Roots of Matrices

Boonrod Yuttanan\textsuperscript{1} and Chaufah Nilrat\textsuperscript{2}

Abstract

Yuttanan, B. and Nilrat, C.
Roots of Matrices

A matrix $S$ is said to be an $n$\textsuperscript{th} root of a matrix $A$ if $S^n = A$, where $n$ is a positive integer greater than or equal to 2. If there is no such matrix for any integer $n \geq 2$, $A$ is called a rootless matrix. After investigating the properties of these matrices, we conclude that we always find an $n$\textsuperscript{th} root of a non-singular matrix and a diagonalizable matrix for any positive integer $n$. On the other hand, we find some matrix having an $n$\textsuperscript{th} root for some positive integer $n$. We call it $p$-nilpotent matrix.

Key words: roots of matrices, rootless matrix, nilpotent matrix, non-singular matrix, diagonalizable matrix

\textsuperscript{1}Student in Mathematics, \textsuperscript{2}M.S.(Mathematics), Assoc. Prof., Department of Mathematics, Faculty of Science, Prince of Songkla University, Hat Yai, Songkhla 90112 Thailand.
Corresponding e-mail: chaufah.n@psu.ac.th
Received, 18 May 2004 Accepted, 11 October 2004
An $m \times m$ matrix $A$ is called nilpotent if $A^r = 0$ for some positive integer $r \geq 2$. Yood (2002) showed that any nilpotent $m \times m$ matrix $A$ such that $A^{m-1} \neq 0$ is rootless. Such a matrix is called principal nilpotent. After we finished reading this article, we raised the question of which matrices always have an $n^{th}$ root for any positive integer $n$ and which have an $n^{th}$ root only for some positive integer $n$. In this paper, we give the answer for these questions.

1. Roots of non-singular matrices

In this section, we prove that every non-singular matrix has an $n^{th}$ root for any positive integer. Before discussing on a non-singular matrix, we start with a property of upper triangular matrices.

**Lemma 1.1** If $A = [a_{ij}]$ is an upper triangular matrix, then so is $A^n = [\alpha_{ij}]$ and

$$
\alpha_{ij} = \begin{cases} 
0 & \text{if } i > j, \\
\alpha_{ij}^n & \text{if } i = j, \\
\sum_{i_1 < i_2 < \cdots < i_k} a_{i_1i_2} a_{i_2i_3} \cdots a_{i_{k}i_{k+1}} & \text{if } i < j.
\end{cases}
$$

**Proof.** We give a proof by mathematical induction. For $n = 2$, we have

$$
A^2 = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\
0 & a_{22} & \cdots & a_{2m} \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & a_{nn} \end{pmatrix}^2 = \begin{pmatrix} a_{11}^2 & a_{11}a_{12} + a_{12}a_{21} & \cdots & a_{11}a_{1m} + \cdots + a_{1m}a_{m1} \\
0 & a_{22}^2 & \cdots & a_{22}a_{2m} + \cdots + a_{2m}a_{mm} \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & a_{nn}^2 \end{pmatrix} = \begin{pmatrix} \alpha_{ij} \end{pmatrix}
$$

Apparently, $A^2$ is an upper triangular matrix such that for each $i, j, 1 \leq i, j \leq m$,

$$
\alpha_{ij} = \begin{cases} 
0 & \text{if } i > j, \\
\alpha_{ij}^2 & \text{if } i = j, \\
\sum_{i_1 < i_2 < \cdots < i_k} a_{i_1i_2} a_{i_2i_3} \cdots a_{i_{k}i_{k+1}} & \text{if } i < j.
\end{cases}
$$
Now, we assume that $A' = [\alpha_{ij}]$ where

$$\alpha_{ij} = \begin{cases} 
0 & \text{if } i > j, \\
a_{ii} & \text{if } i = j, \\
\sum_{i \in \mathbb{Z}, j \in \mathbb{Z}, i \neq j} a_{ij} a_{ki} a_{li} \cdots a_{i-1j} & \text{if } i < j.
\end{cases}$$

Then

$$A^{k+j} = A^k A$$

$$= \begin{pmatrix} 
\alpha_{11} & \alpha_{12} & \ldots & \alpha_{1m} \\
0 & \alpha_{22} & \ldots & \alpha_{2m} \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & \alpha_{mm}
\end{pmatrix} \begin{pmatrix} 
a_{11} & a_{12} & \ldots & a_{1m} \\
0 & a_{22} & \ldots & a_{2m} \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & a_{mm}
\end{pmatrix}$$

$$= \begin{pmatrix} 
\alpha_{11} a_{11} & \alpha_{11} a_{12} + \alpha_{12} a_{22} & \cdots & \alpha_{11} a_{1m} + \alpha_{12} a_{2m} + \cdots + \alpha_{mm} a_{mm} \\
0 & \alpha_{22} a_{22} & \cdots & \alpha_{22} a_{2m} + \alpha_{23} a_{3m} + \cdots + \alpha_{mm} a_{mm} \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \alpha_{mm} a_{mm}
\end{pmatrix}$$

$$= \left[ \alpha'_{ij} \right]$$

It is clear that $\alpha'_{ij} = 0$ if $i > j$. For each integer $i$, $1 \leq i \leq m$, $\alpha'_{ii} = a_{ii} a_{ii} = a_{ii}^2$. We also obtain

$$\alpha'_{ij} = \sum_{k=1}^{j} \alpha_{ik} a_{kj} = \sum_{k=1}^{j} \left( \sum_{l=1}^{m} a_{il} a_{kj} \right) a_{ij} = \sum_{l=1}^{m} a_{il} a_{kj} a_{ij}$$

for all integers $i$ and $j$, $1 \leq i < j \leq m$. □

**Theorem 1.2** Let $A$ be an $m \times m$ complex matrix. If $A$ is non-singular, then $A$ always has an $n^\text{th}$ root for any positive integer $n$.

**Proof.** Let $A$ be non-singular. By Schur’s theorem (Strang, 1988), there exists a non-singular matrix $S$ such that $A = S B S^\dagger$ where $B$ is upper triangular. Let $B = \left[ b_{ij} \right]_{i \neq j}$. We have $\det(B) \neq 0$; that is, $b_{ii} \neq 0$ for $i = 1, 2, \ldots, m$. Let $b_{ii}^*$ be any $n^\text{th}$ root of $b_{ii}$. If $b_{ii} = b_{ii}^*$, we let $b_{ii}^* = b_{ii}$. We define $C = \left[ c_{ij} \right]_{i \neq j}$ as follows.

For each $i$, $c_{ii} = b_{ii}^*$. For $i > j$, let $c_{ij} = 0$. For $j = i+1$, let $c_{ij} = b_{ij} / \sum_{p=0}^{m-1} c_{pp} c_{ij}^p$. For $j = i+k$, where $2 \leq k \leq m-i$, let $c_{ij} = (b_{ij} - R_{ij}) / \sum_{p=0}^{m-i-k} c_{ii} c_{ij}^p$, and $R_{ij}$ be the sum of the products $c_{ii} c_{ij} c_{ij}^2 \cdots c_{ij}^{m-1}$, where the sum is taken over integers $k_1, k_2, \ldots, k_{m-i}$ such that $i \leq k_1 \leq \ldots \leq k_{m-i} \leq j$ and none of the term in the products contains $c_{ij}$.
Since \( b_i \neq 0 \) for \( i = 1, 2, \ldots, m \), we have \( c_i \neq 0 \) for each \( i \) and \( c_i^{n-1} + c_i^{n-2}c_{i+k} + \ldots + c_i^{n-1} \neq 0 \) for \( 1 \leq k \leq m-i \). This guarantees that \( c_{i+k} \) is well-defined. We claim that \( C' = B \).

Let \( C'' = \left[ y_{ij} \right]_{m \times m} \). By Lemma 1.1, we have

\[
y_{ij} = \begin{cases} 
0 & \text{if } i > j, \\
c_i^n & \text{if } i = j, \\
\sum_{i \leq k \leq j \leq i+k} c_{i+k} \ldots c_{k} & \text{if } j = i + k, k = 1, 2, \ldots, m-i.
\end{cases}
\]

If \( i = j, y_{ij} = c_i^n = (b_i^n) = b_i \).

If \( j = i + 1 \), we have

\[
y_{i,i+1} = \sum_{i \leq k \leq j \leq j+k} c_i \ldots c_{k} = b_i^i(c_i^{n-1} + c_i^{n-2}c_{i+1} + \ldots + c_i^{n-1})
\]

\[
= b_i^i \frac{c_i^{n-1} + c_i^{n-2}c_{i+1} + \ldots + c_i^{n-1}}{c_i^{n-1} + c_i^{n-2}c_{i+1} + \ldots + c_i^{n-1}}
\]

\[
= b_i^i.
\]

If \( j = i + k \), when \( k = 2, 3, \ldots, m-i \), we have

\[
y_{i,i+k} = \sum_{i \leq k \leq j \leq j+k} c_i \ldots c_{k} = b_i^i(c_i^{n-1} + c_i^{n-2}c_{i+k} + \ldots + c_i^{n-1}) + R_{i,i+k}
\]

\[
= b_i^i + \frac{(b_i^i - R_{i,i+k})(c_i^{n-1} + c_i^{n-2}c_{i+k} + \ldots + c_i^{n-1}) + R_{i,i+k}}{c_i^{n-1} + c_i^{n-2}c_{i+k} + \ldots + c_i^{n-1}}
\]

\[
= b_i^i.
\]

Then we obtain \( \left[ y_{ij} \right] = \left[ b_{ij} \right] \). Therefore \( A = (SCS^{-1})^n \).

We illustrate the procedure in Theorem 1.2 by the following example. Let

\[
B = \begin{pmatrix} 
8 & -12 & 7 & -8 \\
0 & -1 & 0 & 6 \\
0 & 0 & 1 & -28 \\
0 & 0 & 0 & 8 
\end{pmatrix}
\]

A third root of \( B \) is a matrix \( C = \left[ c_{ij} \right] \) where \( c_{ij} = 0 \), if \( i > j \) and \( c_{11} = 2, c_{22} = -1, c_{33} = 1, c_{44} = 2 \).

\[
c_{12} = b_{12} \left[ c_{11}^2 + c_{11}c_{22} + c_{22}^2 \right] = (-12) \left[ 2^2 + (2)(-1) + (-1)^2 \right] = -4.
\]

\[
c_{23} = b_{23} \left[ c_{22}^2 + c_{22}c_{33} + c_{33}^2 \right] = (0) \left[ (-1)^2 + (2)(1) + 1 \right] = 0.
\]

\[
C = \begin{pmatrix} 
2 & (-1) & 1 & 2 \\
0 & -1 & 0 & 6 \\
0 & 0 & 1 & -28 \\
0 & 0 & 0 & 8 
\end{pmatrix}
\]
\[ c_{34} = b_{34}/\left[ c_{33}^2 + c_{35}c_{44} + c_{44}^2 \right] = (-28)/(1^2 + (1)(2) + 2^2) = -4. \]
\[ c_{13} = \frac{1}{D_{13}} \left[ c_{11}c_{21}c_{33} + c_{12}c_{22}c_{33} + c_{13}c_{23}c_{33} \right] = \frac{1}{2} \left[ -7 - (2)(-4)(0) + (-4)(-1)(0) + (-4)(0)(1) \right]/\left[ 2^2 + (2)(1) + 1^2 \right] = 1, \]
\[ c_{24} = \frac{1}{D_{24}} \left[ c_{11}c_{22}c_{24} + c_{12}c_{22}c_{44} + c_{13}c_{23}c_{34} + c_{22}c_{34}c_{44} \right] = \frac{1}{2} \left[ 6 - (1)(0)(-4) + (0)(1)(-4) + (0)(-4)(2) \right]/\left[ (-1)^2 + (-1)(2) + 2^2 \right] = 2, \]
\[ c_{14} = \frac{1}{D_{14}} \left[ c_{11}c_{13}c_{24} + c_{12}c_{12}c_{24} + c_{13}c_{23}c_{34} + c_{14}c_{24}c_{44} + c_{23}c_{34}c_{44} \right] = \frac{1}{2} \left[ -8 - (2)(-4)(2) + (2)(1)(-4) + (-4)(-1)(2) + (-4)(0)(-4) + (4)(2)(2) + (1)(1)(-4) + (1)(-4)(2) \right]/\left[ 2^2 + (2)(2) + 2^2 \right] = 3. \]

That is \( \begin{pmatrix} 2 & -4 & 1 & 3 \\ 0 & -1 & 0 & 2 \end{pmatrix} \) is a third root of \( B \).

Some singular matrices also have an \( n^{th} \) root such as
\[
\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}^*.
\]

Moreover, we have a singular matrix \( \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \) as a rootless matrix (Yood, 2002).

**Corollary 1.3** If all eigenvalues of \( A \) are not zero, then \( A \) has an \( n^{th} \) root.

**Proof.** Since \( A \) has all non-zero eigenvalues, \( A \) is a non-singular matrix. \( \square \)

**Note.** If only one eigenvalue of \( A \) is zero, in Theorem 1.2, we have \( b_i = 0 \) for only one value of \( i \).

That means we still have \( c_{i1}^{n-1} + c_{i2}^{n-2}c_{i3}^{i1} + \ldots + c_{i4}^{n-1} \neq 0 \). Then we can say that "A matrix with only one zero eigenvalue always has an \( n^{th} \) root".

2. **Roots of diagonalizable matrices**

In this section, we consider an \( n^{th} \) root of a diagonalizable matrix.

**Theorem 2.1** Let \( A \) be an \( m \times m \) complex matrix. If \( A \) is diagonalizable, then \( A \) has an \( n^{th} \) root, for any positive integer \( n \).

**Proof.** Let \( A \) be a diagonalizable matrix, i.e., there exists a non-singular matrix \( S \) such that \( A = SDS^{-1} \) where \( D = \begin{pmatrix} d_{ij}^{k} \end{pmatrix} \) is a diagonal matrix.

Let \( D^{1/n} = \begin{pmatrix} d_{ij}^{1/n} \end{pmatrix} \), where \( d_{ij}^{1/n} \) is an \( n^{th} \) root of \( d_{ij} \). So \( A = S(D^{1/n})S^{-1} = (SD^{1/n}S^{-1})^n \). Therefore an \( n^{th} \) root of \( A \) exists. \( \square \)
However, we have some non-diagonalizable matrices having an \( n \)th root, for example,
\[
\begin{pmatrix}
2 & 0 & 0 \\
0 & 2 & 0 \\
2 & 0 & 2
\end{pmatrix}
\]
has an \( n \)th root because it is a non-singular matrix. Moreover, we see that diagonalizable matrices and non-singular matrices are not the only matrices which have an \( n \)th root, since
\[
\begin{pmatrix}
1 & 1 & 1 \\
0 & 1 & 1 \\
0 & 0 & 0
\end{pmatrix} = \begin{pmatrix}
1 & 1 & 1 \\
0 & 1 & 1 \\
0 & 0 & 0
\end{pmatrix}^n, \quad n \geq 2.
\]

There are some further questions the reader might like to consider. For instance, what is a necessary and sufficient condition for a matrix to have an \( n \)th root?

As an immediate consequence of the above theorem, we can conclude that a matrix \( A \) with all distinct eigenvalues has an \( n \)th root. On the other hand, a real symmetric matrix also has an \( n \)th root for any positive integer \( n \), as well as a complex Hermitian matrix and a normal matrix.

### 3. Roots of \( p \)-nilpotent matrices

In the previous two sections, we considered matrices whose \( n \)th root always exists for any positive integer \( n \). In this section, we consider some kind of matrices which has an \( n \)th root for just some value of \( n \).

An \( m \times m \) matrix \( A \) is called \( p \)-nilpotent if \( A \) is a nilpotent matrix but not principal nilpotent and \( p \) is the least positive integer such that \( A^p = 0 \) but \( A^{p-1} \neq 0 \).

Before discussing on \( p \)-nilpotent matrices, we first give the following lemma.

**Lemma 3.1** Let \( A \) be an \( m \times m \) complex matrix. If \( A^k = 0 \) for some \( k \geq 2 \), then \( A^m = 0 \).

**Proof.** If \( 2 \leq k \leq m \), then we are done. Now we suppose \( k > m \). By Schur’s theorem (Strang, 1988), there exists a non-singular matrix \( S \) such that \( A = SBS^{-1} \) where \( B \) is upper triangular. Since \( A^k = 0 \), we have \( B^k = 0 \).

Let \( B = \begin{pmatrix} b_{ij} \end{pmatrix}_{m \times m} \) and \( B' = \begin{pmatrix} b_i' \end{pmatrix}_{m \times m} \). For \( 1 \leq i \leq m \), we have \( b_i' = b_i^k \), so \( b_i = 0 \). Then \( B \) is strictly upper triangular. It was proved by Yood (2002) that \( B' = 0 \). Therefore \( A' = 0 \).

**Theorem 3.2** Let \( A \) be an \( m \times m \) \( p \)-nilpotent matrix. If an \( n \)th root of \( A \) exists, then \( n \leq m - p + 1 \).

**Proof.** The proof is by contradiction. Suppose that \( A = S' \), for \( r \geq m - p + 2 \). Then \( S'' = A'' = 0 \) so that \( S \) is an \( m \times m \) nilpotent matrix. By Lemma 3.1, the \( m \)th power of \( S \) is zero. Therefore, \( S' = 0 \) for all positive integers \( k \geq 2 \). But we also have \( S''^{r-1} = A''^{r-1} = 0 \). Now \( p \geq 2 \), hence, \( 2r - 2 \leq rp - r \), so that \( r + p - 2 \leq rp - r \). Since \( r \geq m - p + 2 \), \( m \leq r + p - 2 \leq rp - r \). Therefore \( S'^r = 0 \) or \( A'^r = 0 \), which is contrary to the hypotheses on \( A \). Hence, if an \( n \)th root of \( p \)-nilpotent matrix exists, then \( n \leq m - p + 1 \).

Let \( A \) be a 2-nilpotent matrix of size \( 3 \times 3 \), i.e., \( A^2 = 0 \). By Schur’s theorem, \( A \) is of the form \( SBS^{-1} \) where \( S \) is non-singular and \( B \) is upper triangular. Hence \( B' = 0 \). This implies \( B'' = 0 \). By Yood (2002), \( B \) is a strictly upper triangular matrix. It is possible to classify \( B \) which is not principal nilpotent as five different types:
We observe that
\[
\begin{pmatrix}
0 & a & 0 \\
0 & 0 & 0 \\
0 & 0 & 0 \\
\end{pmatrix}
= \begin{pmatrix}
0 & 0 & a \\
0 & 0 & 0 \\
0 & 0 & 1 \\
\end{pmatrix}^2,
\]
\[
\begin{pmatrix}
0 & 0 & a \\
0 & 0 & 0 \\
0 & 0 & 0 \\
\end{pmatrix}
= \begin{pmatrix}
0 & a & 0 \\
0 & 0 & 1 \\
0 & 0 & 0 \\
\end{pmatrix}^2,
\]
\[
\begin{pmatrix}
0 & a & b \\
0 & 0 & 0 \\
0 & 0 & 0 \\
\end{pmatrix}
= \begin{pmatrix}
0 & -a & b \\
0 & -1 & \frac{b}{a} \\
0 & a & 1 \\
\end{pmatrix}^2,
\]
\[
\begin{pmatrix}
0 & 0 & b \\
0 & 0 & a \\
0 & 0 & 0 \\
\end{pmatrix}
= \begin{pmatrix}
1 & \frac{b}{a} & 0 \\
\frac{a}{b} & 1 & -a \\
0 & 0 & 1 \\
\end{pmatrix}^2.
\]

Then we see that all five types of \( B \) has a square root, say \( T \). Therefore \( A = ST^2S^{-1} = (STS)^2 \). This shows that a square root of any 2-nilpotent matrix of size 3×3 always exists.

**Conclusion and Discussion**

According to this article, we obtain a formula for calculating an \( n \)th root of a matrix which is non-singular or diagonalizable.

However, being non-singular or diagonalizable are not necessary for matrices to have \( n \)th roots. The reader may try to find other properties of his own. In addition, a matrix having an \( n \)th root for some positive integer \( n \) is not only a \( p \)-nilpotent matrix.

**References**

