Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 38, pp. 457-462, August 2022.

ON M-TH ROOTS OF COMPLEX MATRICES*

HEGUO LIU† AND JING ZHAO‡

Abstract. For an $n \times n$ matrix M, $\sigma(M)$ denotes the set of all different eigenvalues of M. In this paper, we will prove two results on the m-th ($m \geq 2$) roots of a matrix A. Firstly, let X be an m-th root of A. Then X can be expressed as a polynomial in A if and only if $\text{rank}X^2 = \text{rank}X$ and $|\sigma(X)| = |\sigma(A)|$. Secondly, let X and Y be two m-th roots of A. If both X and Y can be expressed as polynomials in A, then $X = Y$ if and only if $\sigma(X) = \sigma(Y)$.

Key words. Root, Rank, Eigenvalue, Unipotent matrix, Chinese remainder theorem.

AMS subject classifications. 15A18, 15A21.

1. Introduction. Let A be a square matrix, and let m be a positive integer. A matrix X is called an m-th root of a matrix A if $X^m = A$. For a nonsingular complex matrix A, there always exists an m-th root, which is, in general, not representable in the form of a polynomial in A; see [1]. It is well-known that every positive semidefinite Hermitian matrix H has a unique m-th root Y such that Y is also a positive semidefinite Hermitian matrix, and Y can be expressed as a polynomial in H; see [2]. For square root, the following result appears in [2, Theorem 6.4.12].

THEOREM 1.1. *Let A be an $n \times n$ complex matrix. If A is singular and has Jordan canonical form $A = SJS^{-1}$, let $J_{k_1}(0) \oplus J_{k_2}(0) \oplus \cdots \oplus J_{k_p}(0)$ be the singular part of J with the blocks arranged in decreasing order of size:*

$$k_1 \geq k_2 \geq \cdots \geq k_p \geq 1.$$

Define $\triangle_1 = k_1 - k_2$, $\triangle_3 = k_3 - k_4$, $\cdots$ Then A has a square root if and only if $\triangle_i = 0$ or 1 for $i = 1, 3, 5, \cdots$ and , if p is odd, $k_p = 1$. Moreover, A has a square root that is a polynomial in A if and only if $k_1 = 1$, a condition that is equivalent to requiring that $\text{rank}A = \text{rank}A^2$.

Let λ be an eigenvalue of a square matrix A, the dimension of the eigenspace of A corresponding to λ is called the geometric multiplicity of λ, the multiplicity of λ as a zero of the characteristic polynomial of A is called the algebraic multiplicity of λ. It is well-known that $\text{rank}A = \text{rank}A^2$ is equivalent to the geometric multiplicity of the eigenvalue 0 of A is equal to its algebraic multiplicity. More related results on these multiplicities can be found in [4].

2. Main results. Let $\sigma(M)$ be the set of all different eigenvalues of a matrix M. In this paper, we will study when an m-th root of a given matrix A can be expressed as a polynomial in A. Our aim is to prove the following two theorems.

THEOREM 2.1. *Let A be a complex square matrix, and let X be an m-th root of A, $m \geq 2$. Then X can be expressed as a polynomial in A if and only if* $\text{rank}X^2 = \text{rank}X$ *and* $|\sigma(A)| = |\sigma(X)|$.

*Received by the editors on March 26, 2022. Accepted for publication on July 22, 2022. Handling Editor: Panagiotis Psarrakos. Corresponding Author: Jing Zhao

†School of Science, Hainan University, Haikou, 570228, China (liuheguo0@163.com). Supported by the National Natural Science Foundation of China (12171142).

‡School of Mathematics and Statistics, Hubei University, Wuhan, 430062, China (jzhao0@163.com).

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 38, pp. 457-462, August 2022.

THEOREM 2.2. *Suppose that X and Y are two m-th roots of a complex square matrix A which can be expressed as polynomials in A, then $X = Y$ if and only if $\sigma(X) = \sigma(Y)$.*

From these theorems, we can obtain some corollaries.

COROLLARY 2.3. *Let A be an $n \times n$ nonsingular matrix, and let X be an m-th root of A. Then X can be expressed as a polynomial in A if and only if $|\sigma(A)| = |\sigma(X)|$.*

The following first example is a simple one illustrating Theorem 2.1. Other two counterexamples show that the conditions $\text{rank}X^2 = \text{rank}X$ and $|\sigma(A)| = |\sigma(X)|$ in Theorem 2.1 are necessary.

Example 2.4. Let $A = \begin{pmatrix} 1 & -4 & -4 \\ -1 & 4 & 4 \\ 1 & -3 & -3 \end{pmatrix}$, and $X = \begin{pmatrix} -1 & 1 & 1 \\ 1 & -1 & -1 \\ -1 & 0 & 0 \end{pmatrix}$. Then $\text{rank}X^2 = \text{rank}X = 2$ and $\sigma(A) = \{0, 1\}, \sigma(X) = \{0, -1\}$. We can prove that $X^4 = A$ and $X = -\frac{7}{4}A + \frac{3}{4}A^2$.

Example 2.5. Let ω be an m-th primitive roots of unity, and let $X = diag(\omega, \omega^2, \cdots, \omega^n)$. Then $X^m = I$ and $\text{rank}X^2 = \text{rank}X$, but $|\sigma(X)| > |\sigma(I)| = 1$. Clearly, X cannot be expressed as a polynomial in I.

Example 2.6. Let X be a nilpotent matrix of rank 1. Then $X^m = \text{O}$ for any integer $m \geq 2$, and $\sigma(X) = \sigma(\text{O}) = \{0\}$, but $1 = \text{rank}X > \text{rank}X^2 = 0$. Clearly, X cannot be expressed as a polynomial in O.

When $\text{rank}A^2 = \text{rank}A$, there exists an m-th roots of A which can be expressed as a polynomial in A. More accurately, we can obtain the following conclusion from Theorem 2.2.

COROLLARY 2.7. *If* $\text{rank}A^2 = \text{rank}A$, *then*

$$|\{X | X^m = A \text{ and } X = f(A)\}| = m^s,$$

where s is the number of non-zero different eigenvalues of A.

The following Chinese Remainder Theorem is a special form of [3, Theorem 2.25], and it is a key tool in the argument of this paper.

THEOREM 2.8 (Chinese Remainder Theorem). *Suppose that $m_1(\lambda), m_2(\lambda), \cdots, m_s(\lambda)$ are s pairwise relatively prime polynomials over a field, then for any s polynomials $f_1(\lambda), f_2(\lambda), \cdots, f_s(\lambda)$, there exists a unique polynomial $f(\lambda)$ whose degree is less than the sum of the degrees of these $m_i(\lambda)$ $(i = 1, 2, \cdots, s)$, such that*

$$\begin{cases} f(\lambda) \equiv f_1(\lambda) & (mod\ m_1(\lambda)) \\ f(\lambda) \equiv f_2(\lambda) & (mod\ m_2(\lambda)) \\ \qquad \vdots & \\ f(\lambda) \equiv f_s(\lambda) & (mod\ m_s(\lambda)). \end{cases}$$

3. Proof of Theorem 2.1. We first establish a technique lemma about unipotent matrices. A square matrix U is said to be unipotent if $U - I$ is nilpotent.

LEMMA 3.1. *Let U be a unipotent matrix. Then for any nonzero integer m, U can be expressed as a polynomial in U^m.*

Proof. We first deal with the case $m > 0$. Write $U = I + N$, where I is the identity matrix and N is a nilpotent matrix. Then we choose the least positive integer r such that $N^r = \text{O}$. For any positive integer $0 \leq s \leq r - 1$, we have

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 38, pp. 457-462, August 2022.

$$(U^m)^s = (I+N)^{sm} = I + \binom{sm}{1}N + \binom{sm}{2}N^2 + \cdots + \binom{sm}{r-1}N^{r-1}.$$

Furthermore,

$$\begin{pmatrix} I \\ U^m \\ U^{2m} \\ \cdots \\ U^{(r-1)m} \end{pmatrix} = \begin{pmatrix} 1 & \binom{0m}{1} & \binom{0m}{2} & \cdots & \binom{0m}{r-1} \\ 1 & \binom{1m}{1} & \binom{1m}{2} & \cdots & \binom{1m}{r-1} \\ 1 & \binom{2m}{1} & \binom{2m}{2} & \cdots & \binom{2m}{r-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \binom{(r-1)m}{1} & \binom{(r-1)m}{2} & \cdots & \binom{(r-1)m}{r-1} \end{pmatrix} \begin{pmatrix} I \\ N \\ N^2 \\ \cdots \\ N^{r-1} \end{pmatrix}.$$

Since the above transition matrix is nonsingular, it follows that N is a linear combination of I, U^m, U^{2m}, $\cdots$, $U^{(r-1)m}$. Thus, $U = I + N$ can be expressed as a polynomial in U^m.

Secondly, assume that $m < 0$. By the above argument, U^{-1} can be expressed as a polynomial in U^m. Since U can be expressed as a polynomial in U^{-1}, U can be expressed as a polynomial in U^m. $\square$

Proof of Theorem 2.1. There exists a nonsingular matrix P such that

$$A = P\begin{pmatrix} N & \\ & B \end{pmatrix}P^{-1},$$

where N is a nilpotent matrix and B is a nonsingular matrix. Let r be the least positive integer satisfying $N^r = \mathrm{O}$.

Since $X^m = A$, we have $AX = XA$. Write

$$X = P\begin{pmatrix} X_N & X_{12} \\ X_{21} & X_B \end{pmatrix}P^{-1},$$

where the order of X_N is the same as that of N. We have

$$\begin{pmatrix} N & \\ & B \end{pmatrix}\begin{pmatrix} X_N & X_{12} \\ X_{21} & X_B \end{pmatrix} = \begin{pmatrix} X_N & X_{12} \\ X_{21} & X_B \end{pmatrix}\begin{pmatrix} N & \\ & B \end{pmatrix},$$

which implies that

$$\begin{cases} NX_{12} = X_{12}B \\ BX_{21} = X_{21}N. \end{cases}$$

Then $X_{12} = \mathrm{O}$ because $X_{12}B^r = N^rX_{12} = \mathrm{O}$, and $X_{21} = \mathrm{O}$ because $B^rX_{21} = X_{21}N^r = \mathrm{O}$. It follows that

$$X = P\begin{pmatrix} X_N & \\ & X_B \end{pmatrix}P^{-1}.$$

Note that $X^m = A$, i.e., $X_N^m = N$ and $X_B^m = B$.

Assume that $X = f(A)$ for some polynomial $f(\lambda)$. Then $X_N = f(N)$ and $X_B = f(B)$. We claim that $N = \mathrm{O}$. Suppose that this is false and $N \neq \mathrm{O}$. Since $X_N^{mr} = N^r = \mathrm{O}$, we have

$$\mathrm{rank}X_N > \mathrm{rank}X_N^2 \geq \cdots \geq \mathrm{rank}X_N^m = \mathrm{rank}N.$$

On the other hand, note that

$$X_N = f(N) = k_0I + k_1N + \cdots + k_{r-1}N^{r-1},$$

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 38, pp. 457-462, August 2022.

then $k_0 = 0$ since k_0 is the eigenvalue of nilpotent matrix X_N. This means that

$$X_N = N(k_1 I + k_2 N + \cdots + k_{r-1} N^{r-2}),$$

and $\text{rank}X_N \leq \text{rank}N$, a contradiction. Therefore, $N = X_N = \text{O}$, and $\text{rank}A^2 = \text{rank}A$. Note that B is a nonsingular matrix and $X_B^m = B$, so $\text{rank}X^2 = \text{rank}X_B^2 = \text{rank}X_B = \text{rank}X$.

Since $A = X^m$, we have $|\sigma(A)| \leq |\sigma(X)|$. It follows from $X = f(A)$ that $|\sigma(X)| \leq |\sigma(A)|$. Thus, $|\sigma(A)| = |\sigma(X)|$.

Conversely, suppose that $|\sigma(A)| = |\sigma(X)|$ and $\text{rank}X^2 = \text{rank}X$. Let $\lambda_0 = 0, \lambda_1, \lambda_2, \cdots, \lambda_s$ be all different eigenvalues of X. Then there exists a nonsingular matrix Q such that

$$X = Q \begin{pmatrix} \text{O} & & & & \\ & \lambda_1 U_1 & & & \\ & & \lambda_2 U_2 & & \\ & & & \ddots & \\ & & & & \lambda_s U_s \end{pmatrix} Q^{-1},$$

where U_i is a unipotent matrix of order n_i, $1 \leq i \leq s$. It follows from $X^m = A$ that

$$X^m = Q \begin{pmatrix} \text{O} & & & & \\ & (\lambda_1 U_1)^m & & & \\ & & (\lambda_2 U_2)^m & & \\ & & & \ddots & \\ & & & & (\lambda_s U_s)^m \end{pmatrix} Q^{-1} = A.$$

Note that U_i is a unipotent matrix. By Lemma 3.1, U_i can be expressed as a polynomial in U_i^m. Therefore, $\lambda_i U_i$ can be expressed as a polynomial in $(\lambda_i U_i)^m$. Write $\lambda_i U_i = g_i((\lambda_i U_i)^m)$.

Note also that the characteristic polynomial of $\lambda_i^m U_i^m$ is equal to $(\lambda - \lambda_i^m)^{n_i}$. Since $|\sigma(X)| = |\sigma(A)|$, we have $\lambda_0^m = 0, \lambda_1^m, \lambda_2^m, \cdots, \lambda_s^m$ are all different eigenvalues of A. Hence, $\lambda, (\lambda - \lambda_1^m)^{n_1}, (\lambda - \lambda_2^m)^{n_2}, \cdots, (\lambda - \lambda_s^m)^{n_s}$ are $s+1$ pairwise relatively prime polynomials. According to the Chinese Remainder Theorem, there exists a polynomial $f(\lambda)$ such that

$$\begin{cases} f(\lambda) \equiv 0 & (\ mod\ \lambda\ ) \\ f(\lambda) \equiv g_1(\lambda) & (\ mod\ (\lambda - \lambda_1^m)^{n_1}\ ) \\ f(\lambda) \equiv g_2(\lambda) & (\ mod\ (\lambda - \lambda_2^m)^{n_2}\ ) \\ \quad\vdots & \\ f(\lambda) \equiv g_s(\lambda) & (\ mod\ (\lambda - \lambda_s^m)^{n_s}\ ). \end{cases}$$

Therefore, X can be expressed as a polynomial in A. □

4. Proof of Theorem 2.2. For the proof, we require a lemma.

LEMMA 4.1. *Let U and V be two unipotent matrices. If $U^m = V^m$ for some nonzero integer m, then $U = V$.*

Proof. Without loss of generality, assume that $m > 0$. By induction on the order of U and V. Since 1 is the unique eigenvalue of U, there exists a nonzero vector α such that $U\alpha = \alpha$. Thus, $U^m \alpha = \alpha = V^m \alpha$ and

$$(I - V^m)\alpha = (I + V + \cdots + V^{m-1})(I - V)\alpha = 0.$$

Note that 1 is the unique eigenvalue of V, so $I + V + \cdots + V^{m-1}$ is nonsingular. Thus, $V\alpha = \alpha$.

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 38, pp. 457-462, August 2022.

Let $P = (\alpha, \alpha_2, \cdots, \alpha_n)$ be a nonsingular matrix. Then

$$P^{-1}UP = \begin{pmatrix} 1 & X \\ & U_1 \end{pmatrix}, \qquad P^{-1}VP = \begin{pmatrix} 1 & Y \\ & V_1 \end{pmatrix},$$

where U_1 and V_1 are two unipotent matrices. We deduce that

$$P^{-1}U^mP = \begin{pmatrix} 1 & X(I + U_1 + \cdots + U_1^{m-1}) \\ & U_1^m \end{pmatrix},$$

$$P^{-1}V^mP = \begin{pmatrix} 1 & Y(I + V_1 + \cdots + V_1^{m-1}) \\ & V_1^m \end{pmatrix}.$$

It follows from $U^m = V^m$ that $U_1^m = V_1^m$, and

$$X(I + U_1 + \cdots + U_1^{m-1}) = Y(I + V_1 + \cdots + V_1^{m-1}).$$

So $U_1 = V_1$ by the induction hypothesis. Furthermore, $X = Y$ because $I + U_1 + \cdots + U_1^{m-1}$ is nonsingular. Hence, $U = V$. □

The proof of Theorem 2.2 depends on that of Theorem 2.1.

Proof of Theorem 2.2. Only the necessary of the condition is in question. Assume that $\sigma(X) = \sigma(Y)$, we will prove that $X = Y$. Since X can be expressed as a polynomial in A, it follows by Theorem 2.1 that $\text{rank}X^2 = \text{rank}X$. Then, there exists a nonsingular matrix P such that

$$X = P \begin{pmatrix} \mathrm{O} & & & & \\ & \lambda_1 U_1 & & & \\ & & \lambda_2 U_2 & & \\ & & & \ddots & \\ & & & & \lambda_s U_s \end{pmatrix} P^{-1},$$

where U_i is a unipotent matrix of order n_i, $1 \leq i \leq s$, and $\lambda_1, \lambda_2, \cdots, \lambda_s$ are all nonzero different eigenvalues of X.

Note that both X and Y can be expressed as polynomials in A, so $XY = YX$. Furthermore,

$$Y = P \begin{pmatrix} Y_0 & & & & \\ & Y_1 & & & \\ & & Y_2 & & \\ & & & \ddots & \\ & & & & Y_s \end{pmatrix} P^{-1},$$

where the size of Y_i is the same as that of U_i for $1 \leq i \leq s$. Since $X^m = Y^m = A$, we have $Y_0^m = \mathrm{O}$ and $Y_i^m = (\lambda_i U_i)^m$. By Theorem 2.1 again, $\text{rank}Y^2 = \text{rank}Y$. So $Y_0 = \mathrm{O}$.

Next Y_i has a unique eigenvalue because $|\sigma(Y)| = |\sigma(A)|$, and we assume that μ_i be the eigenvalue of Y_i. Then $\mu_i^m = \lambda_i^m$. Moreover, $\mu_i = \lambda_i$ because $\sigma(X) = \sigma(Y)$. Note that

$$\lambda_i^m U_i^m = (\lambda_i U_i)^m = Y_i^m = \lambda_i^m \left(\frac{1}{\lambda_i} Y_i\right)^m,$$

so $U_i^m = (\frac{1}{\lambda_i} Y_i)^m$. Note also that $\frac{1}{\lambda_i} Y_i$ is a unipotent matrix, and thus $U_i = \frac{1}{\lambda_i} Y_i$ by Lemma 4.1. Hence, $Y_i = \lambda_i U_i$, and $X = Y$. □

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 38, pp. 457-462, August 2022.

Proof of Corollary 2.7. Since $X^m = A$ and $X = f(A)$, so by Theorem 2.1 we have $\mathrm{rank}X^2 = \mathrm{rank}X$, $\mathrm{rank}A^2 = \mathrm{rank}A$ and $|\sigma(X)| = |\sigma(A)|$. There exists a nonsingular matrix P such that

$$A = P \begin{pmatrix} \mathrm{O} & & & & \\ & \lambda_1 U_1 & & & \\ & & \lambda_2 U_2 & & \\ & & & \ddots & \\ & & & & \lambda_s U_s \end{pmatrix} P^{-1},$$

where U_i is a unipotent matrix of order n_i, $1 \le i \le s$, and $\lambda_1, \lambda_2, \cdots, \lambda_s$ are all nonzero different eigenvalues of A.

It is easy to prove that

$$X = P \begin{pmatrix} \mathrm{O} & & & & \\ & X_1 & & & \\ & & X_2 & & \\ & & & \ddots & \\ & & & & X_s \end{pmatrix} P^{-1},$$

where the size of X_i is same as that of U_i for $1 \le i \le s$. Then $X_i^m = \lambda_i U_i$, and X_i only has a eigenvalue μ_i such that $\mu_i^m = \lambda_i$. By Theorem 2.2, X is uniquely determined by $\mu_1, \mu_2, \cdots, \mu_s$. Hence, $|\{X | X^m = A \text{ and } X = f(A)\}| = m^s$. $\square$

Acknowledgment. The authors would like to thank the referee for his/her helpful comments and suggestions.

REFERENCES

[1] F.R. Gantmacher. *The Theory of Matrices,* 2 *vols.* Chelsea, New York, 1959.
[2] R.A. Horn and C.R. Johnson. *Topics in Matrix Analysis.* Cambridge University Press, Cambridge, 1985.
[3] T.W. Hungerford. *Algebra.* Springer-Verlag, New York, 1974.
[4] J. Liao, H.G. Liu, M.F. Shao, and X.Z. Xu. A matrix identity and its applications. *Linear Algebra Appl.*, 471:346–352, 2015.