Orthogonal Tensor Decomposition

Elina Robeva

July 2, 2014

Abstract

In symmetric tensor decomposition one expresses a given symmetric tensor $T$ a sum of tensor powers of a number of vectors: $T = v_1^d + \cdots + v_k^d$. Orthogonal decomposition is a special type of symmetric tensor decomposition in which in addition the vectors $v_1, \ldots, v_k$ are required to be pairwise orthogonal. We study the properties of orthogonally decomposable tensors. In particular, we give a formula for all of the eigenvectors of an orthogonally decomposable tensor. Moreover, we give a conjecture for the defining equations of the set of orthogonally decomposable tensors.

1 Introduction

In this paper we will be concerned with the following general problem: given a homogeneous polynomial of degree $d$ in $n$ variables, under what conditions can this polynomial be written as a sum of $n$ $d$-th powers of linear forms such that the linear forms are orthogonal to each other?

Consider the space $\mathbb{C}^n$ and let $S^d(\mathbb{C}^n)$ denote the space of $d$-th order symmetric tensors. Each such tensor $T \in S^d(\mathbb{C}^n)$ can equivalently be represented by a homogeneous polynomial $f \in \mathbb{C}[x_1, \ldots, x_n]$ of degree $d$. More precisely,

$$f(x_1, \ldots, x_n) = \sum_{i_1, \ldots, i_d=1}^{n} T_{i_1, \ldots, i_d} x_{i_1} x_{i_2} \cdots x_{i_d}.$$ 

Symmetric tensor decomposition has been of much interest in the recent years. Given a polynomial $f \in \mathbb{C}[x_1, \ldots, x_n]$ homogeneous of degree $d$, the aim is to decompose it as

$$f = \sum_{i=1}^{r} \lambda_i (v_{i1} x_1 + \cdots + v_{in} x_n)^d$$

and to find the smallest $r$ possible for which such a decomposition exists. The minimum such $r$ is called the (symmetric) rank of $f$. Symmetric tensor decomposition is a hard problem and algorithms for it have been proposed by several authors, for example [6] and [3].

A very important role in symmetric tensor decomposition is played by eigenvectors of tensors. A vector $x \in \mathbb{C}^n$ is an eigenvector of the tensor $T$ if there exists $\lambda \in \mathbb{C}$ such that

$$Tx^{d-1} := \left( \sum_{i_2, \ldots, i_d=1}^{n} T_{i_2, \ldots, i_d} x_{i_2} \cdots x_{i_d} \right)_i = \lambda x.$$
Equivalently, $x$ is an eigenvector for the polynomial $f$ if
\[ \nabla f(x) = d\lambda x. \]

In this paper we are going to be concerned with a special type of symmetric tensor decomposition that does not exist for all symmetric tensors.

**Definition 1.1.** An orthogonal decomposition of a homogeneous polynomial $f \in \mathbb{C}[x_1, \ldots, x_n]$ of degree $d$ is a decomposition
\[ f = \sum_{i=1}^{r} \lambda_i (v_1 x_1 + \cdots + v_m x_m)^d \]
such that $r \leq n$ and the vectors $v_1, \ldots, v_r$ are orthonormal. An orthogonal decomposition of a symmetric tensor $T \in \bigotimes^d \mathbb{R}^n$ is a collection of orthonormal vectors $\{v_1, v_2, \ldots, v_k\}$ together with corresponding scalars $\lambda_i$ such that
\[ T = \sum_{i=1}^{k} \lambda_i v_i \otimes \cdots \otimes v_i. \]

As mentioned above, not all symmetric tensors can be decomposed in this way.

**Definition 1.2.** A symmetric tensor is orthogonally decomposable, in short, odeco, if it possesses an orthogonal decomposition.

In Section 2 we are going to give motivation from machine learning for studying orthogonal tensor decomposition. In Section 3 we give a clear algebraic characterization of all of the eigenvectors of odeco tensors. In Section 4 we study the variety of all odeco tensors and describe (some of) the generators of its ideal.

## 2 Motivation

### 2.1 The Method of Moments

The method of moments is a classical parameter estimation technique from statistics. The basic setup is as follows. Let $X_1, X_2, \ldots, X_k$ be i.i.d. random variables coming from the model distribution whose parameters we would like to compute. Assume that these variables take values in $\{1, 2, \ldots, m\}$. Compute samples $X_1^{(i)}, X_2^{(i)}, \ldots, X_k^{(i)}$ for $i = 1, 2, \ldots, N$. Let $y_1^{(i)}, \ldots, y_k^{(i)}$ be vectors in $\mathbb{R}^m$ such that $y_j^{(i)}$ has coordinate 1 at $X_j^{(i)}$ and has all other coordinates 0, i.e. the $y_j^{(i)}$’s are standard basis vectors. Then, compute the $k$-th empirical moment as
\[ \mathbb{E}[X_1 \otimes X_2 \otimes \cdots \otimes X_k] := \sum_{i=1}^{N} y_1^{(i)} \otimes y_2^{(i)} \otimes \cdots \otimes y_k^{(i)}. \]

Computing the empirical moments for $k = 1, 2, 3$ is often times enough to estimate the model parameters efficiently [1], [2].
The authors of [1] show that the parameter estimation task from the observed moments can be reduced to extracting a symmetric orthogonal decomposition after slightly changing the original tensors to orthogonally decomposable ones. They solve the problem of orthogonal tensor decomposition in a variety of approaches, including fixed-point and variational methods. In the following subsections we describe these methods as well as some of the theory about orthogonal tensor decomposition developed in [1].

2.2 Fixed-Points of the Tensor Power Method

Let $T \in S^d(\mathbb{R}^n)$. If $T$ is orthogonally decomposable, i.e. $T = \sum_{i=1}^{k} \lambda_i v_i \otimes d_i$, then $T v_i^2 = \lambda_i v_i$, for all $i = 1, 2, ..., k$. Therefore, each $(v_i, \lambda_i)$ is an eigenvector-eigenvalue pair.

**Definition 2.1.** A unit vector $u \in \mathbb{R}^n$ is a robust eigenvector of $T$ if there exists $\epsilon > 0$ such that for all $\theta \in \{u' \in \mathbb{R}^n : ||u - u'|| < \epsilon\}$, repeated iteration of the map

$$\overline{\theta} \mapsto \frac{T\overline{\theta}^{d-1}}{||T\overline{\theta}^{d-1}||},$$

starting from $\theta$ converges to $u$.

The following theorem implies that if $T$ has an orthogonal decomposition, then the set of robust eigenvectors of $T$ is precisely the set $\{v_1, v_2, ..., v_k\}$, implying that the orthogonal decomposition is unique.

**Theorem 2.2.** Let $T$ have an orthogonal decomposition $T = \sum_{i=1}^{k} \lambda_i v_i \otimes d_i$ as in the definition.

1. The set of $\theta \in \mathbb{R}^n$ which do not converge to some $v_i$ under repeated iteration of (2.1) has measure 0.

2. The set of robust eigenvectors of $T$ is equal to $\{v_1, v_2, ..., v_k\}$.

Therefore, to recover the orthogonal decomposition of $T$, one needs to find the robust eigenvectors. The definition of robust eigenvectors suggests an algorithm to compute them, namely repeated iteration of the map (2.1). Starting with random vectors $u \in \mathbb{R}^m$. This is called the tensor power method. The formal algorithm is as follows

1. Start with an orthogonally decomposable tensor $T$.

2. Repeat until $T = 0$.

   Choose random $u \in \mathbb{R}^m$.

   Let $v$ be the result of applying the tensor power method to $u$.

   Compute the eigenvalue $\lambda$ corresponding to $v$, from the equation $Tv^{d-1} = \lambda v$.

   Set $T = T - \lambda v \otimes d$.

3. End Repeat
2.3 Variational Characterization

Theorem 2.3. Let \( n \geq 2 \). Let \( T \) have an orthogonal decomposition \( T = \sum_{i=1}^{k} \lambda_i v_i \otimes v_i \otimes v_i \) as before, and consider the optimization problem
\[
\max_{u \in \mathbb{R}^n} T(u, u, u) \text{ such that } ||u|| = 1.
\]
1. The stationary points of \( T \) are eigenvectors of \( T \).
2. A stationary point \( u \) is an isolated local maximizer if and only if \( u = v_i \) for some \( i \in [k] \).

The generalized Rayleigh quotient for a third-order tensor is
\[
u \mapsto T(u, u, u) \frac{u^T u}{(u^T u)^{3/2}}.
\]

The above theorem shows that a vector \( u \in \mathbb{R}^n \) is an isolated local maximizer of the generalized Rayleigh quotient if and only if \( u = v_i \) for some \( i = 1, 2, ..., k \). Therefore, one can compute the robust eigenvectors by computing the isolated local maximizers or the generalized Rayleigh quotient.

2.4 Reducing to Orthogonal Decomposition

In certain cases, recovering the orthogonal decomposition of the empirical moments recovers all of the model parameters. For example, the authors of [1] show that the following structure of the second and third moments:
\[
M_2 = \sum_{i=1}^{k} \omega_i \mu_i \otimes \mu_i \\
M_3 = \sum_{i=1}^{k} \omega_i \mu_i \otimes \mu_i \otimes \mu_i,
\]
is valid for a family of statistical models. We would like to transform \( M_2 \) and \( M_3 \) into tensors whose orthogonal decomposition will help us recover \( \omega_i \) and \( \mu_i \). We will show that this can be done under the following degeneracy condition.

Condition 2.4 (non-degeneracy). The vectors \( \mu_1, ..., \mu_k \in \mathbb{R}^d \) are linearly independent and \( \omega_1, ..., \omega_k > 0 \).

The transformation goes as follows. Let \( W \in \mathbb{R}^{d \times k} \) be such that
\[
M_2(W, W) = W^T M_2 W = I_{k \times k},
\]
i.e. let \( W := UD^{-1} \), where \( M_2 = U D U^T \) and \( U \) is the \( d \times k \) matrix of orthogonal eigenvectors of \( M_2 \) with nonnegative eigenvalues, and \( D \in \mathbb{R}^{k \times k} \) is the diagonal matrix of positive eigenvalues of \( M_2 \). Set \( \hat{\mu}_i := \sqrt{\omega_i} W^T \mu_i \). Then, we have that
\[
\tilde{M}_2 := M_2(W, W) = \sum_{i=1}^{k} W^T (\sqrt{\omega_i} \mu_i) (\sqrt{\omega_i} \mu_i)^T W = \sum_{i=1}^{k} \hat{\mu}_i \hat{\mu}_i^T = I.
\]
Therefore, the $\tilde{\mu}_i$ are orthonormal. Moreover,

$$\tilde{M}_3 := M_3(W, W, W) = \sum_{i=1}^{k} \omega_i (W^T \mu_i)^{\otimes 3} = \sum_{i=1}^{k} \frac{1}{\sqrt{\omega_i}} \tilde{\mu}_i^{\otimes 3}.$$ 

Thus, using orthogonal tensor decomposition for $\tilde{M}_3$, we can recover $\frac{1}{\sqrt{\omega_i}}$ and $\tilde{\mu}_i$, which means that we can recover $\omega_i$ and $\mu_i$. Therefore, orthogonal tensor decomposition as well as more knowledge about the eigenvectors of odeco tensors are quite relevant to this application.

### 3 The Variety of Eigenvectors of Tensors

In this section, we are going to study the variety of all eigenvectors of a given orthogonally decomposable tensor.

As we mentioned in the Introduction, a symmetric tensor $T \in S^m(\mathbb{R}^n)$ can equivalently be represented by a homogeneous polynomial $f \in \mathbb{R}[x_1, \ldots, x_n]$ of degree $m$. Indeed, given such a tensor $T$, we obtain $f$ by

$$f = \sum_{i_1 + \cdots + i_n = m} \binom{m}{i_1, \ldots, i_n} T_{i_1, \ldots, i_n} x_1^{i_1} \cdots x_n^{i_n}.$$ 

Then, $Tx^{m-1} = \lambda x$ is equivalent to $\nabla f(x) = m\lambda x$, i.e. $\nabla f(x)$ and $x$ are parallel to each other. This is equivalent to the vanishing of the $2 \times 2$ minors of the matrix $[\nabla f(x)|x]$.

**Definition 3.1.** The variety of eigenvectors $V_f$ of a given symmetric tensor $f$ is given by the $2 \times 2$ minors of the matrix $[\nabla f(x)|x]$.

We are going to be interested in the eigenvectors and, in general, in the space of tensors that have an orthogonal decomposition.

**Definition 3.2.** A symmetric tensor $f \in S^m(\mathbb{R}^n)$ is odeco if there exists an $n \times n$ matrix $V$ whose rows are pairwise orthogonal and such that

$$f(x) = \sum_{i=1}^{n} ((Vx)_i)^m$$

The aim of this section is to prove the following theorem.

**Theorem 3.3.** An odeco tensor $f \in S^m(\mathbb{R}^n)$ has $\frac{(m-1)n-1}{m-2}$ eigenvectors, which are exactly the the fixed points of the gradient map in $\mathbb{P}^n$, given explicitly in terms of the $(m-2)^{nd}$ roots of the entries of $V$. More precisely, the eigenvectors of $f$ are

$$(x_1 : \cdots : x_n) =$$

$$V^T(\eta_1 \lambda_1^{\frac{1}{m-2}} : \cdots : \eta_{k-1} \lambda_{k-1}^{\frac{1}{m-2}} : \lambda_k^{\frac{1}{m-2}} : 0 : \cdots : 0)^T,$$

where $k = 1, \ldots, n$ and $\eta_1, \ldots, \eta_{k-1}$ are $d - 2^{nd}$ roots of unity.
We illustrate this theorem by a simple example.

**Example 3.4.** Let \( m = n = 3 \) and consider the matrix

\[
V = \begin{pmatrix} 1 & 3 & 1 \\ -2 & 1 & -1 \\ 4 & 1 & -7 \end{pmatrix}.
\]

\( V \) has orthogonal rows and

\[
f(x, y, z) = (x + 3y + z)^3 + (-2x + y - z)^3 + (4x + y - 7z)^3.
\]

We use Macaulay2 to decompose the ideal \( I \) given by the \( 2 \times 2 \) minors of \( \left[ \nabla f(x) \right] \). We obtain the following decomposition

\[
I = \langle -y + 3z, x - z \rangle \cap \langle y + z, x - 2z \rangle \cap \langle 7y + z, 7x + 4z \rangle \\
\cap \langle 64y + 61z, 64x - 119z \rangle \cap \langle -29y + 109z, 29x - 40z \rangle \\
\cap \langle 2y + 5z, 46x - 101z \rangle \cap \langle 85y + 229z, 85x - 206z \rangle.
\]

The first three ideals correspond to the rows of \( V \), which are the robust eigenvectors of \( T \). The 7 eigenvectors that we get are

\[
(1 : 3 : 1), (2 : -1 : 1), (4 : 1 : -7),
\]
\[
(119 : -61 : 64), (40 : 109 : 29),
\]
\[
(101 : 115 : 46), (206 : -229 : 85)
\]

**Lemma 3.5.** The above theorem holds for the case \( f(x_1, \ldots, x_n) = a_1x_1^m + a_2x_2^m + \cdots + a_nx_n^m \). In particular, the decomposition of the ideal \( I \) is as follows.

\[
I = \bigcap_{k=1}^{n} \text{intersection of ideals of type } (k),
\]

where an ideal is of type \( (k) \) if up to permutation of \( x_1, x_2, \ldots, x_n \) and of \( a_1, a_2, \ldots, a_n \), it is equal to

\[
\langle x_1 - \eta_1 \left( \frac{a_k}{a_1} \right)^{\frac{1}{m-2}} x_k, \ldots, x_{k-1} - \eta_{k-1} \left( \frac{a_k}{a_{k-1}} \right)^{\frac{1}{m-2}} x_k, x_{k+1}, \ldots, x_n \rangle,
\]

where \( \eta_1, \ldots, \eta_{k-1} \) are \((m-2)^{nd} \) roots of unity. In particular, there are \( \binom{n}{k} (m-2)^{k-1} \) ideals of type \( (k) \). Each ideal as above has exactly one solution in \( \mathbb{CP}^n \), representing one eigenvector, namely

\[
(x_1 : \ldots : x_n) = (\eta_1 a_1^{\frac{1}{m-2}} : \ldots : \eta_{k-1} a_{k-1}^{\frac{1}{m-2}} : a_k^{\frac{1}{m-2}} : 0 : \ldots : 0).
\]

The total number of such solutions is \( \frac{(m-1)^n - 1}{m-2} \).
Then, we have that

\[ \nabla f(x) | x = \begin{bmatrix} a_1 x_1^{m-1} & x_1 \\ a_2 x_2^{m-1} & x_2 \\ \vdots & \vdots \\ a_n x_n^{m-1} & x_{n-1} \end{bmatrix} \]

Therefore, the ideal of \(2 \times 2\) minors is given by

\[ I = \langle x_i x_j (a_i x_i^{m-2} - a_j x_j^{m-2}) : i \neq j \rangle. \]

We would like to decompose this ideal. Note that any primary ideal \(P \supseteq I\) would either contain \(x_i x_j\) or \(a_i x_i^{m-2} - a_j x_j^{m-2}\) for all \(i \neq j\). Suppose that for a given \(P\), \(\sqrt{P}\) doesn’t contain exactly \(k\) of the variables \(x_1, \ldots, x_n\). We call such a prime \(\sqrt{P}\) of type \((k)\). So, say \(\sqrt{P}\) contains \((\text{WLOG}) x_{k+1}, \ldots, x_n\) for some \(k \in \{1, \ldots, n\}\) and it doesn’t contain \(x_1, \ldots, x_k\). Then, we have that \(\sqrt{P}\) equals \(\langle x_{n-k+1}, \ldots, x_n \rangle + \) an associated prime of

\[ I_k = \langle a_i x_i^{m-2} - a_j x_j^{m-2} : i \neq j, i, j \leq k \rangle = \langle a_i x_i^{m-2} - a_{i+1} x_{i+1}^{m-2} : i = 1, \ldots, k-1 \rangle. \]

The decomposition of \(I_k\) is the same as the decomposition of the lattice ideal associated to the lattice \(L_\rho = \langle (m-2)(e_i - e_k) : i = 1, \ldots, k-1 \rangle\) with partial character \(\rho : L_\rho \to \mathbb{C}^*\) given by

\[ \rho((m-2)(e_i - e_k)) = \frac{a_k}{a_i}. \]

By [Theorem 2.1.(d), [5]], we know that \(I(\rho) = \langle x^m - \rho(m) : m \in L_\rho \rangle = \langle x_i^{m-2} x_j^{(m-2)} - \frac{a_i}{a_j} \rangle\) has the following decomposition

\[ I(\rho) = \bigcap_{\rho' \text{ extends } \rho \text{ to } L} I(\rho'), \]

where \(L_\rho \subseteq L \subseteq \mathbb{Z}^n\) and \(|L/L_\rho|\) is finite. In this case, we can choose

\[ L = \langle e_i - e_k : i = 1, \ldots, k-1 \rangle. \]

Then, \(|L/L_\rho| = (m-2)^{k-1}\). Moreover, by the same theorem, the number of \(\rho'\) extending \(\rho\) is exactly \(|L/L_\rho| = (m-2)^{k-1}\). Also, note that each such \(\rho' : L \to \mathbb{C}^*\) is uniquely defined by the values

\[ \eta_i \left( \frac{a_k}{a_i} \right)^{\frac{1}{m-2}} := \rho'(e_i - e_k) \]

and each \(\eta_i\) is an \((m-2)\)th root of unity. Therefore,

\[ I(\rho') = \langle x_i - \eta_i \left( \frac{a_k}{a_i} \right)^{\frac{1}{m-2}} x_k : i = 1, 2, \ldots, k-1 \rangle. \]

This gives an explicit description of the set

\[ V(I(\rho')) = \{(\eta_1 a_1^{-\frac{1}{m-2}} : \eta_2 a_2^{-\frac{1}{m-2}} : \cdots : \eta_{k-1} a_{k-1}^{-\frac{1}{m-2}} : a_k^{-\frac{1}{m-2}})\}. \]
We have that
\[ I_k = I(\rho) \cap \mathbb{C}[x_1, \ldots, x_k] = \bigcap_{\eta_1, \ldots, \eta_{k-1}} \langle x_i - \eta_i \left( \frac{a_k}{a_i} \right)^{\frac{i-1}{m-2}} x_k : i = 1, 2, \ldots, k - 1 \rangle, \]
where \( \eta_1, \ldots, \eta_{k-1} \) vary over the \((m - 2)^{th}\) roots of unity. Note that \(|V(I(\rho))| = |\{\rho' : \rho' \text{ extends } \rho\}| = (m - 2)^{k-1}\), because there are \(m - 2\) options for each of the \((m - 2)^{th}\) roots of unity \(\eta_i\) for \(i = 1, 2, \ldots, k - 1\). Each element of \(V(I(\rho))\) defines one eigenvector of \(f\), namely
\[ (x_1 : \ldots : x_n) = (\eta_1 a_1^{-\frac{1}{m-2}} : \ldots : \eta_{k-1} a_{k-1}^{-\frac{1}{m-2}} : a_k^{-\frac{1}{m-2}} : 0 : \ldots : 0). \]

The other eigenvectors of type \((k)\) have their \(n - k\) zeros permuted.

To sum up, we have that
\[ I = \bigcap_{k=1}^{n} \bigcap_{1 \leq i_1 < i_2 < \ldots < i_k \leq n} \bigcap_{\eta_1, \ldots, \eta_{k-1}} \langle x_{i_j} - \eta_j \left( \frac{a_k}{a_{i_j}} \right)^{\frac{i_j-1}{m-2}} x_k : j = 1, 2, \ldots, k - 1 \rangle + \langle x_i : i \neq i_1, \ldots, i_k \rangle. \]

Therefore, the number of eigenvectors of \(f\) is
\[ \sum_{k=1}^{n} \binom{n}{k} (m - 2)^{k-1} = \frac{1}{m - 2} \sum_{k=1}^{n} \binom{n}{k} (m - 2)^k = \frac{1}{m - 2} ((m - 2 + 1)^n - 1) = \frac{(m - 1)^n - 1}{m - 2}, \]
recovering the formula expected by [4]. □

Now, we proceed to the proof of Theorem 3.3.

**Proof of Theorem 3.3.** Now, let \(V = \begin{bmatrix} -v_1 & - \\ \vdots & \vdots \\ -v_n & - \end{bmatrix} \in \mathbb{C}^{n \times n}\) be an orthogonal matrix, i.e. the vectors \(v_1, \ldots, v_n\) form an orthonormal basis. Then, we have that
\[ f(x) = \sum_{i=1}^{n} a_i (v_i \cdot x)^m \]
and
\[ \frac{1}{m} \nabla f(x) = \sum_{i=1}^{n} a_i (v_i \cdot x)^{m-1} v_i. \]

If \(x\) is an eigenvector, then we have that
\[ \frac{1}{m} \nabla f(x) = \sum_{i=1}^{n} a_i (v_i \cdot x)^{m-1} v_i = \lambda x. \]
But note that \( v_1, v_2, \ldots, v_n \) are an orthonormal basis of \( \mathbb{R}^n \) and therefore \( x = \sum_{i=1}^{n} (v_i \cdot x) v_i \).

Since the \( v_i \) are linearly independent and
\[
\sum_{i=1}^{n} a_i (v_i \cdot x)^{m-1} v_i = \lambda \sum_{i=1}^{n} (v_i \cdot x) v_i,
\]
then an equivalent description of the above equality is that for all \( i \neq j \), we have that
\[
\frac{a_i (v_i \cdot x)^{m-1}}{a_j (v_j \cdot x)^{m-1}} = \frac{(v_i \cdot x)}{(v_j \cdot x)}.
\]

Let
\[
y_i = (v_i \cdot x), \text{ i.e. } y = Vx.
\]

Then, an equivalent description of \( x \) being an eigenvector is that for all \( i \neq j \),
\[
a_i y_i^{m-1} y_j = a_j y_j^{m-1} y_i
\]
and the ideal describing the eigenvectors is given by
\[
I = \langle a_i y_i^{m-1} y_j - a_j y_j^{m-1} y_i : i \neq j \rangle.
\]

By Lemma 3.5, we have that our ideal
\[
I = \bigcap_{k=1}^{n} \text{intersection of ideals of type (k)}
\]
and each ideal of type (k), up to a permutation of \( y_1, \ldots, y_n \) and \( a_1, \ldots, a_n \) has the form
\[
\langle y_1 - \eta_1 \left( \frac{a_k}{a_1} \right)^{\frac{1}{m-2}} y_k, \ldots, y_k-1 - \eta_{k-1} \left( \frac{a_k}{a_{k-1}} \right)^{\frac{1}{m-2}} y_k, y_{k+1}, \ldots, y_n \rangle,
\]
where \( \eta_1, \ldots, \eta_{k-1} \) are \((m-2)^{nd}\) roots of unity. Each such ideal has exactly one solution in \( \mathbb{CP}^n \), representing one eigenvector, namely
\[
(y_1 : \ldots : y_n) = (\eta_1 a_1^{-\frac{1}{m-2}} : \ldots : \eta_{k-1} a_{k-1}^{-\frac{1}{m-2}} : a_k^{-\frac{1}{m-2}} : 0 : \ldots : 0).
\]

Since \( y_i = v_i \cdot x_i \), then, the corresponding ideal in the \( x \)-variables would be obtained by plugging in that expression of the \( y_i \) into 3.1.

But \( y = Vx \) and \( V \) is an orthogonal matrix. Therefore,
\[
x = V^T y = V^T (\eta_1 a_1^{-\frac{1}{m-2}} : \ldots : \eta_{k-1} a_{k-1}^{-\frac{1}{m-2}} : a_k^{-\frac{1}{m-2}} : 0 : \ldots : 0).
\]

By Lemma 3.5, we know that there are \( \binom{n}{k} (m-2)^{k-1} \) eigenvectors of type (k), which makes for a total of \( \frac{(m-1)^n - 1}{m-2} \) eigenvectors.
4 The Odeco Variety

The Odeco Variety is the Zariski closure of the set of all tensors $T \in S^m(\mathbb{C}^n)$ which are orthogonally decomposable. If a tensor is odeco, then, in particular, it is decomposable as a sum of $n$ powers of linear forms, i.e. it lies in the $n$-th secant variety of the $m$-th Veronese variety, denoted by $\sigma_n(v_m(\mathbb{C}^n))$.

When $m = n = 3$, the equation of $\sigma_3(v_3(\mathbb{C}^3))$ is called the Aronhold invariant and it is given by the Pfaffian of a certain skew-symmetric matrix.

We are going to describe all of the quadratic equations cutting out the Odeco variety. Consider a tensor $f(x_1, x_2, ..., x_n)$, i.e. a homogeneous polynomial in $\mathbb{C}[x_1, ..., x_n]$ of degree $m$.

Let $d_1, d_2, ..., d_N$ be all the possible $n - 3$rd partial derivatives of our $f$.

Lemma 4.1. If $f$ is orthogonally decomposable, i.e. $f(x_1, ..., x_n) = \sum_{i=1}^{n}((Vx)_i)^m$, where the rows of $V$ are pairwise orthogonal, then,

$$\sum_{s=1}^{n} \left( \frac{\partial^4 f}{\partial x_s \partial x_v} \frac{\partial^4 f}{\partial x_s \partial x_v} - \frac{\partial^4 f}{\partial x_s \partial x_v} \frac{\partial^4 f}{\partial x_s \partial x_v} \right) = 0 \quad (4.1)$$

for all $y, v, w, z \in \mathbb{Z}_{\geq 0}$ such that $\sum_i y_i = \sum_i v_i = \sum_i z_i = \sum_i w_i = d - 1$ and $y + v = z + w$.

Proof. Assume that the rows of $V$ are the orthogonal vectors $v_1, ..., v_n$ so that $f(x_1, ..., x_n) = \sum_{i=1}^{n} (v_i \cdot x)^m$.

Then,

$$\sum_{s=1}^{n} \left( \frac{\partial^4 f}{\partial x_s \partial x_v} \frac{\partial^4 f}{\partial x_s \partial x_v} - \frac{\partial^4 f}{\partial x_s \partial x_v} \frac{\partial^4 f}{\partial x_s \partial x_v} \right) =$$

$$= \sum_{s=1}^{n} \left( \sum_{r=1}^{n} v_r v_r^y v_r^v \sum_{r=1}^{n} v_r v_r^w v_r^w - \sum_{r=1}^{n} v_r v_r^w v_r^w \sum_{r=1}^{n} v_r v_r^w v_r^w \right)$$

$$= \sum_{s=1}^{n} \left( \sum_{r=1}^{n} (v_r^y v_r^v - v_r^w v_r^w) + \sum_{r \neq t} (v_r v_t v_t v_t v_t - v_r v_t v_t v_t v_t) \right)$$

$$= \sum_{r \neq t} \left( v_r^y v_t^w v_t^w \sum_{s=1}^{n} v_r v_t - v_r^w v_t^w \sum_{s=1}^{n} v_r v_t \right) = 0.$$

Note that each of the equations in (4.1) is a quadratic equation in the coefficients of $f$, i.e. it doesn’t contain any of the $x$ variables. More precisely,

$$\sum_{s=1}^{n} \left( \frac{\partial^4 f}{\partial x_s \partial x_v} \frac{\partial^4 f}{\partial x_s \partial x_v} - \frac{\partial^4 f}{\partial x_s \partial x_v} \frac{\partial^4 f}{\partial x_s \partial x_v} \right) = \sum_{s=1}^{n} u_{y+e_s} u_{v+e_s} - u_{w+e_s} u_{z+e_s},$$

where $e_1, ..., e_n$ are the standard basis vectors.

The aim of the future work on this project is to prove or disprove the following conjecture.
Conjecture 4.2. The equations in (4.1) cut out the Odeco variety.

The following examples give a justification of this Conjecture for small cases of $m$ and $n$.

Example 4.3 ($m = n = 3$).
The 6 quadrics from (4.1) form a prime ideal of dimension 5 and degree 10 that contains the Aronhold invariant. Since the the degrees of freedom for choosing the orthogonal matrix $V$ and the are exactly 5, then, this has to be exactly the ideal of the Odeco variety.

Example 4.4 ($m = 4, n = 3$).
A computation of all of the quadrics vanishing on the Odeco variety for $m = 4$ and $n = 3$ shows that the ideal $I$ generated by these has dimension 5 and degree 35.
The ideal $J$ generated by the quadrics in (4.1) is contained inside $I$. The calculation of its exact dimension and degree did not terminate. However, computing its Gröbner basis up to 200 elements and taking the ideal $J_1$ generated by these 200 elements shows that $J_1$ also has dimension 5 and degree 35. We believe that 200 is a large enough number and that the dimension of $J$ is, indeed 5 and its degree is 35, showing that it does cut out the Odeco variety in this case. [I am not sure how this could be true].

References


