An algorithm for the calculation of the minimal polynomial

S. BIAŁAS 1* and M. BIAŁAS 2

1 The School of Banking and Management, 4 Armii Krajowej St., 30-150 Kraków, Poland
2 Faculty of Management, AGH University of Science and Technology, 30 Mickiewicza St., 30-059 Kraków, Poland

Abstract. This paper gives the simple algorithm for calculation of the degree and coefficients of the minimal polynomial for the complex matrix \( A = [a_{ij}]_{n \times n} \).

Key words: matrix, minimal polynomial, characteristic polynomial.

1. Introduction

We use the standard notation. We denote by \( M_{m,n} \) the set of \( m \times n \) real or complex matrices. In case \( n = m \) we will write \( M_n \) instead of \( M_{n,n} \).

A complex polynomial \( f(\lambda) \) is called an annihilating polynomial for a matrix \( A \in M_n \) if \( f(\lambda) \neq 0 \) and \( f(A) = 0 \). The complex polynomial \( \psi(\lambda) \) of least degree for which \( \psi(A) = 0 \) is called the minimal polynomial of the matrix \( A \in M_n \).

The properties and the applications of the minimal polynomials in the control theory have been presented in [1, 2].

In this paper the simple algorithm is given for calculation of the degree and coefficients of the minimal polynomial.

For the matrix \( A = [a_{ij}] \in M_n \) we will use the following notations:

- \( \varphi(\lambda) = \det(\lambda I - A) \) – characteristic polynomial of the matrix \( A \),
- \( \psi(\lambda) \) – minimal polynomial of the matrix \( A \),
- \( \text{vec} A = [a_{11} \ a_{12} \ a_{13} \ ... \ a_{n1} \ a_{21} \ a_{22} \ ... \ a_{n2} \ ... \ a_{m1} \ a_{m2} \ ... \ a_{mn}]^T \),
- \( A^0 = I \in M_n \),
- \( A^k = A^{k-1}A \ (k = 1, 2, \ldots) \),
- \( A^{(k)} = a_{ij}^{(k)} \ (k = 0, 1, 2, \ldots) \),
- \( B_k = [a^{(0)} a^{(1)} \ ... \ a^{(k)}] \ (k = 0, 1, 2, \ldots) \) where \( a^{(k)} = k + 1 \)-th column of the matrix \( B_k \in M_{n^2,k+1} \),
- \( \text{rank} (B) \) – rank of the matrix \( B \),
- \( \text{deg} f(x) \) – degree of the polynomial \( f(\lambda) \),
- \( \text{unit matrix} \),
- \( \varnothing \) – empty set.

Example 1. For the matrix \( A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \) we have:

\[
\begin{align*}
\varphi(\lambda) &= \det(\lambda I - A) = \lambda^2 - (a_{11} + a_{22})\lambda + (a_{11}a_{22} - a_{12}a_{21}), \\
\psi(\lambda) &= \lambda - (a_{11} + a_{22}) = 0 , \\
A^{(0)} &= \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \\
A^{(1)} &= \begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix}, \\
B_1 &= \begin{bmatrix} 1 & a_{11} \\ 0 & a_{21} \end{bmatrix}.
\end{align*}
\]

2. An algorithm for the calculation of the degree and the coefficients of the minimal polynomial

For the matrix \( A = [a_{ij}] \in M_n \) we will prove the following Lemma.

Lemma 1. If the matrix \( A = [a_{ij}] \in M_n \), the matrix \( B_k \) is defined by (1), then

\( K = \{k \in N : \text{rank} B_k = \text{rank} B_{k-1} \neq 0 \ \text{and} \ n \in K \} \).

Proof. We see that if

\[
\varphi(\lambda) = \det(\lambda I - A) = \lambda^n + b_{n-1}\lambda^{n-1} + \cdots + b_1\lambda + b_0,
\]

then

\[
A^n + b_{n-1}A^{n-1} + \cdots + b_1A + b_0I = 0 \in M_n,
\]

\[
a^{(n)} = -[b_{n-1}a^{(n-1)} + \cdots + b_1a^{(1)} + b_0a^{(0)}],
\]

\[
\text{rank} B_n = \text{rank} [a^{(0)} a^{(1)} \ ... \ a^{(n)}] = \text{rank} [a^{(0)} a^{(1)} \ ... \ a^{(n-1)} 0] = \text{rank} B_{n-1},
\]

where \( 0 = [0 0 \ldots 0]^T \in M_{n^2,1} \). Therefore \( n \in K \) and \( K \neq \emptyset \).

Definition 1. A number \( k_0 = \min K \) is called the associated rank of the matrix \( A = [a_{ij}] \in M_n \).

Theorem 1. If \( k_0 \) is the associated rank of the matrix \( A = [a_{ij}] \in M_n \) and \( \psi(\lambda) \) is the minimal polynomial of this matrix then:
1) \( \text{rank } B_k = k + 1 \) \( (k = 0, 1, \ldots, k_0 - 1) \),
2) \( \text{rank } B_k = k_0 \) \( (k \geq k_0) \),
3) \( \deg \psi(\lambda) = k_0 \),

where the matrix \( B_k \) is defined by the relation (1).

**Proof.** Let \( B_k = [a^{(0)}(1) \ldots a^{(k_0 - 1)}] \).

First we prove that \( \text{rank } B_k = k + 1 \) \( (k = 0, 1, \ldots, k_0 - 1) \). From the definition of \( k_0 \) it follows that \( \text{rank } B_{k_0} = \text{rank } B_{k_0 - 1} \).
For \( k = 1 \) \( \text{rank } B_1 = \text{rank } B_0 = 1 \).

However, for \( k > 0 \) we have:

\[
\begin{align*}
\text{rank } B_1 & > \text{rank } B_0 = 1 \Rightarrow \text{rank } B_1 = 2, \\
\text{rank } B_2 & > \text{rank } B_1 = 2 \Rightarrow \text{rank } B_2 = 3,
\end{align*}
\]

\[
\begin{align*}
\text{rank } B_{k_0 - 1} & > \text{rank } B_{k_0 - 2} = k_0 - 1 \Rightarrow \text{rank } B_{k_0 - 1} = k_0.
\end{align*}
\]

Therefore \( \text{rank } B_k = k + 1 \) \( (k \in \{0, 1, 2, \ldots, k_0 - 1\} \)
and \( \text{rank } B_{k_0} = \text{rank } B_{k_0 - 1} = k_0 \).
Hence it follows that the columns \( a^{(0)}, a^{(1)}, \ldots, a^{(k_0 - 1)} \) are linear independent and the column \( a^{(k_0)} \) can be written as the linear combination of the columns \( a^{(0)}, a^{(1)}, \ldots, a^{(k_0 - 1)} \), so there exists \( \alpha = (\alpha_0, \alpha_1, \ldots, \alpha_{k_0 - 1}) \in C^{k_0} \) such that

\[
\begin{align*}
\alpha_0 a^{(0)} + \alpha_1 a^{(1)} + \cdots + \alpha_{k_0 - 1} a^{(k_0 - 1)} &= -a^{(k_0)}.
\end{align*}
\]

It denotes that

\[
\alpha_0 I + \alpha_1 A + \cdots + \alpha_{k_0 - 1} A^{k_0 - 1} + A^{k_0} = 0 \in M_n
\]

and the polynomial \( f(\lambda) = \alpha_0 + \alpha_1 \lambda + \cdots + \lambda^{k_0} \) is the annihilatory polynomial of the matrix \( A \).

For \( k > k_0 \), \( m = k - k_0 \) and any arbitrary numbers \( \beta_0, \beta_1, \ldots, \beta_m - 1 \in C \) the polynomial \( g(\lambda) = f(\lambda)(\beta_0 + \beta_1 \lambda + \cdots + \beta_m - 1 \lambda^{m - 1} + \lambda^m) = \gamma_0 + \gamma_1 \lambda + \cdots + \gamma_{k_0 - 1} \lambda^{k_0 - 1} + \lambda^k \) is the annihilatory polynomial of the matrix \( A \), too.

Therefore

\[
\begin{align*}
\gamma_0 I + \gamma_1 A + \cdots + \gamma_{k_0 - 1} A^{k_0 - 1} + A^k &= 0 \in M_n,
\end{align*}
\]

\[
\begin{align*}
\gamma_0 a^{(0)} + \gamma_1 a^{(1)} + \cdots + \gamma_{k_0 - 1} a^{(k_0 - 1)} + a^{(k)} &= 0 \in M_{n^2}.
\end{align*}
\]

(2)

In the matrix \( B_k = [a^{(0)}(1) \ldots a^{(k_0 - 1)}(1) a^{(k_0)} a^{(k_0 + 1)} \ldots a^{(k)}] \)
the column \( a^{(j)} \) can be multiplied by \(-\gamma_j (j = 0, 1, \ldots, k - 1)\)
and added to the column \( a^{(k)} \). Hence (2) we have

\[
\begin{align*}
\text{rank } B_k &= \text{rank } [a^{(0)}(1) a^{(1)} \ldots a^{(k_0 - 1)}(1) a^{(k_0)}(0) \cdots 0] = \text{rank } B_{k - 1}.
\end{align*}
\]

Similarly transformation can be used to the matrix \( B_{k - 1} \).
At the end, we have

\[
\begin{align*}
\text{rank } B_k &= \text{rank } [a^{(0)}(1) a^{(1)} \ldots a^{(k_0 - 1)}(1) a^{(k_0)}(0) \cdots 0] = \text{rank } B_{k_0} = k_0
\end{align*}
\]

for \( k \geq k_0 \). This finishes the proof of 2) of the Theorem 1.

Now we prove that if \( \psi(\lambda) \) is the minimal polynomial of the matrix \( A \) then \( \deg \psi(\lambda) = k_0 \).

Hence that \( \text{rank } B_{k_0} = \text{rank } B_{k_0 - 1} = k_0 \) it follows that the set of equations

\[
B_{k_0 - 1} \alpha = -a^{(k_0)},
\]

with the unknown \( \alpha = [\alpha_0 \alpha_1 \ldots \alpha_{k_0 - 1}]^T \in C^{k_0} \), has only one solution and

\[
\alpha_0 I + \alpha_1 A + \cdots + \alpha_{k_0 - 1} A^{k_0 - 1} + A^{k_0} = 0 \in M_n,
\]

besides

\[
\alpha_0 + \alpha_1 \lambda + \cdots + \alpha_{k_0 - 1} \lambda^{k_0 - 1} + \lambda^k,
\]

is the annihilatory polynomial of the matrix \( A \).

Hence that \( \text{rank } B_k = k + 1 \) \( (k = 0, 1, \ldots, k_0 - 1) \) it follows that the set of equations

\[
B_{k - 1} \alpha = -a^{(k)},
\]

with the unknown \( \alpha = [\alpha_0 \alpha_1 \ldots \alpha_{k - 1}]^T \in C^{k} \), has not the solutions.

This denotes that the polynomial (3) is the minimal polynomial of the matrix \( A \) and \( \deg \psi(\lambda) = k_0 \).

Now, we give the algorithm for the calculation of the degree and coefficients of the minimal polynomial of the matrix \( A = [a_{ij}] \in M_n \).

Consider the matrix

\[
B_n = [a^{(0)}(1) \ldots a^{(n)}] \in M_{n^2}, n + 1.
\]

which is defined in (1).

The elements of the matrix \( B_n \) are denoted by \( b_{ij} \), therefore \( B_n = [b_{ij}] \in M_{n^2, n + 1} \), where \( b_{11} = 1, b_{12} = a^{(1)}, \ldots, b_{1,n + 1} = a^{(n)} \), \( b_{n+1,1}, \ldots, b_{n+1,n} = 0 \).

We will calculate the rank of the matrix \( B_n \) by Gaussian elimination, except interchange and cancel of the null columns.

We obtain

\[
\begin{pmatrix}
1 & b_{12} & \cdots & b_{1,n + 1} \\
0 & b_{22} & \cdots & b_{2,n + 1} \\
& \ddots & \ddots & \ddots \\
0 & b_{n^2,2} & \cdots & b_{n^2, n + 1}
\end{pmatrix}
\]

rank \( B_n \) = rank

where, for example \( b_{22}^{(1)} = b_{22}, \ldots, b_{2,n + 1}^{(1)} = b_{2,n + 1}, b_{n^2,2}^{(1)} = b_{n^2,2} - b_{12} \).

From the Lemma 1 it follows that \( n \in K = \{k \in N : \text{rank } B_k = \text{rank } B_{k - 1}\} \).
An algorithm for the calculation of the minimal polynomial

Therefore there exists \( r \in N \) such that \( r \leq n \) and

\[
\begin{pmatrix}
1 & b_{12} & b_{13} & \ldots & \ldots & \ldots & \ldots & \ldots & b_{1,n+1} \\
0 & b_{12}^{(1)} & b_{23}^{(1)} & \ldots & \ldots & \ldots & \ldots & \ldots & b_{2,n+1}^{(1)} \\
0 & 0 & b_{33}^{(2)} & \ldots & \ldots & \ldots & \ldots & \ldots & b_{3,n+1}^{(2)} \\
\vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\
0 & 0 & 0 & \ldots & \ldots & \ldots & \ldots & \ldots & b_{r,n+1}^{(r-1)}
\end{pmatrix},
\]

where \( b_{ij}^{(r-1)} \neq 0 \) (\( i = 1, 2, \ldots, r \)).

From this it follows that \( \text{rank}_B j = j \) (\( j = 1, 2, \ldots, n \)), \( \text{rank}_B r = r \).

Therefore \( k_0 = \min K = r \) and \( \deg \psi(\lambda) = r = k_0 \).

Thus, by Gaussian elimination we can compute the degree of the minimal polynomial of the matrix \( A \).

Hence that \( \det B_{r-1} = \det B_{k_0-1} \neq 0 \) and \( \text{rank}_B k_0 = \text{rank}_B k_0-1 = k_0 \) it follows that the set of equations

\[
B_{k_0-1} \alpha = -\delta^{(k_0)},
\]

with the unknown \( \alpha = [\alpha_0 \alpha_1 \ldots \alpha_{k_0-1}]^T \in C^{k_0} \), has only one solution and

\[
\alpha + \alpha_1 A + \cdots + \alpha_{k_0-1} A^{k_0-1} + A^{k_0} = 0 \in M_n.
\]

Therefore \( \alpha_0, \alpha_1, \ldots, \alpha_{k_0-1} \) are the coefficients of the minimal polynomial of the matrix \( A \). The set of Eq. (4) is equivalent to the set of equations

\[
\tilde{B} \alpha = \tilde{b},
\]

where

\[
\tilde{B} = \begin{pmatrix}
1 & b_{11} & \ldots & \ldots & b_{1r} \\
0 & b_{21}^{(1)} & \ldots & \ldots & b_{2r}^{(1)} \\
0 & 0 & b_{33}^{(2)} & \ldots & b_{3r}^{(2)} \\
\vdots & \vdots & \vdots & \ddots & \ddots \\
0 & 0 & 0 & \ldots & b_{r,r}^{(r-1)} \end{pmatrix}, \quad \tilde{b} = \begin{pmatrix}
b_{1,r+1} \\
b_{2,r+1}^{(1)} \\
b_{2,r+1}^{(2)} \\
\vdots \\
b_{r,r+1}^{(r-1)} \end{pmatrix},
\]

\( \alpha = [\alpha_0 \alpha_1 \ldots \alpha_{k_0-1}]^T, r = k_0 \).

**Example 2.** We will calculate the minimal polynomial of the matrix

\[
A = \begin{pmatrix}
3 & -3 & 2 \\
-1 & 5 & -2 \\
-1 & 3 & 0
\end{pmatrix}.
\]

In this example we have

\[
A^2 = \begin{pmatrix}
10 & -18 & 12 \\
-6 & 22 & -12 \\
-6 & 18 & -8
\end{pmatrix},
\]

\[
A^3 = \begin{pmatrix}
36 & -84 & 56 \\
-28 & 92 & -56 \\
-28 & 84 & -48
\end{pmatrix},
\]

\[
\text{rank}_B 3 = \begin{pmatrix}
1 & 3 & 10 & 36 \\
0 & -3 & -18 & -84 \\
0 & 2 & 12 & 56 \\
0 & -1 & -6 & -28
\end{pmatrix},
\]

\[
\text{rank}_B 3 = 2.
\]

\[
A = [\alpha_0 \alpha_1] = [8 -6]^T.
\]

Therefore, \( \psi(\lambda) = \lambda^2 - 6\lambda + 8 \) is the minimal polynomial of the matrix \( A \).

**REFERENCES**
