k+p} = 0$, for some r, s $(n \geq r > k)$, then $P_k = I_n + E_{k,r}$;

4. $a^k_{s,k+p} \neq 0, a^k_{k+p,s} = 0, a^k_{k+p,r} \neq 0, a^k_{r,k+p} = 0$, for some r, s $(n \geq r, s > k)$, then $P_k = I_n + E_{k,r} + E_{k,s}$;
5. $a^k_{k,k+p} = 0, h_k = 0$, i.e. zero row k, then $\widetilde{\mathbf{U}}_k = \mathbf{U}_k = \mathbf{C}_k = I_n$;
6. $a^k_{k+p,k} = 0, v_k = 0$, i.e. zero column k, then $\widetilde{\mathbf{L}}_k = \mathbf{L}_k = \mathbf{D}_k = I_n$;

If we have z zero rows $r - z, r - z + 1, \dots, r - 1$ and nonzero row r than $\mathbf{C}_r = \mathrm{diag}\,(I_{r-z}, \delta_{r-z} I_{n-r+z})$.

If we have z zero columns $c - z, c - z + 1, \dots, c - 1$ and nonzero row c than $\mathbf{D}_c = \mathrm{diag}\,(I_{c-z}, \delta_{c-z} I_{n-c+z})$.

An example of the tridiagonal matrix computation see in Appendix B.

2.4. Computation of characteristic polynomial for the thridiagonal matrix.

The computation of characteristic polynomial may be fulfilled in two steps: to compute the thridiagonal matrix D for $p = 1$ and then to compute the determinant of the matrix $\lambda I_n - \mathcal{T}^{-1} D = \mathcal{T}^{-1}(\lambda \mathcal{T} - D)$.

For the thridiagonal matrix it is easy to receive the recurrent relation for the computation of the determinant with the help of the Silvester identity (2.4):

$$a^{k+1}_{k+1,k+1} = a^k_{k,k} a^k_{k+1,k+1} - a^{k-1}_{k-1,k-1} a_{k,k+1} a_{k+1,k}, \quad k = 2, \dots, n-1, \qquad (2.11)$$

as for $(i,j) = (k, k+1), (k+1, k), (k+1, k+1)$ we have $a^k_{i,j} = a_{i,j} a^{k-1}_{k-1,k-1}$.

The recurrent formula (2.11) for the matrix $\lambda\mathcal{T} - D$ with the elements $a_{i,j}(\lambda)$ for $\mathcal{T} = \mathrm{diag}\,(d_{1,1}, \cdots, d_{n,n}), D = (a_{i,j})$ is:

$$a^1_{1,1}(\lambda) = \lambda d_{1,1} - a_{1,1},$$

$$a^2_{2,2}(\lambda) = a^1_{1,1}(\lambda)(\lambda d_{2,2} - a_{2,2}) - a_{1,2} a_{2,1}, \qquad (2.12)$$

$$a^{k+1}_{k+1,k+1}(\lambda) = a^k_{k,k}(\lambda)(\lambda d_{k+1,k+1} - a_{k+1,k+1}) - a^{k-1}_{k-1,k-1}(\lambda) a_{k,k+1} a_{k+1,k}. \qquad (2.13)$$

3. The dichotomous method

The other method of computation of the characteristic polynomial is based on the recurrent method from [5]. This method allows to compute the factorization of the matrix A^* adjoint for the matrix A. We give the description of this factorization.

3.1. Factorization of the adjoint matrix.

For the matrix A of size $n \times m$, $n < m$ we denote by A_R the right square $n \times n$ block, A_L - the left $n \times (m-n)$ block. Denote by B^* the matrix adjoint for the matrix B.

Let A be the matrix of size $n \times m$, $n \leq m$. For the integer s, $2 \leq s \leq n$,

$$s = \sum_{i=1}^{r} 2^{p_i}, \quad 0 < p_1 < \ldots < p_r$$

we denote by

$$\begin{aligned} s^o &= p_1, \\ \widehat{s} &= 2^{p_1}, \\ \overline{s} &= s - \widehat{s}, \\ s_t &= s - 2^t. \end{aligned}$$

The matrix $A_s = (a_{ij}^{\overline{s}+1})$ have the size $\hat{s} \times (m - s + \hat{s})$, $i = s+1, \ldots, s+\hat{s}$, $j = s - \hat{s} + 1, \ldots, m$. The element a_{ij}^k is the bordering minor of matrix A, as in \$2.1.

The matrix $G_s = (\delta_{ij}^s)$ have the size $\hat{s} \times (m-s)$, $i = s - \hat{s} + 1, \ldots, s$, $j = s+1, \ldots, m$. $\delta_k = a_{kk}^k$, $\delta_0 = 1$, δ_{ij}^s is the minor, received by the change of i column by j column in the minor δ_s.

Let

$$F_2 = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad \widetilde{F}_2 = \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix},$$

$$L^s = \begin{pmatrix} \delta_{s_1} I_{\widehat{s}_1} & 0 \\ A_{s_1,L} & I_{\widehat{s}_1} \end{pmatrix}, \quad \widetilde{L}^s = \begin{pmatrix} I_{\widehat{s}_1} & 0 \\ -A_{s_1,L} & \delta_{s_1} I_{\widehat{s}_1} \end{pmatrix},$$

$$M^s = \begin{pmatrix} a_{s-1,s-1}^{s-1} & a_{s-1,s}^{s-1} \\ a_{s,s-1}^{s-1} & a_{s,s}^{s-1} \end{pmatrix}, \quad \widetilde{M}^s = \begin{pmatrix} a_{s,s}^{s-1} & -a_{s-1,s}^{s-1} \\ -a_{s,s-1}^{s-1} & a_{s-1,s-1}^{s-1} \end{pmatrix},$$

$$U_t^s = \begin{pmatrix} I_{2^t} & G_{s_t,L} \\ 0 & \delta_s I_{2^t} \end{pmatrix}, \quad \widetilde{U}_t^s = \begin{pmatrix} \delta_s I_{2^t} & -G_{s_t,L} \\ 0 & I_{2^t} \end{pmatrix},$$

$t = 1, 2, \ldots, s^o - 1$.

Denote

$$\mathbf{L}^s = \operatorname{diag}(I_{\overline{s}_1}, L^s, I_{n-s-\widehat{s}_1}), \widetilde{\mathbf{L}}^s = \operatorname{diag}(I_{\overline{s}_1}, \widetilde{L}^s, I_{n-s-\widehat{s}_1}),$$

$$\mathbf{M}^s = \operatorname{diag}(I_{s_1}, M^s, I_{n-s}), \widetilde{\mathbf{M}}^s = \operatorname{diag}(I_{s_1}, \widetilde{M}^s, I_{n-s}),$$

$$\mathbf{D}_l^s = \operatorname{diag}(I_{s_1}, \delta_{\overline{s}_1}^{-1} I_{\widehat{s}_1}, I_{n-s_1-\widehat{s}_1}), \mathbf{D}_m^s = \operatorname{diag}(I_{s_1}, \delta_{s_1}^{-1} I_2, I_{n-s}),$$

$$\mathbf{U}_t^s = \operatorname{diag}(I_{s_t-2^t}, U_t^s, I_{n-s}), \widetilde{\mathbf{U}}_t^s = \operatorname{diag}(I_{s_t-2^t}, \widetilde{U}_t^s, I_{n-s}),$$

$$\mathbf{D}_t^s = \operatorname{diag}(I_{s_t-2^t}, \delta_{s_t}^{-1} I_{2^t}, I_{n-s_t}), \quad t = 1, 2, \ldots, s^o - 1.$$

Theorem 1. On the dichotomous factorization

Let A be the matrix of order $n = 2^r$. If all its diagonal minors δ_k $(k = 1, 2, \ldots, n-1)$ are nonzero, then for the adjoint matrix A^* there exists the factorization

$$A^* = \overline{\mathbf{F}}_n \cdots \overline{\mathbf{F}}_4 \overline{\mathbf{F}}_2,$$

$$\overline{\mathbf{F}}_2 = \operatorname{diag}(\widetilde{F}_2, I_{n-2}), \overline{\mathbf{F}}_s = \mathbf{D}^s_{s^o-1} \widetilde{\mathbf{U}}^s_{s^o-1} \cdots \mathbf{D}^s_1 \widetilde{\mathbf{U}}^s_1 \mathbf{D}^s_m \widetilde{\mathbf{M}}^s \mathbf{D}^s_l \widetilde{\mathbf{L}}^s,$$

$s = 2, 4, \ldots, n.$

The proof of the theorem is based on the determinant identity from the recursive method [5].

Further $\overline{\mathbf{F}}$ denote $\overline{F}(A)$ for the matrix A. If A is not the square matrix, then the factorization is fulfilled for the maximal left upper square block of A.

To reduce the matrix A of size $n \times m$ to the diagonal form we must multiply it by the adjoint matrix A^*_L: $A^*_L A = (\delta_n I_n, G_n)$. We may use the dichotomous factorization of the matrix A^*_L. The complexity of such computations ([5]) has the same order as the matrix multiplications, that participate in them.

If the multiplication of the matrices is fulfilled by the Strassen method [6], then the complexity of such computations is $O(n^{\log 7})$.

3.2. Computation of the surrounded diagonal matrix.

Let A be the matrix of order $n+1$ and A' – its submatrix, received by the crossing out the first row, $E = \operatorname{diag}(1, \mathbf{E})$ – a matrix of order $n+1$. Let $\overline{\mathbf{F}} = \overline{F}(A'E)$ be the dichotomous factorization of the left block of the matrix $A'E$. If $\mathbf{E} = \mathbf{F}_2 \mathbf{F}_4 \ldots \mathbf{F}_n$, $\mathbf{F}_2 = \operatorname{diag}(F_2, I_{n-2})$, $\mathbf{F}_s = \mathbf{L}^s \mathbf{M}^s \mathbf{U}^s_1 \ldots \mathbf{U}^s_{s^o-1}$, then $\widetilde{\overline{\mathbf{F}}} \mathbf{E} = \mathbf{T}$, where $\mathbf{T}$ is the diagonal matrix.

Moreover, if $\overline{F} = \operatorname{diag}(1, \overline{\mathbf{F}})$, then $A_b = \overline{F} A E$ is the matrix that has the nonzero elements only in the second diagonal (the diagonal of the block A'), in the first row and in the last column. The matrix A_b is similar to $\mathbf{T}A$. The computations are fulfilled according to the scheme, analogous of the scheme in **2.3**.

3.3. Computation of the characteristic polynomial. For the surrounded diagonal matrix D' of order n with nonzero elements in the first row $d_{1,i}$, last column $d_{i,n}$ and in the second diagonal $d_i = d_{i+1,i}$, the characteristic polynomial $\det(\lambda I_n - D')$ is easily computed:

$$\det(\lambda I_n - D') = (-1)^{n+1} d_{1,n} d_1 \cdots d_{n-1} + (-1)^{n+2} d_{2,n} \Delta_1 d_2 \cdots d_{n-1} + \ldots$$

$$+\dots(-1)^{2n}d_{n,n}\Delta_{n-1},$$

where

$$\Delta_1=\lambda-d_{1,1},$$
$$\Delta_2=\Delta_1\lambda-d_1d_{1,2},$$
$$\Delta_{k+1}=\Delta_k\lambda+(-1)^k d_1\dots d_k d_{1,k+1},$$
$$k\geq 2.$$

4. Appendix A

Example for the computation of a similar p-triangular matrix

Let us have a matrix

$$A=\begin{pmatrix}0&2&0&1&1\\2&1&1&0&2\\0&3&1&0&0\\1&0&0&1&0\\2&1&0&0&2\end{pmatrix},$$

then (see the section 2.2) we obtain:

$$\widetilde{\mathcal{L}}_1=\begin{pmatrix}1&0&0&0&0\\0&1&0&0&0\\0&0&2&0&0\\0&-1&0&2&0\\0&-2&0&0&2\end{pmatrix},\ \mathcal{L}_1=\begin{pmatrix}1&0&0&0&0\\0&2&0&0&0\\0&0&1&0&0\\0&1&0&1&0\\0&2&0&0&1\end{pmatrix},$$

$$A^{(2)}=\widetilde{\mathcal{L}}_1A\mathcal{L}_1=\begin{pmatrix}0&7&0&1&1\\2&6&1&0&2\\0&12&2&0&0\\0&-4&-1&2&-2\\0&0&-2&0&0\end{pmatrix},$$

$$\widetilde{\mathcal{L}}_2=\begin{pmatrix}1&0&0&0&0\\0&1&0&0&0\\0&0&1&0&0\\0&0&4&12&0\\0&0&0&0&12\end{pmatrix},\ \mathcal{L}_2=\begin{pmatrix}1&0&0&0&0\\0&1&0&0&0\\0&0&12&0&0\\0&0&-4&1&0\\0&0&0&0&1\end{pmatrix},$$

$$A^{(3)} = \operatorname{diag}^{-1}(1,1,1,2,2)\widetilde{\mathcal{L}}_2 A^{(2)}\mathcal{L}_2 = \begin{pmatrix} 0 & 7 & -4 & 1 & 1 \\ 2 & 6 & 12 & 0 & 2 \\ 0 & 12 & 24 & 0 & 0 \\ 0 & 0 & -72 & 12 & -12 \\ 0 & 0 & -144 & 0 & 0 \end{pmatrix},$$

$$\widetilde{\mathcal{L}}_3 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 144 & -72 \end{pmatrix}, \quad \mathcal{L}_3 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -72 & 0 \\ 0 & 0 & 0 & -144 & 1 \end{pmatrix},$$

$$A^{(4)} = \operatorname{diag}^{-1}(1,1,1,1,12)\widetilde{\mathcal{L}}_2 A^{(2)}\mathcal{L}_2 = \begin{pmatrix} 0 & 7 & -4 & -216 & 1 \\ 2 & 6 & 12 & -288 & 2 \\ 0 & 12 & 24 & 0 & 0 \\ 0 & 0 & -72 & 864 & -12 \\ 0 & 0 & 0 & 10368 & -144 \end{pmatrix}.$$

The 1-triangular matrix $A^{(4)}$ is similar to $\operatorname{diag}(1,2,24,-864,-72)A$.

The characteristic polynomial of the matrix A and the matrix

$$\operatorname{diag}(1, \frac{1}{2}, \frac{1}{24}, \frac{-1}{864}, \frac{-1}{72})A^{(4)} =$$

$$= \operatorname{diag}(1, 1, \frac{1}{2}, \frac{-1}{72}, -1)\begin{pmatrix} 0 & 7 & -4 & -216 & 1 \\ 1 & 3 & 6 & -144 & 1 \\ 0 & 1 & 2 & 0 & 0 \\ 0 & 0 & -6 & 72 & -1 \\ 0 & 0 & 0 & 144 & -2 \end{pmatrix}$$

is

$$-x^5 + x^4 + 3x^3 - 22x^2 + 6x + 12. \qquad (A.1)$$

5. Appendix B

Example for the computation of a similar tridiagonal matrix

Let we have the same initial matrix as in Appendix A.

$$\begin{pmatrix} 0 & 2 & 0 & 1 & 1 \\ 2 & 1 & 1 & 0 & 2 \\ 0 & 3 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 2 & 1 & 0 & 0 & 2 \end{pmatrix} \rightarrow_{1(L)} \begin{pmatrix} 0 & 7 & 0 & 1 & 1 \\ 2 & 6 & 1 & 0 & 2 \\ 0 & 12 & 2 & 0 & 0 \\ 0 & -4 & -1 & 2 & -2 \\ 0 & 0 & -2 & 0 & 0 \end{pmatrix} \rightarrow_{1(U)}$$

$$\begin{pmatrix} 0 & 7 & 0 & 0 & 0 \\ 14 & 38 & 28 & -24 & 46 \\ 0 & 12 & 14 & -12 & -12 \\ 0 & -4 & -7 & 18 & -10 \\ 0 & 0 & -14 & 0 & 0 \end{pmatrix} \to_{2(L)}$$

$$\begin{pmatrix} 0 & 7 & 0 & 0 & 0 \\ 7 & 19 & 216 & -12 & 23 \\ 0 & 6 & 108 & -6 & -6 \\ 0 & 0 & -504 & 84 & -84 \\ 0 & 0 & -1008 & 0 & 0 \end{pmatrix} \to_{2(U)}$$

$$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 7 & \frac{19}{7} & \frac{216}{7} & 0 & 0 \\ 0 & \frac{1296}{7} & \frac{6192}{7} & -60480 & -29232 \\ 0 & 0 & -72 & 1728 & -936 \\ 0 & 0 & -144 & -1728 & 3312 \end{pmatrix} \to_{3(L)}$$

$$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 7 & \frac{19}{7} & \frac{216}{7} & 0 & 0 \\ 0 & \frac{108}{7} & \frac{516}{7} & 713664 & -2436 \\ 0 & 0 & -6 & 864 & -78 \\ 0 & 0 & 0 & 2239488 & -31104 \end{pmatrix} \to_{3(U)}$$

$$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 7 & \frac{19}{7} & \frac{1}{7} & 0 & 0 \\ 0 & \frac{108}{7} & \frac{43}{126} & 3304 & 0 \\ 0 & 0 & -19824 & -22401792 & 11851370496 \\ 0 & 0 & 0 & 10368 & -77511168 \end{pmatrix} \to_{4(D)}$$

$$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 7 & \frac{19}{7} & \frac{1}{7} & 0 & 0 \\ 0 & \frac{108}{7} & \frac{43}{126} & \frac{1}{216} & 0 \\ 0 & 0 & \frac{826}{3} & \frac{463}{1062} & -\frac{13608}{59} \\ 0 & 0 & 0 & -\frac{1}{4956} & \frac{89}{59} \end{pmatrix}$$

This tridiagonal matrix is similar to initial matrix A and has the same characteristic polynomial (A.1).

References

1. J. Abdeljaoued, *Algorithmes rapides pour le calcul du polynôme caractéristique.* Thèse de l'Université de Franche-Comté, Mars 1997.

2. C. L. Dodgson, *Condensation of determinants, being a new and brief method for computing their arithmetic values.* Proc. Royal Soc. Lond. **A.15** (1866), 150–155.
3. D. K. Faddeev and V. N. Faddeeva, *Computational methods of linear algebra.* W. H. Freeman & Co., San Francisco, 1963.
4. G. I. Malaschonok, *Algorithms for the solution of systems of linear equations in commutative rings.* Effective Methods in Algebraic Geometry, ed. by T. Mora and C. Traverso, Progress in Mathematics 94, Birkhauser, 1991, 289–298.
5. G. I. Malaschonok, *Recursive Method for the Solution of systems of Linear Equations.* Computational Mathematics (A. Sydow Ed, Proceedings of the 15th IMACS World Congress, Vol. I, Berlin, August 1997), Wissenschaft & Technik Verlag, Berlin 1997, 475–480.
6. V. Strassen, *Gaussian Elimination is not optimal.* Numerische Mathematik, **13** (1969), 354–356.

Tambov State University,
392622 Tambov, Russia
e-mail: malaschonok@main.tsu.tambov.ru

Поступило 15 мая 1999 г.