Matrix Square Roots of Polynomials

Kosmas Diveris * Trung Vu †

October 6, 2017

Abstract
In this article we consider matrix factorizations of a polynomial where the two matrices appearing in the factorization are the same, which we call “matrix square roots.” The main result is that any polynomial in \( \mathbb{R}[x_1, \ldots, x_n] \) admits a matrix square root. Our proof is constructive and provides an algorithm for constructing these matrices.

1 Introduction

In this article we consider matrix factorizations of a polynomial \( f \) in the ring \( S = \mathbb{R}[x_1, \ldots, x_n] \). These were first introduced by David Eisenbud in [2] to study modules over the quotient ring \( S/(f) \). These are known as hypersurface rings as they are coordinate rings of the zero-locus of the polynomial \( f \), which is a hypersurface in \( \mathbb{R}^n \).

Definition 1 [2] An \( n \times n \) matrix factorization of a polynomial \( f \in S \) is a pair of \( n \times n \) matrices \((A, B)\) such that \( AB = fI_n \), where \( I_n \) is the \( n \times n \) identity matrix.

A 1 \( \times \) 1 matrix factorization, \( [a] \cdot [b] = [f] \), is simply a factorization of \( f \) into a product of polynomials, so we see that matrix factorizations generalize the classical notion of factorization. In [2], it is shown that all polynomials, even irreducible polynomials, admit \( n \times n \) matrix factorizations for some \( n \).

Example 2 The polynomial \( f = x_1^2 + x_2^2 \) is irreducible in \( \mathbb{R}[x_1, x_2] \), but it has a 2 \( \times \) 2 matrix factorization:

\[
\begin{bmatrix}
    x_1 & x_2 \\
    x_2 & -x_1
\end{bmatrix}
\begin{bmatrix}
    x_1 & x_2 \\
    x_2 & -x_1
\end{bmatrix}
= \begin{bmatrix}
    x_1^2 + x_2^2 & 0 \\
    0 & x_1^2 + x_2^2
\end{bmatrix}.
\]

Note that in this example, the two matrices used to factor \( f \) were actually the same. This phenomenon motivates the following definition:

\*St. Olaf College
†St. Olaf College
Definition 3 An $n \times n$ matrix square root of a polynomial $f \in S$ is an $n \times n$ matrix $A$ such that $A^2 = fI_n$.

The main result of this paper, Theorem 7, is that any polynomial $f \in S$ admits a matrix square root. Our proof of this fact uses only techniques from elementary linear algebra. It is also constructive and provides an algorithm for building matrix square roots of any polynomial $f$. If $n$ is any natural number and the polynomial $f$ is expressed as the sum on $k$ monomials, then we are able to build matrix square roots of size $2^k n \times 2^k n$. If, however, one of the summands of $f$ is a perfect square, then we are able to build matrix square roots of half of this size.

2 Main Results

We open with a some preliminary observations about products of special block matrices that allow us to build matrix square roots. The necessary background on block matrices can be found in most introductory linear algebra textbooks. We also make use of the fact that the matrices appearing in a matrix factorization commute with each other, i.e. if $AB = fI_n$, then $BA = fI_n$, see Proposition 4 of [1].

Proposition 4 Assume $A$ is a $n \times n$ square root of $f$ and $(B, C)$ is a $n \times n$ matrix factorization of $g$. If $A$ commutes with both $B$ and $C$, then

$$
\begin{bmatrix}
A & B \\
C & -A
\end{bmatrix}
$$

is a $2n \times 2n$ square root of $f + g$.

Proof: The hypotheses give that $A^2 = fI_n$, $BC = gI_n = CB$, $AB - BA = 0$, and $CA - AC = 0$. Direct computation then gives the desired result:

$$
\begin{bmatrix}
A & B \\
C & -A
\end{bmatrix}^2 = \begin{bmatrix}
A^2 + BC & AB - BA \\
CA - AC & CB + A^2
\end{bmatrix} = \begin{bmatrix}
(f + g)I_n & 0 \\
0 & (f + g)I_n
\end{bmatrix}.
$$

QED

We highlight a special case of this proposition in the following corollary, which is key ingredient in our construction of matrix square roots.

Corollary 5 Assume that $f, g$, and $h$ are polynomials in $S$ and that $A$ is a $n \times n$ square root of $f$. Then

$$
\begin{bmatrix}
A & gI_n \\
hI_n & -A
\end{bmatrix}
$$

is a $2n \times 2n$ square root of $f + gh$. 

2
Proof: Since the matrices $B = gI_n$ and $C = hI_n$ commute with all $n \times n$ matrices, this follows immediately from the previous proposition. QED

Any polynomial in $f \in S$ can be expressed in the form $f = g_1h_1 + \cdots + g_kh_k$ for some $g_i, h_i \in S$. The next result provides a means to construct a matrix square root of the polynomial $f$ using the polynomials $g_i$ and $h_i$. Of course, such an expression for $f$ is not unique and different representations of $f$ will produce different matrix square roots.

**Theorem 6** Let $f_k = g_1h_1 + \cdots + g_kh_k$ be a polynomial in $S$ and $n \in \mathbb{N}$. Then $f_k$ has a $2^k n \times 2^k n$ matrix square root whose entries are the polynomials 0, $g_i$, and $h_i$ where $1 \leq i \leq k$.

**Proof:** We proceed by induction on $k$, the number of summands of $f_k$. In the case when $k = 1$, observe that the $2^n \times 2^n$ matrix

$$A_1 = \begin{bmatrix} 0 & g_1I_n \\ h_1I_n & 0 \end{bmatrix}$$

is a square root of the polynomial $f_1 = g_1h_1$. This follows from Corollary 5, by taking $A = 0$. Note that the size of $A_1$ is $2n \times 2n$, and its only entries are 0, $g_1$, and $h_1$.

Assume that for some $j \geq 1$ we have constructed a $2^j n \times 2^j n$ matrix square root, $A_j$, of the polynomial $f_j$, where whose entries consist solely of the polynomials 0, $g_i$, and $h_i$ where $1 \leq i \leq j$. Then the $2^{j+1} n \times 2^{j+1} n$ matrix

$$A_{j+1} = \begin{bmatrix} A_j \\ \frac{g_{j+1}I_{2^n}}{h_{j+1}I_{2^n}} & -A_j \end{bmatrix}$$

is a square root matrix of the polynomial $f_{j+1}$, again by Corollary 5. The only new entries in $A_{j+1}$ that were not also entries of $A_j$ are $g_{j+1}$ and $h_{j+1}$.

By induction, after $k$ steps we will obtain a square root matrix $A_k$ of $f_k$ whose size is $2^k n \times 2^k n$ and whose only entries are 0, $g_i$, and $h_i$ where $1 \leq i \leq k$, as claimed. QED

The polynomial $f = x_1^2 + x_2^2$ from Example 2 is the sum of two terms. The matrix factorization given in this example has size $2 \times 2$. The previous Theorem, however, would only generate factorizations of $f$ of size at least $4 \times 4$. In this case, we can improve upon the Theorem and build a smaller matrix square root because one of the summands of $f$ actually has a $(1 \times 1)$ square root in $S$. In fact, any time the polynomial $f_k$ has a summand that is a perfect square, then a slight modification of the above proof will yield a matrix square root of half the size.

**Theorem 7** Let $f_k = g_1^2 + g_2h_2 + \cdots + g_kh_k$ be a polynomial in $S$ and $n \in \mathbb{N}$. Then $f_k$ has a $2^{k-1} n \times 2^{k-1} n$ matrix square root whose entries are the polynomials 0, $g_i$, and $h_i$ where $1 \leq i \leq k$.

**Proof:** The induction step in the proof proceeds exactly the same as that of Theorem 7. Here we treat only the base case, which is where the reduction in
size occurs. This time, the \( n \times n \) matrix \( A_1 = g_1I_n \) is easily seen to be a square root of the polynomial \( f_1 = g_1^2 \). QED

The reader familiar with ring theory may note that the only property of the polynomial ring \( S \) that we have used in our arguments so far is that it is a commutative ring. Thus our methods actually allow one to construct a matrix square root of any element in any commutative ring.

3 An Application.

A monomial in \( S = \mathbb{R}[x_1, ..., x_n] \) is a polynomial of the form \( m = cx_1^{a_1}x_2^{a_2} \cdots x_n^{a_n} \) where \( c \in \mathbb{R} \) and each \( a_i \) is a non-negative integer. The degree of \( m \) is defined to be \( \deg m = \sum a_i \). We say that \( m \) is constant if its degree is zero, and linear if its degree is one.

Every polynomial in \( S \) can be expressed as a sum of monomials. We denote by \( \mathfrak{m} \) the subset of \( S \) consisting of all polynomials with no constant summand. This is actually a maximal ideal of \( S \). We denote by \( \mathfrak{m}^2 \) the subset of \( \mathfrak{m} \) consisting of polynomials having no linear terms and no constant terms. This is the product of the ideal \( \mathfrak{m} \) with itself. As such, its elements consist of sums of products of elements in \( \mathfrak{m} \).

**Corollary 8** If \( f \in \mathfrak{m}^2 \), then \( f \) has a square root matrix whose entries are all elements of \( \mathfrak{m} \).

**Proof:** If \( f \in \mathfrak{m}^2 \), then \( f \) may be expressed as a sum of products of elements of \( \mathfrak{m} \). That is, \( f \) can be expressed in the form

\[
    f = g_1h_1 + \cdots + g_kh_k,
\]

where each \( g_i \) and \( h_i \) is in \( \mathfrak{m} \). Applying Theorem 7 to \( f \) yields a matrix square root of \( f \) whose entries are \( g_i \) and \( h_i \), which are in \( \mathfrak{m} \). QED

This shows that when \( f \) is non-linear, i.e. when \( f \in \mathfrak{m}^2 \), the matrix square root constructed in the proof of Theorem 7 is reduced, in the sense of [2]. Thus it gives rise to a maximal Cohen-Macaulay module, without free summands, over the ring \( S/(f) \) whose free resolution is periodic of period 1. It would be interesting to know when this matrix factorization is indecomposable, again in the sense of [2], and thus corresponds to an indecomposable module with this property.

References
