MATRIX SQUARE ROOT OF POLYNOMIALS

KOSMAS DIVERIS AND HIEU TRUNG VU

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

CONTENTS

1. INTRODUCTION

In this article we consider matrix factorizations of a polynomial f in the ring

$$S = \mathbb{R}[x_1, \cdots, x_n]$$

These were first introduced b David Eisenbud in [2] to study modules over the quotient ring $S/(f)$. These are known as the hypersurface rings as they are coordinate rings of the zero-locus of the polynomial f, which is a hypersurface in $\mathbb{R}^n$, denoted by $Z(f)$.

Definition 1.1 ([2]). *An $n \times n$ matrix factorization of a polynomial $f \in S$ is a pair of $n \times n$ matrices $(A, B) \in M_n(S)$ so that $AB = f I_n$, where I_n is the $n \times n$ identity matrix*

Note that the entries of the matrices A and B in the definition are polynomials from S. A 1×1 matrix factorization, $[g] \ldots [h] = [f]$ ios 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, admits $n \times n$ matrix factorizations for some n.

Example 1.2. *The polynomial $f = x_1^2 + x_2^2$ is irreducible in $\mathbb{R}[x_1, x_2]$, but it has a 2×2 matrix factorization:*

$$\begin{bmatrix} x_1 & x_2 \\ x_2 & -x_1 \end{bmatrix} \begin{bmatrix} x_1 & x_2 \\ x_2 & -x_1 \end{bmatrix} = \begin{bmatrix} x_1^2 + x_2^2 & 0 \\ 0 & x_1^2 + x_2^2 \end{bmatrix} \tag{1.1}$$

Algebraic geometry provides a variety of tools to connect geometric properties of a hypersurface $Z(f)$ to its coordinates ring $S/(f)$, and modules over this ring. In [2], Eisenbud takes these connections a step further by establishing a correspondence between certain $S/(f)$-modules and matrix factorizations of f. For the reader with some background in commutative algebra, we recommend the texts [4] and [4] to see the ultility of this approach to examining singularities of hypersurfaces illustrated. One of the most powerful aspects of this approach is that it allows one to work with matrices, instead of modules, and thus to bring in more techniques from linear algebra with fewer prerequisites from abstract algebra.

We would like to draw attention to the fact that in Example 1.2, the two matrices used to factor f were actually the same. This phenomenon motivates the following definition:

Definition 1.3 (Matrix Square Root)**.** *An $n \times n$ matrix square root of a polynomial $f \in S$ is an $n \times n$ matrix $A \in M_n(S)$ such that $A^2 = f I_n$.*

In addition to their connections with algebraic geometry, matrix square roots have also arisen in the context of orthogonal designs, we refer the reader to [3] for these connections. While polynomials in S rarely have square roots in S, all polynomials admit matrix square roots, as the next example illustrates.

Example 1.4. *If $f \in S$ is any polynomial, then f has a 2×2 matrix square root:*

$$\begin{bmatrix} 0 & f \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & f \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} f & 0 \\ 0 & f \end{bmatrix}$$

This matrix square root, however, is not particularly "interesting" in that it does not shed any new light on the polynomial f (or the hypersurface $Z(f)$). Fortunately, this needs not be the only matrix square root of a polynomial. Indeed, while square roots of polynomials in S are quite rare, and when they do exist they are often rather unique, a polynomial will often admit many different matrix square square roots. This stems, in part, from the fact that S is a unique factorization domain, whereas $M_n(S)$ is not. The next example is provided to show two distinct 4×4 matrix square roots of the sample polynomial.

Example 1.5. *The following matrices are both 4×4 square roots of $x^3y + xy^3$:*

$$\begin{bmatrix} 0 & 0 & x^2+y^2 & 0 \\ 0 & 0 & 0 & x^2+y^2 \\ xy & 0 & 0 & 0 \\ 0 & xy & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & x^3 & x & 0 \\ y & 0 & 0 & x \\ y^3 & 0 & 0 & x \\ 0 & y^3 & -y & 0 \end{bmatrix}$$

The main result of this paper, theorem 2.3, is that any polynomial $f \in S$ admits "interesting" matrix square roots. Our proof of this fact uses only techniques from elementary algebra. It is also constructive and provides an algorithm for building matrix square roots of any polynomial f. The size of the matrix square roots we construct is also easily determined from the polynomial. If n is any natural number and the polynomial f is expressed as the sum of 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 summads of f is a perfect square, then we are able to build matrix square roots of half of this size.

2. Main Result

We open with 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 the matrix factorization commute with each other,i.e. if $AB = fI_n$, then $BA = fI_n$. This is not assumed in the defintion of a matrix factorization, but a proof of this can be found, for example, in Proposition 4 of [1].

Proposition 2.1. *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 the equations

$$A^2 = fI_n, BC = gI_n = CB, AB - BA = 0, CA - AC = 0$$

Direct computation then gives the desired result:

$$\left[\begin{array}{c|c} A & B \\ \hline C & -A \end{array}\right]^2 = \left[\begin{array}{c|c} A^2 + BC & AB - BA \\ \hline CA - AC & BC + A^2 \end{array}\right] = \left[\begin{array}{c|c} (f+g)I_n & 0 \\ \hline 0 & (f+g)I_n \end{array}\right] \tag{2.1}$$

□

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

Corollary 2.2. *Assume that f, g_1, and h_1 are polynomials in S and that A is a $n \times n$ square root of f. Then:*

$$\left[\begin{array}{c|c} A & gI_n \\ \hline hI_n & -A \end{array}\right]$$

is a $2n \times 2n$ square root of $f + g_1h_1$.

Proof. Since the matrices $B = g_1I_n$ and $C = h_1I_n$ commute with all $n \times n$ matrices, and $BG = g_1h_1I_n$, this follows from the previous proposition. □

Any polynomial in $f \in S$ can be expressed in the form $f = g_1h_1 + \cdots + g_nh_n$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 of f will produce different matrix square roots.

Theorem 2.3. *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 $2^kn \times 2^kn$ matrix square root whose enries are the polynomials $0, g_i$, and h_i where $1 \leq i < k$*

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

$$A_1 = \left[\begin{array}{c|c} 0 & g_1I_n \\ \hline h_1I_n & 0 \end{array}\right]$$

is a square root of the polynomial $f_1 = g_1h_1$. This follows from Corollary 2.2, by taking $A = 0$. Note that the size of A_1 is $2n \times 2n$, and its only entries are $0, g_1, h_1$. Assume that for some $j \geq 1$

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

$$A_{j+1} = \left[\begin{array}{c|c} A_j & g_{j+1} I_{2^j n} \\ \hline h_{j+1} I_{2^j n} & -A_j \end{array}\right]$$

is a square root of the polynomial f_{j+1}, again by Corollary 2.2. 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 obtain a square root matrix A_k of f_k whose size is $2^k n \times 2^k n$ and whose entries are $0, g_i$, and h_i where $1 \leq i \leq k$, as claimed. □

The polynomial $x_1^2 + x_2^2$ from example 1.2 is the sum of two terms. The matrix factorization given in this example has size 2×2. Theorem 2.3, however, would only generate factorizations of f of size at least 4×4. In this case, we can improve and build a smaller matrix square root because one of the summands of f actually has a 1×1 square roots 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 2.4. *Let $f_k = g_1^2 + g_2 h_2 + \cdots + g_k h_k$ 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 polynomial 0, g_i and h_i where $1 \leq i \leq k$*

Proof. The induction step proceeds exactly the same as in the proof of theorem 2.3. Here we treat only the base case, which is where the reduction in size occurs. This time, the $n \times n$ matrix $A_1 = g_1 I_n$ is easily seen to be a square root of the polynomial $f_1 = g_1^2$. □

Remark 2.5. *The readers 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. "INTERESTING" MATRIX SQUARE ROOTS

In the introduction we claimed that the matrix square root given in example 1.5 was not "interesting". In fact, this is called *trivial factorization* in [2]. While is it evident that this factorization does not shed new light on f, the sense in which it is trivial can be made precise. Under Eisenbud's correspondence between matrix factorization of f and $S/(f)$-modules, this factorization corresponds to the free $S/(f)$-module $S/(f) \oplus 0$. The free $S/(f)$-modules do not reveal much information about $Z(f)$, so the algebraic geometers are most often interested in the modules without the free summands. The matrix factorizations corresponding to these modules are called *reduced*, and these are the ones we consider most interesting. Essentially, a matrix factorization is reduced if one cannot perform invertible row and column operations, over $M_n(S)$, to transform the matrices into new pair having a block in the form I_k, fI_k for some k. These reduced factorizations are the ones we consider "interesting".

In an important case, there is a very efficient test to determine if a matrix factorization is reduced, which we now explain. Let $\mathfrak{m} \subset S$ denote the polynomials whose constant term is zero. This is actually a maximal ideal of S, $\mathfrak{m} = (x_1, \ldots, x_n)$, consisting of each polynomial $f \in S$ whose hypersurface, $Z(f)$, passes through the origin in $\mathbb{R}^n$. If $f \in \mathfrak{m}$, then a matrix factorization of f is reduced if and only if all of the entries appearing in the matrices are also in $\mathfrak{m}$, [2]. Indeed, if all of the entries of a matrix A lies in $\mathfrak{m}$, then any matrix obtained from A by performing row and columns operations must also contain entries in $\mathfrak{m}$, because it is an ideal. Thus, one cannot obtain 1 as an entry in such a matrix, which would be necessary if A were not reduced.

The next result refers to the product ideal of $\mathfrak{m}$ with itself, $\mathfrak{m}^2$. This consists of all polynomials in S whose constant term and linear term are both zero. These are the elements of $\mathfrak{m}$ whose corresponding hypersurface has a singularity at the origin (as its Jacobian matrix will be the zero matrix). The takeaway of the next corollary is that any polynomial corresponding to these hypersurfaces has a reduced matrix square root.

Corollary 3.1. *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 maybe 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 are in $\mathfrak{m}$. Applying theorem 2.3 to f yields a matrix square root of f whose entries are g_i and h_i, which are in $\mathfrak{m}$. □

REFERENCES

[1] Kosmas Diveris and David Crisler. *Matrix factorizations of sums of squares polynomials*. Pi Mu Epsilon J., 2016. Hadamard matrices, quadratic forms and algebras.

[2] David Eisenbud. Homological algebra on a complete intersection, with an application to group representations. *Trans. Amer. Math. Soc.*, 260(1):35–64, 1980.

[3] Jennifer Seberry. *Orthogonal designs*. Springer, Cham, updated edition, 2017. Hadamard matrices, quadratic forms and algebras.

[4] Yuji Yoshino. *Cohen-Macaulay modules over Cohen-Macaulay rings*, volume 146 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 1990.

DEPARTMENT OF MATHEMATICS, ST. OLAF COLLEGE, NORTHFIELD, MN

DEPARTMENT OF MATHEMATICS, ST. OLAF COLLEGE, NORTHFIELD, MN