The Square and nth Roots of Square Matrices
$$B = \sqrt{A}$$

Taran Lynn

2014*

Contents

*Exact date lost due accidental to LaTeX recompilation.

1 Abstract

This paper details methods for finding the nth roots of matrices and their relation to linear transformations. The nth primary root of matrix A is matrix B that has all positive eigenvalues and meets the condition that $B^2 = A$. The main method of computation is through the diagonalization of matrices. Other methods include using the Jordan normal form and the Babylonian method. nth roots of matrices can also be used to find linear transformations whose repeated application to a vector results in a desired linear transformation. This use for finding linear transformations may have potential applications in computer graphics and other areas of science.

2 Introduction to Matrix Roots

The multiplication of matrices is quite often taught near the beginning of linear algebra courses. After learning how to multiply row and column vectors it is quite easy to multiply two matrices together. Each entry in the product of two matrices AB is found by multiplying the corresponding row of the first matrix A by the corresponding column of the second matrix B. We can expand this definition and find the square of a square matrix A as that matrix times itself, that is $A^2 = AA$. We may then wounder if it is possible to reverse this process. Given a matrix A is it possible to find a matrix B such that $B^2 = A$? It is logical to interpret such matrices as the square roots of other matrices and define them thusly.

Definition 1: The Square Root of a Matrix

Let A be an $n \times n$ matrix and B be an $n \times n$ matrix such that $B^2 = A$. B is called **a square root** of A. Also let $A^{1/2}$ be the principal square root of A (if it exists). For some matrices this would be a square root with non-negative eigenvalues.

You may notice my use of the term "the principal square root". For real non-negative numbers this is the positive square root. For a negative real number x it is $i\sqrt{|x|}$. For matrices this number becomes much harder to define, as their is no (obvious) guarantee that square roots of a matrix with non-negative eigenvalues exist. We can also extend the idea of square roots to nth roots.

Definition 2: The nth Root of a Matrix

Let A be an $n \times n$ matrix and B be an $n \times n$ matrix such that $B^n = A$. B is called **an nth root** of A. Also let $A^{1/n}$ be the principal nth root of A (if it exists). For some matrices this would be an nth root with non-negative eigenvalues, where such a matrix would exist.

We can use these ideas of roots to introduce rational powers to matrices. For example let A be an $n \times n$ matrix, then

$$a, b \in \mathbb{Z}$$
$$a, b > 0$$
$$A^{a/b} = \left(A^{1/b}\right)^a$$

$$A^{-a/b} = \left(\left(A^{1/b}\right)^a\right)^{-1}$$

These powers of square matrices are thus relatively simple to find after finding the matrix's roots.

3 Roots of Diagonal Matrices

The roots of matrices can be rather complicated to find at first glance, so lets look at the simple example of diagonal matrices. To understand the patterns behind the roots of diagonal matrices we should first understand their whole powers. Observe the following solutions.

$$\begin{aligned}
D &= \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix} \\
D^2 &= \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix} \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix} \\
&= \begin{bmatrix} \left(d_1^2 + 0d_1 + \cdots + 0d_1\right) & 0 & \cdots & 0 \\ 0 & \left(0d_2 + d_2^2 + \cdots + 0d_2\right) & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \left(0d_3 + 0d_3 + \cdots + d_n^2\right) \end{bmatrix} \\
&= \begin{bmatrix} d_1^2 & 0 & \cdots & 0 \\ 0 & d_2^2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & d_n^2 \end{bmatrix} \\
D^{a+1} &= \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix} \begin{bmatrix} d_1^a & 0 & \cdots & 0 \\ 0 & d_2^a & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & d_n^a \end{bmatrix}
\end{aligned}$$

$$= \begin{bmatrix} \left(d_1^{a+1} + 0d_1 + \cdots + 0d_1\right) & 0 & \cdots & 0 \\ 0 & \left(0d_2 + d_2^{a+1} + \cdots + 0d_2\right) & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \left(0d_n + 0d_n + \cdots + d_n^{a+1}\right) \end{bmatrix}$$

$$= \begin{bmatrix} d_1^{a+1} & 0 & \cdots & 0 \\ 0 & d_2^{a+1} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & d_n^{a+1} \end{bmatrix}$$

Therefore we can find that the following theorem is true

Theorem 1: Roots of Diagonal Matrices

Let D and E be $n \times n$ diagonal matrices. Every entry of E is an nth root of a corresponding entry in D if and only if E is an nth root of D. That is

$$\forall (0 < i \leq n) : (e_i^a = d_i) \iff (E^a = D)$$

The following definition is also useful

Definition 3: Primary Root of a Diagonal Matrix

The primary nth root of a diagonal matrix D is a matrix E whose entries are the primary nth roots of D. That is

$$D = \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix}$$

$$D^{1/a} = \begin{bmatrix} \sqrt[a]{d_1} & 0 & \cdots & 0 \\ 0 & \sqrt[a]{d_2} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \sqrt[a]{d_n} \end{bmatrix}$$

4 Finding Roots of Diagonalizable Matrices

Since we know how to find roots of diagonal matrices the next place to look would be diagonalizable matrices. Lets start by looking at the powers of diagonalizable matrices. Let D be a diagonal matrix.

$$
\begin{aligned}
A &= PDP^{-1} \\
A^2 &= (PDP^{-1})(PDP^{-1}) \\
&= PD(P^{-1}P)DP^{-1} \\
&= PDIDP^{-1} \\
&= PD^2P^{-1} \\
A^3 &= (PD^2P^{-1})(PDP^{-1}) \\
&= PD^2(P^{-1}P)DP^{-1} \\
&= PD^2IDP^{-1} \\
&= PD^3P^{-1} \\
&\vdots \\
A^n &= (PD^{n-1}P^{-1})(PDP^{-1}) \\
&= PD^{n-1}(P^{-1}P)DP^{-1} \\
&= PD^{n-1}IDP^{-1} \\
&= PD^nP^{-1}
\end{aligned}
$$

Based on this we could postulate that if $A = PDP^{-1}$, $B = PEP^{-1}$, and $E^n = D$, where D and E are diagonal matrices, then $B^n = A$. We can show this to be true.

$$
\begin{aligned}
B &= PEP^{-1} \\
B^n &= PE^nP^{-1} \\
&= PDP^{-1} \\
&= A
\end{aligned}
$$

We can now state a new theorem.

Theorem 2: Roots of Diagonalizable Matrices

If $A = PDP^{-1}$, $B = PEP^{-1}$, and $E^n = D$, where D and E are diagonal matrices, then $B^n = A$.

We can also add a new definition.

Definition 4: Principal Roots of Diagonalizable Matrices

If $A = PDP^{-1}$, where D is a diagonal matrix, then the principal nth root of A is $A^{1/n} = PD^{1/n}P^{-1}$.

5 Other Methods for Computing The Roots of Matrices

5.1 Calculation by Jordan Normal Form

The Jordan normal form (a.k.a. Jordan canonical form) of a matrix A is some special matrix J that is similar to A, that is $A = PJP^{-1}$ for some P. If A a is an $n \times n$ matrix with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$, then a Jordan matrix similar to it is an $n \times n$ matrix

$$J = \begin{bmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_s \end{bmatrix}$$

Where J_i is an $m_i \times m_i$ matrix such that

$$J_i = \begin{bmatrix} \lambda_i & 1 & \cdots & 0 \\ 0 & \lambda_i & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \\ 0 & 0 & \cdots & \lambda_i \end{bmatrix}$$

Every m_i for every J_i in J must also add up to n. That is $m_1 + m_2 + \cdots + m_s = n$. Here is an example

$$\begin{aligned} A &= \begin{bmatrix} 1 & 1 \\ -1 & 3 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix} \\ J &= \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix} \end{aligned}$$

We can use a similar approach to the one we used with diagonalization to find the following theorem to be true

Theorem 3

If an $n \times n$ matrix A is similar to E, an $n \times n$ matrix B is similar to F, F is an nth root of E, and A and B have the same similarity matrix P, then B is an nth root of A. A more direct notation for this is

$$\left(A = PEP^{-1} \wedge B = PFP^{-1} \wedge F^n = E\right) \implies (B^n = A)$$

Jordan forms are a special case of this, where if $A = PJP^{-1}$ and $B = PJ^{1/n}P^{-1}$, then $B^n = A$.

We can also find that an $n \times n$ Jordan matrix J taken to the ath power is equal to its Jordan blocks taken to the ath power. That is

$$J = \begin{bmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_s \end{bmatrix}$$
$$J^a = \begin{bmatrix} J_1^a & & & \\ & J_2^a & & \\ & & \ddots & \\ & & & J_s^a \end{bmatrix}$$

and we have our next theorem.

Theorem 4

Let J is an $n \times n$ Jordan matrix with Jordan blocks $J_1, \ldots J_s$, and D is a diagonal block matrix with blocks $D_1, \ldots, D_s$ such that $D_i^n = J_i$ for all $1 \leq i \leq s$. It must then be true that $D^n = J$. D is thus an nth root of J.

Coupled with the fact that any matrix is similar to some Jordan matrix this makes it easy to find roots of non-diagonalizable matrices. To see why this is true we can square a matrix

$$\begin{aligned} A &= \begin{bmatrix} a & b \\ 0 & a \end{bmatrix} \\ A^2 &= \begin{bmatrix} \left(a^2+0\right) & (ab+ab) \\ 0 & \left(a^2+0\right) \end{bmatrix} \\ &= \begin{bmatrix} a^2 & 2ab \\ 0 & a^2 \end{bmatrix} \end{aligned}$$

So if we take

$$\begin{aligned} A^2 &= B \\ B &= \begin{bmatrix} c & 1 \\ 0 & c \end{bmatrix} \end{aligned}$$

then

$$\begin{aligned} a^2 &= c \\ a &= \sqrt{c} \\ 2ab &= 1 \\ b &= \frac{1}{2a} \\ &= \frac{1}{2\sqrt{c}} \\ A &= \begin{bmatrix} \sqrt{c} & 1/\left(2\sqrt{c}\right) \\ 0 & \sqrt{c} \end{bmatrix} \end{aligned}$$

We can now find square roots of any square matrix with eigenvalues that have a maximum duplicity of 2, since we can find the square root of its Jordan matrix. Note that the previous case only works for matrices with one unique eigenvalue. This doesn't matter, however, since matrices with two unique eigenvalues can already be diagonalized. Similar methods can be used for higher order matrices and roots.

5.2 Babylonian Method

The Babylonian method is an iterative method used to calculate the square roots of real numbers, and is given as

$$x \in \mathbb{R}$$

$$\begin{aligned}
x &\geq 0 \\
a_0 &\approx x \\
a_{n+1} &= \frac{1}{2}\left(a_n + \frac{x}{a_n}\right) \\
\lim_{n\to\infty} a_n &= \sqrt{x}
\end{aligned}$$

A similar approach for matrices is

$$\begin{aligned}
X &\in \mathbb{R}^n \\
A_0 &= I_n \\
A_{n+1} &= \frac{1}{2}\left(A_n + \frac{X}{A_n}\right) \\
\lim_{n\to\infty} A_n &= X^{1/2}
\end{aligned}$$

This method is very volatile and can diverge from the square root under many circumstances. These circumstances are very hard to predict. However, this method is also very fast since it requires only a few operations each iteration.

6 Matrix Roots and Linear Transformations

One application of the roots of matrices is in conjunction with linear transformations. Imagine this scenario, a programmer has a linear transformation applied to a figure in a video game. The programmer might then want to animate this transformation. To do this he/she would need to apply several linear transformations to the figure which culminate in the final transformation. To have a smooth animation we would also want the sub-transformations to be the same. We can simulate this with two linear transformations $T : \mathbb{R}^2 \mapsto \mathbb{R}^2$ and $S : \mathbb{R}^2 \mapsto \mathbb{R}^2$, with $T(\vec{x}) = A\vec{x}$ and $S(\vec{x}) = B\vec{x}$. Now lets say we wanted $\vec{x}$ to map to $T(\vec{x})$ when S was applied to it 5 times. We would find that

$$\begin{aligned}
S\left(S\left(S\left(S\left(S\left(\vec{x}\right)\right)\right)\right)\right) &= T(\vec{x}) \\
B\left(B\left(B\left(B\left(B\left(\vec{x}\right)\right)\right)\right)\right) &= A\vec{x} \\
B^5\vec{x} &= A\vec{x}
\end{aligned}$$

$$B^5 = A$$

Thus B is a 5th root of A.
In general we find that if

$$\begin{aligned} T &: \mathbb{R}^n \mapsto \mathbb{R}^n \\ T(\vec{x}) &= A\vec{x} \\ S &: \mathbb{R}^n \mapsto \mathbb{R}^n \\ S(\vec{x}) &= A^{1/n}\vec{x} \end{aligned}$$

then

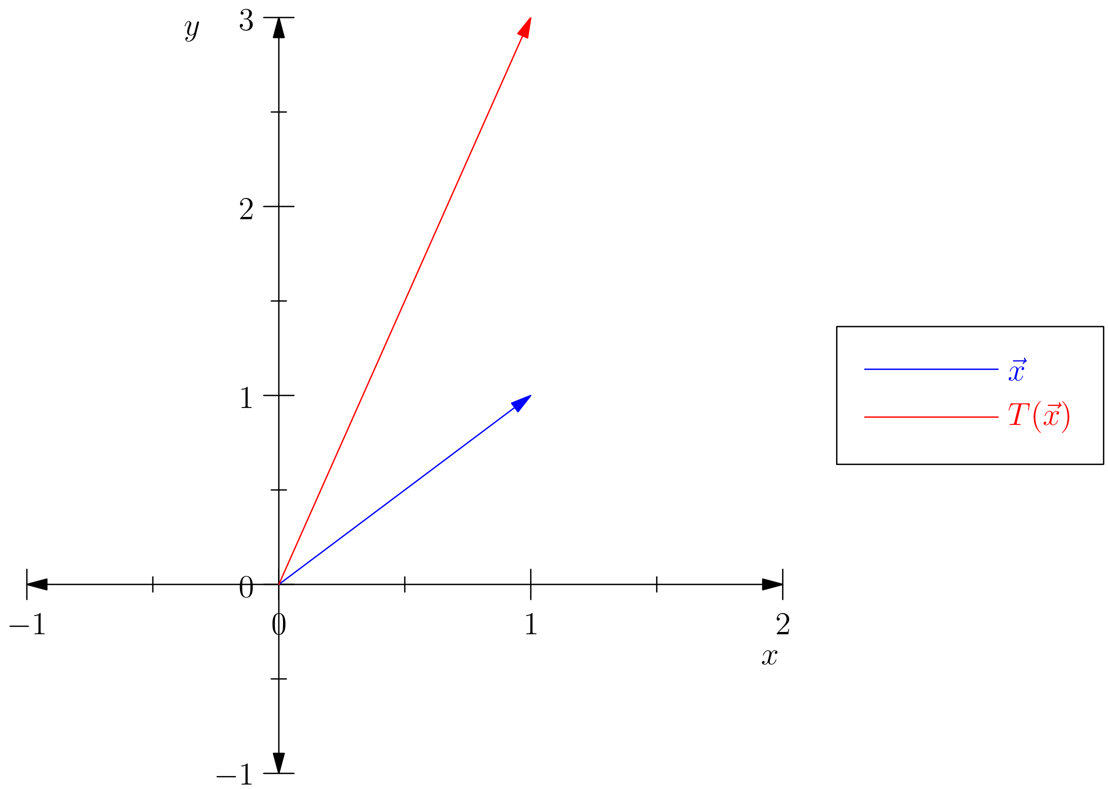
$$\begin{aligned} S\left(S\left(\cdots S\left(\vec{x}\right)\right)\right) &= A^{1/n}\left(A^{1/n}\left(\cdots A^{1/n}\left(\vec{x}\right)\right)\right) \\ &= (A^{1/n})^n\vec{x} \\ &= A\vec{x} \\ &= T(\vec{x}) \end{aligned}$$

And so we can find a transformation S that when applied to $\vec{x}$ n times results in $T(\vec{x})$.

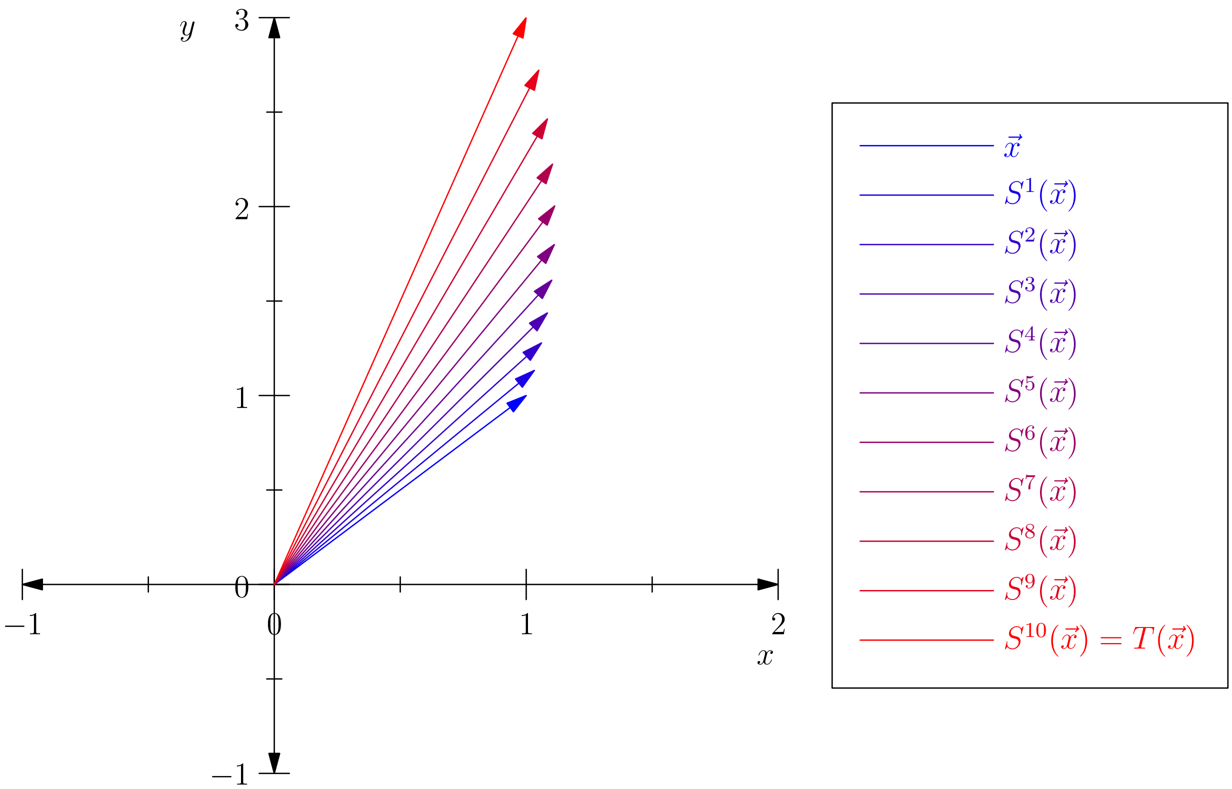
6.0.1 Example

It would probably be helpful to demonstrate this effect in action.

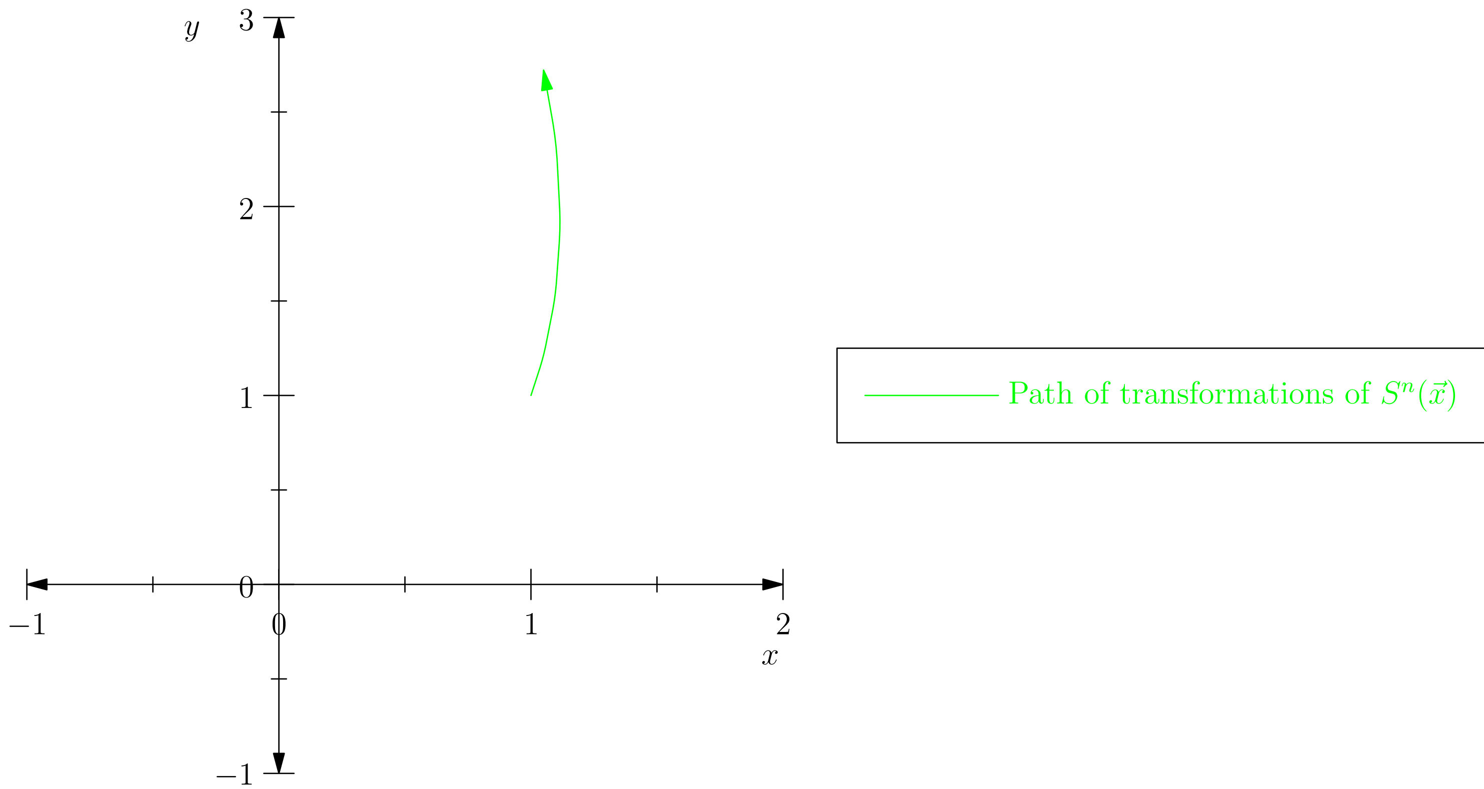
$$\begin{aligned} T &: \mathbb{R}^2 \mapsto \mathbb{R}^2 \\ T(\vec{x}) &= A\vec{x} \\ A &= \begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 1 \\ i & -i \end{bmatrix} \begin{bmatrix} 2-i & 0 \\ 0 & 2+i \end{bmatrix} \begin{bmatrix} 0.5 & -0.5i \\ 0.5 & 0.5i \end{bmatrix} \end{aligned}$$



$$
\begin{aligned}
S &: \mathbb{R}^2 \mapsto \mathbb{R}^2 \\
A(\vec{x}) &= A^{1/10}\vec{x} \\
A^{1/10} &= \begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 1 \\ i & -i \end{bmatrix} \begin{bmatrix} 2-i & 0 \\ 0 & 2+i \end{bmatrix}^{1/10} \begin{bmatrix} 0.5 & -0.5i \\ 0.5 & 0.5i \end{bmatrix} \\
&= \begin{bmatrix} 1 & 1 \\ i & -i \end{bmatrix} \begin{bmatrix} 1.08263 - 0.05023i & 0 \\ 0 & 1.08263 + 0.05023i \end{bmatrix}^{1/10} \begin{bmatrix} 0.5 & -0.5i \\ 0.5 & 0.5i \end{bmatrix} \\
&= \begin{bmatrix} 1.082634 & -0.050232 \\ 0.050232 & 1.082634 \end{bmatrix}
\end{aligned}
$$



As you can see the series of linear transformations S^n forms a smooth path for the vector $\vec{x}$ to travel along, giving rise to potential uses in graphics computing and physics. We can see this path by connecting the intermediate vectors.



7 Conclusion

As we can see it is rather simple to calculate the roots of many matrices. This includes matrices that are diagonalizable or converge under the Babylonian method. For other matrices we would have to use the Jordan normal form or other methods that are beyond the the authors current level of mathematics. Using these methods we can reverse most formulas that use the square of a matrix. We can also estimate the path that a vector takes when a linear transformation is applied to it. The topic of matrix roots is not usually taught in first year linear algebra classes, and so I hope that this paper will prove useful to be a useful guide to the roots of matrices.

8 Bibliography

References

[1] Jean Gallier. *Logarithms and Square Roots of Real Matrices Existence, Uniqueness, and Applications in Medical Imaging*. report. University of Pennsylvania, Oct. 21, 2013.

[2] Eric Jarman. *Nth Roots of Matrices*. thesis. University of Central Missouri, June 2012.

[3] Sam Northshield. *Square Roots of 2x2 Matrices*. report. Plattsburgh State Univ.