Matrix Factorizations of Sums of Squares Polynomials

David Crisler *  Kosmas Diveris †

October 21, 2016

Abstract

We examine matrix factorizations of polynomials in the ring $S = \mathbb{R}[x_1, x_2, \ldots, x_m]$ using only techniques from elementary linear algebra. Special focus is placed on factorizations of sums of squares polynomials.

1 Introduction

In 1980 David Eisenbud introduced a new approach to factoring polynomials using matrices [3]. These matrix factorizations, defined below, generalize the classical notion of factorization of polynomials and allow one to factor polynomials that are irreducible in the classical sense. In this note we consider polynomials in the ring $S = \mathbb{R}[x_1, x_2, \ldots, x_m]$.

Definition 1 An $n \times n$ matrix factorization of a polynomial $f \in S$ is a pair of $n \times n$ matrices $A$ and $B$ such that $AB = f I_n$, where $I_n$ is the $n \times n$ identity matrix.

Note that a $1 \times 1$ matrix factorization $[f] = [g][h] = [gh]$ is simply a factorization of the polynomial $f$ in the classical sense. The next example illustrates how matrices can be used to factor an irreducible polynomial.

Example 1 The polynomial $f = x_1^2 + x_2^2$ is irreducible over $\mathbb{R}$, but the following pair of matrices give a $2 \times 2$ matrix factorization of $f$:

$$
\begin{bmatrix}
  x_1 & -x_2 \\
  x_2 & x_1
\end{bmatrix},
\begin{bmatrix}
  x_1 & x_2 \\
  -x_2 & x_1
\end{bmatrix}.
$$

Eisenbud proved that all polynomials in $S$ admit matrix factorizations. He discovered that the matrix factorizations of the polynomial $f$ are intimately related to homological properties of modules over the quotient ring $S/(f)$. These

*St. Olaf College
†St. Olaf College
quotient rings are known as \textit{hypersurface rings}, as they encode geometric properties of zero-locus of \( f \), \( Z(f) = \{ p \in \mathbb{R}^n \mid f(p) = 0 \} \), which is a hypersurface in \( \mathbb{R}^n \). For more background on the connection between matrix factorizations and algebraic geometry we refer the reader to [5], [4], [2], and [1]. In this paper, we describe some elementary constructions of matrix factorizations without resorting to the homological methods that Eisenbud introduced.

\section{Properties of matrix factorizations}

In this section we collect several properties of matrix factorizations that we will make use of later in the paper. As all of these results are known to the experts in the field, we only include proofs for those for which an appropriate reference is lacking. We continue to denote by \( S \) the polynomial ring \( \mathbb{R}[x_1, \ldots, x_n] \), and we let \( F \) denote its field of fractions, i.e., \( F = \mathbb{R}(x_1, \ldots, x_n) \), the field of rational functions over \( \mathbb{R} \).

Our first observation is that the determinant of a matrix that may appear in a matrix factorization of \( f \) must divide a power of \( f \).

\textbf{Lemma 2} If \( AB = fI_n \), then \( \det(A) \) divides \( f^n \). If, in addition, \( f \) is irreducible in \( S \), then \( \det(A) \) is a power of \( f \).

\textbf{Proof:} For the first claim, note that \( AB = fI_n \) implies

\[ \det(A) \det(B) = \det(AB) = \det(fI_n) = f^n. \]

The second claim now follows from the first because if \( f \) is irreducible, then the only divisors of \( f^n \) are powers of \( f \). QED

According to this lemma, we see that to build a \( n \times n \) matrix factorization of a non-zero polynomial \( f \in S \), a first step could be to construct a \( n \times n \) matrix \( A \) whose determinant divides \( f^n \). Once this is done, we can pass to \( F \), the fraction field of \( S \), to find a companion matrix \( B \) so that \( AB = fI_n \). Indeed, the lemma guarantees that \( \det(A) \) is nonzero because it divides \( f^n \), so over \( F \) the matrix \( A \) is invertible. Therefore equation \( AB = fI_n \) has a unique solution over \( F \), namely

\[ B = A^{-1}fI_n = \frac{f}{\det(A)} \text{adj}(A). \]

Since \( \text{adj}(A) \) is a matrix over \( S \), if \( \det(A) \mid f \), then this matrix \( B \) will also be a matrix over \( S \). It is possible, however, that this matrix \( B \) is not a matrix with entries in \( S \), and therefore the matrix \( A \) does not appear in any matrix factorizations of \( f \) over \( S \).

\textbf{Example 3} Let \( f = x^2 - y^2 \) and consider the \( 4 \times 4 \) matrix

\[ A = \begin{bmatrix} x & 2x & -y & 0 \\ y & x & 0 & -y \\ y & 0 & x & 2x \\ 0 & y & -y & x \end{bmatrix}. \]
Then \( \det(A) = f^2 \) (Lemma 5, below, provides a shortcut for verifying this) but the matrix \( B = fA^{-1} \) contains rational functions that are not polynomials.

The next observation is that the matrices appearing in a matrix factorization commute with each other.

**Proposition 4** If \( 0 \neq f \in S \) and \( A, B \) are \( n \times n \) matrices so that \( AB = fI_n \), then \( BA = fI_n \).

**Proof:** Equation (1) gives \( B = A^{-1}fI_n \). Since the matrix \( fI_n \) commutes with all \( n \times n \) matrices over \( F \), we have

\[
BA = (A^{-1}fI_n)A = (fI_nA^{-1})A = fI_n(A^{-1}A) = fI_n,
\]
as claimed. QED

In all of the constructions of matrix factorizations in this paper we will make use of properties of block matrices. A block matrix is a partition of a matrix

\[
A = \begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix}
\]

where each \( A_i \) is a matrix, called a block of \( A \). Multiplication of block matrices behaves exactly like standard matrix multiplication. That is, if \( A_i \) and \( B_i \) are \( n \times n \) matrices for \( i = 1, 2, 3, 4 \), then:

\[
\begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix} \begin{bmatrix} B_1 & B_2 \\ B_3 & B_4 \end{bmatrix} = \begin{bmatrix} A_1B_1 + A_2B_3 & A_1B_2 + A_2B_4 \\ A_3B_1 + A_4B_3 & A_3B_2 + A_4B_4 \end{bmatrix}.
\]

Schur’s determinant formula, stated below, is useful for computing the determinant of a block matrix, provided there is some commutativity within the blocks. A proof, along with a thorough historical discussion of this result, is given in [6].

**Lemma 5** Assume \( A, B, C, D \) are all \( n \times n \) matrices with \( AC = CA \). Then

\[
\det \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \det(AD - CB).
\]

### 3 Standard Method

In this section we discuss a standard technique for factoring polynomials that dates to the 1980s when Knörrer exploited it to prove his celebrated Periodicity Theorem [4]. The following observation describes one situation in which it is possible to build matrix factorizations of sums of polynomials from factorizations of their summands.
**Proposition 6** For \( i = 1, 2 \), let \( A_i, B_i \) denote an \( n \times n \) matrix factorization of the polynomial \( f_i \in S \). In addition, assume that the matrices \( A_i \) and \( B_j \) commute when \( i \neq j \). Then the matrices

\[
\begin{bmatrix}
A_1 & -B_2 \\
A_2 & -A_2 \\
B_1 & A_1
\end{bmatrix}
\begin{bmatrix}
B_1 \\
B_2 \\
-A_1
\end{bmatrix}
\]

give a \( 2n \times 2n \) matrix factorization of \( f_1 + f_2 \).

**Proof:** According to the hypotheses and Proposition 4 we actually have \( A_i B_j = B_j A_i \) for all \( i, j \). So we see

\[
\begin{bmatrix}
A_1 & -B_2 \\
A_2 & -A_2 \\
B_1 & A_1
\end{bmatrix}
\begin{bmatrix}
B_1 \\
B_2 \\
-A_1
\end{bmatrix}
= \begin{bmatrix}
A_1 B_1 + B_2 A_2 & A_1 B_2 - B_2 A_1 \\
A_2 B_1 - B_1 A_2 & A_2 B_2 + B_1 A_1
\end{bmatrix}
\begin{bmatrix}
f_1 + f_2 I_n \\
0
\end{bmatrix}
= (f_1 + f_2) I_{2n},
\]

as claimed. QED

The following special case of this result is the basis for the main construction of this section.

**Corollary 7** If \( A, B \) is a \( n \times n \) matrix factorization of \( f \), then

\[
\begin{bmatrix}
A & -gI_n \\
hI_n & B
\end{bmatrix}
\begin{bmatrix}
B & gI_n \\
-hI_n & A
\end{bmatrix}
\]

is a \( 2n \times 2n \) matrix factorization of \( f + gh \).

**Proof:** The matrices \( gI_n \) and \( hI_n \) commute with all \( n \times n \) matrices, so the claim follows immediately from the previous proposition. QED

This corollary allows us to inductively construct matrix factorizations of polynomials of the form:

\[
f_k = g_1 h_1 + g_2 h_2 + \ldots + g_k h_k.
\]  \hspace{1cm} (2)

In the case \( k = 1 \), we have \( f_1 = gh \) and \([g][h] = [gh] = [f_1]\) is a \( 1 \times 1 \) factorization. Assume now that \( A \) and \( B \) are matrix factorizations of \( f_{k-1} \), that is \( AB = I f_{k-1} \). The corollary gives that the following matrices are a matrix factorization of \( f_k \):

\[
\begin{bmatrix}
A & -g_k I \\
h_k I & B
\end{bmatrix}
\begin{bmatrix}
B & g_k I \\
-h_k I & A
\end{bmatrix}
\]

We call this algorithm the standard method of constructing matrix factorizations. Since any polynomial can expressed as a sum of finitely many monomials, this can be used to produce matrix factorizations of any polynomial.
Although reliable and proven to work for any polynomial, this algorithm does have a downside: for every new $g_n h_n$ added to the polynomial, the factorizations must double in size. That is, to factor a polynomial $f_n$ in (2) that is the sum of $n$ monomials with this method, we obtain matrices of size $2^{n-1} \times 2^{n-1}$. As such, these factorizations can grow extremely large extremely quickly. As we will see in the next section, it is sometimes possible to produce matrix factorizations with smaller matrices.

## 4 Special Square Factorizations

In this section, we generate smaller factorizations of “sums of squares” polynomials, those of the form:

$$f_n = x_1^2 + x_2^2 + \cdots + x_n^2,$$

for $n \leq 8$. For $4 \leq n \leq 8$, the resulting factorizations have smaller matrices than one would obtain using the standard algorithm from the previous section.

The paper [2] is a thorough study of matrix factorizations over quadratic hypersurfaces and contains factorizations of the polynomials $f_n$ in (3). The authors first demonstrate that there is an equivalence of categories between matrix factorizations of $f_n$ and graded modules over a Clifford algebra associated to $f_n$. They then exploit this equivalence to generate matrix factorizations. This technique can be used to generate minimal matrix factorizations of the polynomials $f_n$ for all $n \geq 1$, see also [1], or [5]. In contrast, our approach is elementary and only relies on the results stated earlier in this paper.

The factorizations of $f_n$ our algorithm generates are given below when $n$ is a power of 2 and at most 8. For $n = 3, 5, 6, 7$, factorizations can be obtained from these factorization by setting $x_i = 0$ for all $i > n$.

1. $n = 1$: We have a $1 \times 1$ factorization given by the pair $A_1, A_1^T$, where $A_1 = [x_1]$.

2. $n = 2$: We have a $2 \times 2$ factorization given by the pair $A_2, A_2^T$, where:

$$A_2 = \begin{bmatrix} x_1 & -x_2 \\ x_2 & x_1 \end{bmatrix}.$$

3. $n = 4$: We have a $4 \times 4$ factorization given by the pair $A_4, A_4^T$, where:

$$A_4 = \begin{bmatrix} x_1 & -x_2 & x_3 & x_4 \\ x_2 & x_1 & -x_4 & x_3 \\ -x_3 & x_4 & x_1 & x_2 \\ -x_4 & -x_3 & -x_2 & x_1 \end{bmatrix}.$$
4. \( n = 8 \): We have an \( 8 \times 8 \) factorization given by the pair \( A_8, A_8^T \), where:

\[
A_8 = \begin{bmatrix}
    x_1 & -x_2 & x_3 & x_4 \\
    x_2 & x_1 & -x_4 & x_3 \\
    -x_3 & x_4 & x_1 & x_2 \\
    -x_4 & -x_3 & -x_2 & x_1 \\
    -x_5 & x_6 & -x_7 & -x_8 \\
    -x_6 & -x_5 & -x_8 & x_7 \\
    -x_7 & -x_8 & x_6 & -x_5 \\
    -x_8 & x_7 & -x_6 & -x_5 \\
\end{bmatrix}
\]

5 Explanation/Process

We now describe our algorithm for obtaining \( A_{2n} \), from \( A_n \), where \( n = 1, 2, \) or 4.

Algorithm.

1. We assume the matrix \( A_n \) has been constructed, and that the pair \( A_n, A_n^T \) is a matrix factorization of \( f_n \). For the initial case, when \( n = 1 \), we may take \( A_1 = [x_1] \).

2. Let \( B_n \) be the matrix obtained by increasing the indices in \( A_n \) by \( n \). For example,

\[
A_2 = \begin{bmatrix}
    x_1 & -x_2 \\
    x_2 & x_1 \\
\end{bmatrix}, \text{ and } B_2 = \begin{bmatrix}
    x_3 & -x_4 \\
    x_4 & x_3 \\
\end{bmatrix}.
\]

Note that \( B_2B_2^T = (x_2^2 + \cdots + x_{2n}^2)I_n \).

3. Denote by \( J_n \) the \( n \times n \) obtained by negating the top left entry of the identity matrix \( J_n \). For example,

\[
J_2 = \begin{bmatrix}
    -1 & 0 \\
    0 & 1 \\
\end{bmatrix}.
\]

Note that \( J_n^2 = I_n \), so \( J_n^{-1} = J_n \).

4. Now let \( C_n = J_nB_nJ_n \). Since \( J_n^{-1} = J_n \), \( C_n \) is the matrix obtained by conjugating \( B_n \) by \( J_n \). Note that this action on \( B_n \) has the effect of negating its first row and column (with the entry in the first row and first column being negated twice and thus retaining its original sign). For example,

\[
C_2 = \begin{bmatrix}
    x_3 & x_4 \\
    -x_4 & x_3 \\
\end{bmatrix}.
\]

Note that \( C_nC_n^T = (x_{n+1}^2 + \cdots + x_{2n}^2)I_n \).
5. Verify directly that the matrices $A_n$ and $C_n$ commute. Indeed, for $n = 2$:

\[
\begin{bmatrix}
  x_1 & -x_2 \\
  x_2 & x_1
\end{bmatrix}
\begin{bmatrix}
  x_3 & x_4 \\
  -x_4 & x_3
\end{bmatrix} =
\begin{bmatrix}
  x_1x_3 + x_2x_4 & x_1x_4 - x_2x_3 \\
  x_2x_3 - x_1x_4 & x_2x_4 + x_1x_3
\end{bmatrix},
\]

\[
\begin{bmatrix}
  x_3 & x_4 \\
  -x_4 & x_3
\end{bmatrix}
\begin{bmatrix}
  x_1 & -x_2 \\
  x_2 & x_1
\end{bmatrix} =
\begin{bmatrix}
  x_1x_3 + x_2x_4 & x_1x_4 - x_2x_3 \\
  x_2x_3 - x_1x_4 & x_2x_4 + x_1x_3
\end{bmatrix}.
\]

We emphasize than in the case $n = 8$, the matrices $A_8$ and $C_8$ obtained in the manner will not commute.

6. Lastly, we set $A_{2n} = \begin{bmatrix} A_n & C_n \\ -C_n & A_n \end{bmatrix}$. Then, if $n \leq 4$, we see

\[
A_{2n}A_{2n}^T = \begin{bmatrix} A_n & C_n \\ -C_n & A_n \end{bmatrix} \begin{bmatrix} A_n^T \\ C_n^T \end{bmatrix} = \begin{bmatrix} A_nA_n^T + C_nC_n^T & -A_nC_n + C_nA_n \\ C_nA_n^T + A_nC_n^T & C_nC_n^T + A_nA_n^T \end{bmatrix} = \begin{bmatrix} x_1^2 + \cdots + x_{2n}^2 & 0 \\ 0 & x_1^2 + \cdots + x_{2n}^2 \end{bmatrix} = (x_1^2 + \cdots + x_{2n}^2)I_{2n}.
\]

Thus, the pair $A_{2n}, A_{2n}^T$ is a matrix factorization of $f_{2n}$, as claimed.

We remark that our algorithm produces a factorization of $f_8$ with $8 \times 8$ matrices. In comparison, the standard method will produce a factorization of size $128 \times 128$. The results in [2] show that our factorizations of $f_n$ have the smallest possible size for $1 \leq n \leq 8$. Indeed, they show that, for $n \geq 8$, the smallest possible matrix factorization for $f_n$ is bounded below by $2^{n/2} \times 2^{n/2}$.

In particular, the smallest size factorizations for $f_{16}$ is $128 \times 128$. This shows that when $n = 8$, the breakdown in commutativity in step (5) of the algorithm is irreparable.

What is more, the authors of [2] show that the factorizations obtained when $n = 1, 2, 4,$ and $8$ are related to the existence of composition algebras over $\mathbb{R}$ of dimension $1, 2, 4,$ and $8$ (namely the complex numbers, quaternions, and octonians). In fact, they deduce Hurwitz’s Theorem that no real composition algebra of dimension $n$ exists for $n \neq 1, 2, 4,$ or $8$. The key ingredient in their argument is the lower bound on the size of the smallest matrix factorizations of $f_n$, stated above. They show that a necessary condition for the existence of a real composition algebra of dimension $n$ is that $f_n$ admits a matrix factorization of size $n \times n$. Since, for all $n > 8$, we have $n < 2^{n/2}$, they deduce that no composition algebra of dimension $n$ exists when $n > 8$.

References


About the authors:

David is a junior mathematics major at St. Olaf College. When not working on his classes, he can be found practicing oboe or leading the St. Olaf team to victory in the video game League of Legends.

Kosmas is an algebraist on the faculty at St. Olaf College. He earned his Ph.D. from Syracuse University in 2012. When not thinking about math, Kosmas enjoys spending his time outdoors, trying new foods, and improving his carpentry skills.

Department of Mathematics, Statistics, and Computer Science St. Olaf College
Northfield, Minnesota, 55057
crisle1@stolaf.edu  diveris@stolaf.edu