Decomposing Tensors into Equiangular Tight Frames

September 2, 2020

Abstract

In this article we explore symmetric tensors which can be decomposed as the sum of rank one terms coming from an equiangular tight frame (ETF). In particular, we study the eigenvectors of such tensors, and show that the elements of the ETF robust eigenvectors for the tensor, and, thus, can be found efficiently via the tensor power method.

1 Introduction

A *tensor* T of *order* d and dimension n is an $n \times n \times \cdots \times n$ (d times) table of numbers. The element in position $i_1, i_2, \ldots, i_d$ of T is denoted by $T_{i_1 i_2 \ldots i_d}$. We denote the set of tensors of order d and dimension n by $\mathbb{R}^n \otimes \cdots \otimes \mathbb{R}^n$. A tensor $T \in \mathbb{R}^n \otimes \cdots \otimes \mathbb{R}^n$ is *symmetric* if for any permutation π of $\{1, \ldots, d\}$ we have

$$T_{i_1 \ldots i_d} = T_{i_{\pi(1)} \ldots i_{\pi(d)}}.$$

We denote the set of symmetric tensors of order d by $S^d(\mathbb{R}^n)$. A *symmetric tensor decomposition* (or *Waring decomposition*) of a symmetric tensor $T \in S^d(\mathbb{R}^n)$ is an expression of T of the form

$$T = \sum_{i=1}^{r} \lambda_i v_i^{\otimes d},$$

where $\alpha_1, \ldots, \alpha_r \in \mathbb{R}$, $v_1, \ldots, v_r \in \mathbb{R}^d$, and $v_i^{\otimes d} = v_i \otimes \cdots \otimes v_i$ (d-times) is a *rank-one tensor*. The smallest r for which such a decomposition exists is called the *symmetric rank* (or, for short, the *rank*) of T.

The rank of any symmetric matrix is always at most n. For tensors $T \in S^d(\mathbb{R}^n)$ this is not the case. The rank can be much much larger. The Alexander-Hirschowitz theorem states that the rank of a random tensor $T \in S^d(\mathbb{R}^n)$ is $\lfloor \frac{1}{n}\binom{n+d-1}{d} \rfloor$ with probability 1, except for a few special values of d and n (not listed here) when the rank is 1 less than this number with probability 1. While decomposing symmetric matrices into their eigenvectors and eigenvalues can be done efficiently, finding the decomposition of a generic tensor T is NP-hard.

A vector $v \in \mathbb{C}^n$ is an eigenvector of T with eigenvalue λ if

$$T \cdot v^{d-1} = \lambda v,$$

where $T \cdot v^{d-1}$ is a vector defined by *contracting* T by v along all of its dimensions except for one, i.e. the i-th entry of $T \cdot v^{d-1}$ is

$$(T \cdot v^{d-1})_i = \sum_{i_1,\ldots,i_{d-1}=1}^{n} T_{i_1\ldots i_{d-1}i} v_{i_1} \cdots v_{i_{d-1}}.$$

Two eigenvector-eigenvalue pairs (v, λ) and (v', λ') are equivalent if there exists a scalar $t \neq 0$ such that $v' = tv$ and $\lambda' = t^{d-2}\lambda$.

A symmetric tensor $T \in S^d(\mathbb{R}^n)$ is *orthogonally decomposable* (or *odeco*) if it has a decomposition of the form

$$T = \sum_{i=1}^{n} \lambda_i v_i^{\otimes d},$$

where $v_1, \ldots, v_n$ form an orthonormal basis of $\mathbb{R}^n$.

Since there are at most n orthogonal vectors in $\mathbb{R}^n$, then the rank of an orthogonally decomposable tensor is at most n.

One reason orthogonally decomposable tensors have been studied [1, 4] is that one can find their decomposition efficiently. One way to do this is via the *tensor power method.* Consider the map

$$v \mapsto \frac{T \cdot v^{d-1}}{||T \cdot v^{d-1}||}.$$

Here, if the numerator is 0, then we consider the fraction to also be 0. This is called the *tensor power iteration.* A vector v is a *fixed point* of the tensor power iteration if it maps to itself (or to the negative of itself) via the power iteration map.

An eigenvector v of a tensor T is called *robust* if there exists $\epsilon > 0$ such that for **any** vector w in the ball with center v and radius ϵ, if we start at w, and repeatedly apply the tensor power iteration map, we always converge to v. If T is orthogonally decomposable, then the elements in its decomposition $v_1, \ldots, v_k$ are precisely its robust eigenvectors [1]. Furthermore, when starting from a random vector $v \in \mathbb{R}^n$, the power method for T converges to one of those robust eigenvectors exponentially fast. Therefore, using the power method, we can decompose T efficiently. Although orthogonally decomposable tensors can be decomposed efficiently, they only constitute a tiny portion of the set of all tensors. The rank of a general tensor is $\lfloor \frac{1}{n}\binom{n+d-1}{d} \rfloor$ whereas the rank of an orthogonally decomposable one is at most n. In this article, we study another set of tensors which can be decomposed efficiently via the tensor power method.We replace the orthonormal basis from the definition of orthogonally decomposable tensors by the more general notion of an *equiangular tight frame.*

Definition 1.1. A *finite unit norm tight frame* is a set of vectors $v_1, \ldots, v_m \in \mathbb{R}^n$, where $m \geq n$, such that

$$\begin{pmatrix} | & & | \\ v_1 & \cdots & v_m \\ | & & | \end{pmatrix} \begin{pmatrix} - & v_1 & - \\ & \vdots & \\ - & v_m & - \end{pmatrix} = \frac{m}{n} I \qquad \text{and} \qquad ||v_i|| = 1, i = 1, \ldots, m,$$

where I is the $n \times n$ identity matrix.

An *equiangular tight frame* (or an *ETF*) is a finite unit norm tight frame $v_1, \ldots, v_m$ with the additional constraint that there exists $\alpha \geq 0$ such that, for all $i \neq j$,

$$|\langle v_i, v_j \rangle| = \alpha.$$

A tensor $T \in S^d(\mathbb{R}^n)$ is *ETF decomposable* if there exists an ETF $v_1, \ldots, v_m$ such that

$$T = \sum_{i=1}^{m} v_i^{\otimes d}.$$

Example 1.2. An orthonormal basis $v_1, \ldots, v_n$ is also an equiangular tight frame. Therefore, any orthodonally decomposable tensor is also ETF decomposable.

Example 1.3. The Mercedes Benz frame

is the set of vectors $\begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} \frac{\sqrt{3}}{2} \\ -\frac{1}{2} \end{pmatrix}, \begin{pmatrix} -\frac{\sqrt{3}}{2} \\ -\frac{1}{2} \end{pmatrix}$.

Example 1.4. The following four vectors in $\mathbb{R}^3$ also form an ETF

$$\frac{1}{\sqrt{27}} \begin{pmatrix} 1 \\ 1 \\ -5 \end{pmatrix}, \frac{1}{\sqrt{27}} \begin{pmatrix} 1 \\ -5 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{27}} \begin{pmatrix} -5 \\ 1 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{27}} \begin{pmatrix} 3 \\ 3 \\ 3 \end{pmatrix}.$$

During the course of the project we will try to solve the following problems.

Problem 1.5. Suppose that T is ETF decomposable. Show that the elements $v_1, \ldots, v_m$ of the ETF are eigenvectors of T.

Problem 1.6. Suppose that T is ETF decomposable. Perform computational experiments to find conditions under which the elements $v_1, \ldots, v_m$ of the ETF are the robust eigenvectors of T.

Problem 1.7. Under what conditions are the elements of the ETF are precisely the robust eigenvectors?

Problem 1.8. From this, show that the tensor power method recovers the decomposition of an ETF decomposable tensor T, and find the rate at which this happens.

In this article, we study the convergence properties of the tensor power method for ETF decomposable tensors. In Section 2, we begin with the simplest case of Mercedes-Benz decomposable tensors in $S^d(\mathbb{R}^2)$. We show that as long as $d \geq 5$, the set of robust eigenvectors is precisely the set of vectors in the Mercedes-Benz frame. In Section 3, we extend our findings to Mercedes-Benz decomposable tensors in $S^d(\mathbb{R}^n)$ (for a definition of a Mecedes-Benz frame in $\mathbb{R}^n$, see Definition **??**). In this case, we also show that the tensor power method recovers the decomposition efficiently. Finally, in Section 4, we study general ETF decomposable tensors. In this case we show conditions under which the elements of the ETF are eigenvectors of the tensor, and we conjecture that they are also robust eigenvectors.

2 Decomposing tensors into the Mercedez-Benz Frame

We begin our study by looking at the simples ETF. The Mercedes Benz frame which is formed by the vectors

$$v_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, v_2 = \begin{pmatrix} \frac{\sqrt{3}}{2} \\ -\frac{1}{2} \end{pmatrix}, v_3 = \begin{pmatrix} -\frac{\sqrt{3}}{2} \\ -\frac{1}{2} \end{pmatrix}.$$

It is the only equiangular tight frame in $\mathbb{R}^2$ with more than 2 vectors (apart from rotations of it). A tensor $T_d \in \mathcal{S}^d(\mathbb{R}^2)$ decomposes into this frame if

$$T_d = v_1^{\otimes d} + v_2^{\otimes d} + v_3^{\otimes d}.$$

2.1 The Tensor Power Method

$$v_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, v_2 = \begin{bmatrix} \frac{\sqrt{3}}{2} \\ -\frac{1}{2} \end{bmatrix}, v_3 = \begin{bmatrix} -\frac{\sqrt{3}}{2} \\ -\frac{1}{2} \end{bmatrix}$$
$$T = v_1^{\otimes d} + v_2^{\otimes d} + v_3^{\otimes d}$$

Tensor Power Method:

$$v^{(i+1)} \leftarrow \frac{T \cdot (v^{(i)})^{d-1}}{||T \cdot (v^{(i)})^{d-1}||}$$

2.2 Proof of Tensor Power Method Convergence

In this section we will prove the notion that the unit vectors in the various colored regions of Fig 1. converge to the corresponding vectors of the Mercedes-Benz ETF in $\mathbb{R}^2$ in the tensor power method.

Theorem 2.1. *The vectors v_1, v_2, v_3 of a given Mercedes-Benz ETF in $\mathbb{R}^2$ are robust eigenvectors for the tensor*

$$T = v_1^{\otimes d} + v_2^{\otimes d} + v_3^{\otimes d}$$

for any $d \geq 5$.

Note that for odd values of d, we observe different convergence behavior than even d in the tensor power method (see Figure 1). Thus, for odd d, we will be demonstrating that if a vector is close enough in angle to a vector in in the Mercedes-Benz ETF, then it will converge to that vector.

For the even case, we can prove a slightly strong statement; Given some vector $v \in \mathbb{R}^2$ and d even, we will show that v converges in the tensor power method to the vector $v_i \in \{v_1, v_2, v_3\}$ in the ETF which minimizes $|\cos(\theta_i)|$, where θ_i is the angle between $v_i \in \{v_1, v_2, v_3\}$ and v.

To do this, we will be analyzing values at sequential iterations of the tensor power method. Given some vector $v^{(i)} = [v_1^{(i)}, v_2^{(i)}]^T$, which represents the i-th iteration of the tensor power

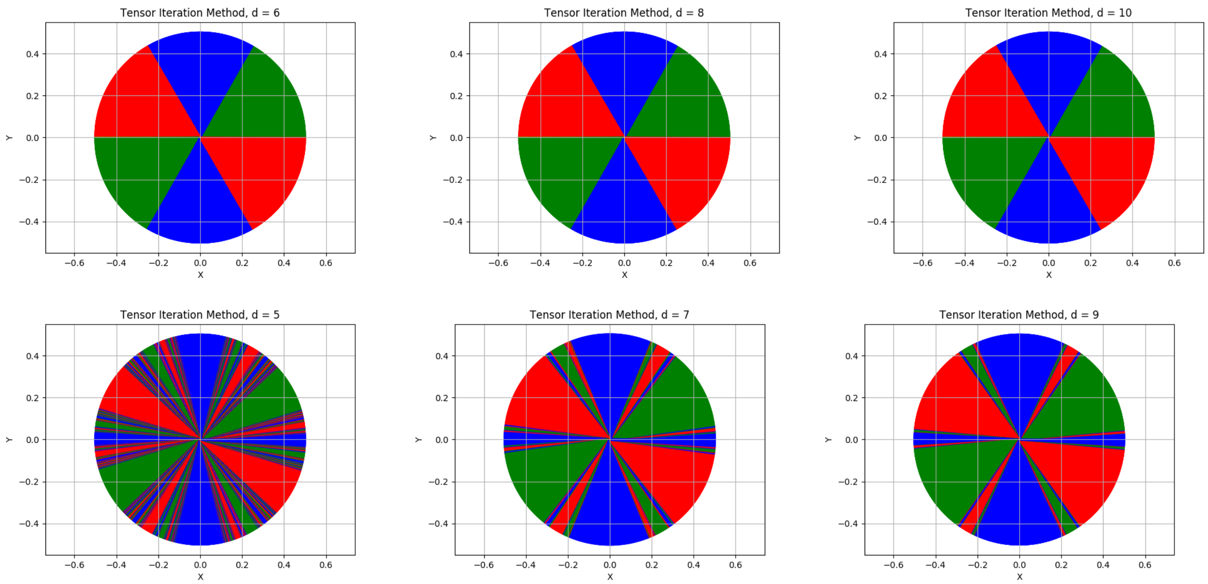

Figure 1: Convergence of vectors in $\mathbb{R}^2$ for different values of d in tensor power method.

method, we want to find a formula for the next vector, $v^{(i+1)}$. Using the definition of the tensor power method we have that:

$$\begin{aligned} v^{(i+1)} &= \begin{bmatrix} 0 \\ 1 \end{bmatrix} v_2^{(i)d-1} + \begin{bmatrix} \frac{\sqrt{3}}{2} \\ -\frac{1}{2} \end{bmatrix} \left(\frac{\sqrt{3}}{2} v_1^{(i)} - \frac{1}{2} v_2^{(i)} \right)^{d-1} + \begin{bmatrix} -\frac{\sqrt{3}}{2} \\ -\frac{1}{2} \end{bmatrix} \left(-\frac{\sqrt{3}}{2} v_1^{(i)} - \frac{1}{2} v_2^{(i)} \right)^{d-1} \\ &= \begin{bmatrix} \frac{\sqrt{3}}{2} \left[\left(\frac{\sqrt{3}}{2} v_1^{(i)} - \frac{1}{2} v_2^{(i)} \right)^{d-1} - \left(-\frac{\sqrt{3}}{2} v_1^{(i)} - \frac{1}{2} v_2^{(i)} \right)^{d-1} \right] \\ v_2^{(i)d-1} - \frac{1}{2} \left[\left(\frac{\sqrt{3}}{2} v_1^{(i)} - \frac{1}{2} v_2^{(i)} \right)^{d-1} - \left(-\frac{\sqrt{3}}{2} v_1^{(i)} - \frac{1}{2} v_2^{(i)} \right)^{d-1} \right] \end{bmatrix} \end{aligned}$$

To help simplify notation, we will define the following:

$$\begin{aligned} a_i &= \left(\frac{\sqrt{3}}{2} v_1^{(i)} - \frac{1}{2} v_2^{(i)} \right) = \langle v^{(i)}, v_2 \rangle \\ b_i &= \left(-\frac{\sqrt{3}}{2} v_1^{(i)} - \frac{1}{2} v_2^{(i)} \right) = \langle v^{(i)}, v_3 \rangle \\ c_i &= v_2^{(i)} = \langle v^{(i)}, v_1 \rangle \end{aligned}$$

Thus, we have the following formula for $v^{(i+1)}$:

$$v^{(i+1)} = \begin{bmatrix} \frac{\sqrt{3}}{2} (a_i^{d-1} - b_i^{d-1}) \\ c_i^{d-1} - \frac{1}{2} (a_i^{d-1} + b_i^{d-1}) \end{bmatrix}$$

To show convergence of the output of tensor power method we will be showing that the ratio of dot products between subsequent vectors from the tensor power method and the vectors of the ETF converge to the values we expect. For simplicity, we will demonstrate convergence for the vectors closer in angle to v_1, but the proof works for the other two cases as well:

$$\lim_{i\to\infty} \frac{\langle v_2, v^{(i+1)}\rangle}{\langle v_1, v^{(i+1)}\rangle} = -\frac{1}{2}$$
$$\lim_{i\to\infty} \frac{\langle v_3, v^{(i+1)}\rangle}{\langle v_1, v^{(i+1)}\rangle} = -\frac{1}{2}$$

We will show that the ratio of the dot products $\langle v_1, v^{(i+1)}\rangle$ and $\langle v_2, v^{(i+1)}\rangle$ converge to 1/2. By symmetry, then the ratio of the dot products $\langle v_1, v^{(i+1)}\rangle$ and $\langle v_3, v^{(i+1)}\rangle$ will also converge to 1/2.

$$\begin{aligned}\frac{\langle v_2, v^{(i+1)}\rangle}{\langle v_1, v^{(i+1)}\rangle} &= \frac{\frac{3}{4}(a_i^{d-1} - b_i^{d-1}) - \frac{1}{2}c_i^{d-1} + \frac{1}{4}(a_i^{d-1} + b_i^{d-1})}{c_i^{d-1} - \frac{1}{2}(a_i^{d-1} + b_i^{d-1})} \\ &= \frac{a_i^{d-1} - \frac{1}{2}b_i^{d-1} - \frac{1}{2}c_i^{d-1}}{-\frac{1}{2}a_i^{d-1} - \frac{1}{2}b_i^{d-1} + c_i^{d-1}} \\ &= \frac{(\frac{a_i}{c_i})^{d-1} - \frac{1}{2}(\frac{b_i}{c_i})^{d-1} - \frac{1}{2}}{-\frac{1}{2}(\frac{a_i}{c_i})^{d-1} - \frac{1}{2}(\frac{b_i}{c_i})^{d-1} + 1}\end{aligned}$$

Define the ratio α_i as follows:

$$\begin{aligned}\alpha_i &= -\frac{a_i}{c_i} \\ 1 - \alpha_i &= -\frac{b_i}{c_i}\end{aligned}$$

Thus:

$$\begin{aligned}\frac{\langle v_2, v^{(i+1)}\rangle}{\langle v_1, v^{(i+1)}\rangle} &= \frac{(\frac{a_i}{c_i})^{d-1} - \frac{1}{2}(\frac{b_i}{c_i})^{d-1} - \frac{1}{2}}{-\frac{1}{2}(\frac{a_i}{c_i})^{d-1} - \frac{1}{2}(\frac{b_i}{c_i})^{d-1} + 1} \\ &= \frac{(-\alpha_i)^{d-1} - \frac{1}{2}(-(1-\alpha_i))^{d-1} - \frac{1}{2}}{1 - \frac{1}{2}(-\alpha_i)^{d-1} - \frac{1}{2}(-(1-\alpha_i))^{d-1}} \\ &= -\alpha_{i+1}\end{aligned}$$

Where the last step holds by the definition of α_i.

We will split the proof of convergence into two cases, when d is even and when d is odd. We will first take care of the d even case. When d is even, for α_1 we have that $0 < \alpha_1 < 1$ and $0 < 1 - \alpha_1 < 1$ since our starting vector $v^{(1)}$ in our problem statement is closest in angle to v_1. In addition, if the previous inequalities hold for some α_i, then they also hold for α_{i+1}. This is because:

$$\begin{aligned}\alpha_{i+1} &= -\frac{(-\alpha_i)^{d-1} - \frac{1}{2}(-(1-\alpha_i))^{d-1} - \frac{1}{2}}{1 + \frac{1}{2}\alpha_i^{d-1} + \frac{1}{2}(1-\alpha_i)^{d-1}} \\ &= \frac{\alpha_i^{d-1} - \frac{1}{2}(1-\alpha_i)^{d-1} + \frac{1}{2}}{1 + \frac{1}{2}\alpha_i^{d-1} + \frac{1}{2}(1-\alpha_i)^{d-1}}\end{aligned}$$

and the numerator is positive when d is even by induction, the denominator is also positive by induction. To see that $\alpha_{i+1} < 1$, note that this is equivalent to

$$\alpha_i^{d-1} - \frac{1}{2}(1-\alpha_i)^{d-1} + \frac{1}{2} < 1 + \frac{1}{2}\alpha_i^{d-1} + \frac{1}{2}(1-\alpha_i)^{d-1},$$

or equivalently

$$0 < \frac{1}{2} - \frac{1}{2}\alpha_i^{d-1} + (1-\alpha_i)^{d-1},$$

which is also true by induction.

Intuitively, the inequalities simply state that if our starting vector $v^{(1)}$ of the tensor power method is closer in angle to v_1 than the other angles in the ETF, then all $v^{(i)}$ (outputs of the tensor power method) will also be closer in angle to v_1 than the other angles in the ETF.

We now show that α_i converges to $\frac{1}{2}$, which means that $v^{(i)}$ converges to $\pm v_1$.

Note that

$$\begin{aligned}
\alpha_{i+1} - \frac{1}{2} &= \frac{\alpha_i^{d-1} - \frac{1}{2}(1-\alpha_i)^{d-1} + \frac{1}{2}}{1 + \frac{1}{2}\alpha_i^{d-1} + \frac{1}{2}(1-\alpha_i)^{d-1}} - \frac{1}{2} \\
&= \frac{2\alpha_i^{d-1} - (1-\alpha_i)^{d-1} + 1 - 1 - \frac{1}{2}\alpha_i^{d-1} - \frac{1}{2}(1-\alpha_i)^{d-1}}{2(1 + \frac{1}{2}\alpha_i^{d-1} + \frac{1}{2}(1-\alpha_i)^{d-1})} \\
&= \frac{\frac{3}{2}\alpha_i^{d-1} - \frac{3}{2}(1-\alpha_i)^{d-1}}{2(1 + \frac{1}{2}\alpha_i^{d-1} + \frac{1}{2}(1-\alpha_i)^{d-1})} = \frac{3}{2}\frac{(2\alpha_i - 1)(\alpha_i^{d-2} + \cdots + (1-\alpha_i)^{d-2})}{2 + \alpha_i^{d-1} + (1-\alpha_i)^{d-1}} \\
&= \left(\alpha_i - \frac{1}{2}\right)\frac{3(\alpha_i^{d-2} + \alpha_i^{d-3}(1-\alpha_i)\cdots + (1-\alpha_i)^{d-2})}{2 + \alpha_i^{d-1} + (1-\alpha_i)^{d-1}}
\end{aligned}$$

Now, observe that $\frac{3(\alpha_i^{d-2} + \alpha_i^{d-3}(1-\alpha_i)\cdots + (1-\alpha_i)^{d-2})}{2 + \alpha_i^{d-1} + (1-\alpha_i)^{d-1}} > 0$. Therefore, if $\alpha_i > \frac{1}{2}$, then $\alpha_{i+1} > \frac{1}{2}$, and if $\alpha_i < \frac{1}{2}$, then $\alpha_{i+1} < \frac{1}{2}$. Thus, the sequence $\{\alpha_i\}$ stays on the same side of $\frac{1}{2}$, meaning that if $\alpha_i > \frac{1}{2}$, then the sequence $\{\alpha_i\}$ is bounded below by $\frac{1}{2}$ and if $\alpha_i < \frac{1}{2}$, then the sequence $\{\alpha_i\}$ is bounded above by $\frac{1}{2}$.

We will now show that $(\alpha_{i+1} - \frac{1}{2}) - (\alpha_i - \frac{1}{2}) = (\alpha_i - \frac{1}{2}) \times (\text{negative number})$. Essentially, this will show that the distance of terms in the sequence $\{\alpha_i\}$ from $\frac{1}{2}$ is decreasing. Thus, this would imply that $\alpha_i \to \frac{1}{2}$ as $i \to \infty$.

$$
\begin{aligned}
&(\alpha_{i+1} - \frac{1}{2}) - (\alpha_i - \frac{1}{2}) \\
&= \frac{3}{2}\frac{\alpha_i^{d-1} - (1-\alpha_i)^{d-1}}{2 + \alpha_i^{d-1} + (1-\alpha_i)^{d-1}} - (\alpha_i - \frac{1}{2}) = \\
&= \frac{\frac{3}{2}\alpha_i^{d-1} - \frac{3}{2}(1-\alpha_i)^{d-1} - 2\alpha_i - \alpha_i^d - \alpha_i(1-\alpha_i)^{d-1} + 1 + \frac{1}{2}\alpha_i^{d-1} + \frac{1}{2}(1-\alpha_i)^{d-1}}{2 + \alpha_i^{d-1} + (1-\alpha_i)^{d-1}} \\
&= \frac{2\alpha_i^{d-1} - (1-\alpha_i)^{d-1} - 2\alpha_i + 1 - \alpha_i^d - \alpha_i(1-\alpha_i)^{d-1}}{2 + \alpha_i^{d-1} + (1-\alpha_i)^{d-1}} \\
&= \left(\alpha_i - \frac{1}{2}\right)\left(\frac{1}{2 + \alpha_i^{d-1} + (1-\alpha_i)^{d-1}}\right)[-2 + 4(\alpha_i^{d-2} + \cdots + (1-\alpha_i)^{d-2}) \\
&\quad - 2(1-\alpha_i)^{d-1} - 2\alpha_i(\alpha_i^{d-2} + \cdots + (1-\alpha_i)^{d-2})] \\
&= 2\left(\alpha_i - \frac{1}{2}\right)\frac{(2-\alpha_i)(\alpha_i^{d-2} + \cdots + (1-\alpha_i)^{d-2}) - (1 + (1-\alpha_i)^{d-1})}{2 + \alpha_i^{d-1} + (1-\alpha_i)^{d-1}} \\
&= 2\left(\alpha_i - \frac{1}{2}\right)(2-\alpha_i)\frac{(\alpha_i^{d-2} + \cdots + (1-\alpha_i)^{d-2}) - (1 - (1-\alpha_i) + \cdots + (1-\alpha_i)^{d-2})}{2 + \alpha_i^{d-1} + (1-\alpha_i)^{d-1}}
\end{aligned}
$$

Now, since $2 - \alpha_i > 0$, we will show that the term $(\alpha_i^{d-2} + \cdots + (1-\alpha_i)^{d-2}) - (1 - (1 - \alpha_i) + \cdots + (1-\alpha_i)^{d-2})$ is negative.

$$
\begin{aligned}
&(\alpha_i^{d-2} + \cdots + (1-\alpha_i)^{d-2}) - (1 - (1-\alpha_i) + \cdots + (1-\alpha_i)^{d-2}) = \\
&= (\underbrace{\alpha_i^{d-2} + \alpha_i^{d-3}(1-\alpha_i)} + \cdots + \underbrace{\alpha_i^2(1-\alpha_i)^{d-4} + \alpha_i(1-\alpha_i)^{d-3}} + \underline{(1-\alpha_i)}^{d-2}) \\
&- (\underbrace{1 - (1-\alpha_i)} + \cdots + \underbrace{(1-\alpha_i)^{d-4} - (1-\alpha_i)^{d-3}} + \underline{(1-\alpha_i)}^{d-2}) \\
&= (\alpha_i^{d-3} + \alpha_i^{d-5}(1-\alpha_i)^2 + \cdots + \alpha_i(1-\alpha_i)^{d-4}) - (\alpha_i + \alpha_i(1-\alpha_i)^2 + \cdots + (1-\alpha_i)^{d-4}) \\
&= \underbrace{\alpha_i^{d-3} - \alpha_i}_{\leq 0} + \underbrace{\alpha_i^{d-5}(1-\alpha_i)^2 - \alpha_i(1-\alpha_i)^2}_{\leq 0} + \cdots + \underbrace{\alpha_i(1-\alpha_i)^{d-4} - (1-\alpha_i)^{d-4}}_{\leq 0} \leq 0,
\end{aligned}
$$

which proves that $(\alpha_i - \frac{1}{2})$ is always decreasing in absolute value for d even.

Now, if $d > 4$, we see that the last expression is strictly less than 0, which means that $\alpha_i \to 0$. However, if $d = 4$, the last expression equals:

$$
\begin{aligned}
&2\left(\alpha_i - \frac{1}{2}\right)(2-\alpha_i)\frac{\alpha_i^2 + \alpha_i(1-\alpha_i) + (1-\alpha_i)^2 - (1 - (1-\alpha_i) + (1-\alpha_i)^2)}{2 + \alpha_i^{d-1} + (1-\alpha_i)^{d-1}} = \\
&= 2\left(\alpha_i - \frac{1}{2}\right)(2-\alpha_i)\frac{\alpha_i^2 - \alpha_i + 1 - (\alpha_i^2 - \alpha_i + 1)}{2 + \alpha_i^{d-1} + (1-\alpha_i)^{d-1}} \\
&= 0
\end{aligned}
$$

Thus, we have that $(\alpha_{i+1} - \frac{1}{2}) - (\alpha_i - \frac{1}{2}) = 0$ for $d = 4$, which implies that $\alpha_{i+1} = \alpha_i$. Recall that we defined α_i to be the ratio of two different dot products between the ith iteration of the tensor power method and the vectors $\left[\frac{\sqrt{3}}{2}, -\frac{1}{2}\right]^T$ and $[0, 1]^T$ from our ETF. Our result implies that the angles between the ith iteration and two different fixed points are unchanging. Thus, when $d = 4$, we have that all points in $\mathbb{R}^2$ become fixed points in the tensor power method (The input vectors converge to themselves). [Kevin, please plug in $d = 4$ in the expression above, and show that it equals 0. Then explain why this leads to every point being fixed for the power method. Please, also compute the actual tensor $T = v_1^{\otimes 4} + v_2^{\otimes 4} + v_3^{\otimes 4}$ by plugging in the Mercedes frame for v_1, v_2, v_3, and maybe see that T is the diagonal "identity" tensor?]

We will now show convergence for the d odd case. Recall our formula for α_{i+1}:

$$\alpha_{i+1} = -\frac{(-\alpha_i)^{d-1} - \frac{1}{2}(-(1-\alpha_i))^{d-1} - \frac{1}{2}}{1 - \frac{1}{2}(-\alpha_i)^{d-1} - \frac{1}{2}(-(1-\alpha_i))^{d-1}}$$

When d is odd, we have that:

$$\begin{aligned}
\alpha_{i+1} &= \frac{\frac{1}{2} - \alpha_i^{d-1} + \frac{1}{2}(1-\alpha_i)^{d-1}}{1 - \frac{1}{2}\alpha_i^{d-1} - \frac{1}{2}(1-\alpha_i)^{d-1}} && (2.1)\\
\implies \alpha_{i+1} - \frac{1}{2} &= \frac{\frac{1}{2} - \alpha_i^{d-1} + \frac{1}{2}(1-\alpha_i)^{d-1}}{1 - \frac{1}{2}\alpha_i^{d-1} - \frac{1}{2}(1-\alpha_i)^{d-1}} - \frac{1}{2} && (2.2)\\
&= \frac{1 - 2\alpha_i^{d-1} + (1-\alpha_i)^{d-1} - (1 - \frac{1}{2}\alpha_i^{d-1} - \frac{1}{2}(1-\alpha_i)^{d-1})}{2 - \alpha_i^{d-1} - (1-\alpha_i)^{d-1}} && (2.3)\\
&= \frac{-\frac{3}{2}\alpha_i^{d-1} + \frac{3}{2}(1-\alpha_i)^{d-1}}{2 - \alpha_i^{d-1} - (1-\alpha_i)^{d-1}} && (2.4)\\
&= \left(\alpha_i - \frac{1}{2}\right)\left(\frac{-3(\alpha_i^{d-2} + \alpha_i^{d-3}(1-\alpha_i)\cdots + (1-\alpha_i)^{d-2})}{2 - \alpha_i^{d-1} - (1-\alpha_i)^{d-1}}\right) && (2.5)
\end{aligned}$$

Because we want to show that the sequence of α_i converges to $\frac{1}{2}$, it would suffice to show that the fractional term in the above formula for α_{i+1} is less than one in absolute value. Unfortunately, unlike the d even case, $0 < \alpha_i < 1$ does not imply that $0 < \alpha_{i+1} < 1$. Instead, we will have a smaller range of starting values for which α_i will converge. Our strategy will be to first show that for $d = 5$, there exists a range of values of α_i for which the fractional term in (1.5) has an absolute value of less than 1, and then show that as d increases, the value of the fraction is strictly decreasing, meaning that we will always have a range of α_i values such that the sequence converges to $\frac{1}{2}$.

As seen in figure 1.2, we have when $\alpha_i \in [0.3, 0.7]$ then $\frac{-3(\alpha_i^{d-2} + \alpha_i^{d-3}(1-\alpha_i)\cdots + (1-\alpha_i)^{d-2})}{2 - \alpha_i^{d-1} - (1-\alpha_i)^{d-1}} < 1$. We will show that this interval widens by showing that $\frac{-3(\alpha_i^{d-2} + \alpha_i^{d-3}(1-\alpha_i)\cdots + (1-\alpha_i)^{d-2})}{2 - \alpha_i^{d-1} - (1-\alpha_i)^{d-1}}$ decreases as d increases.

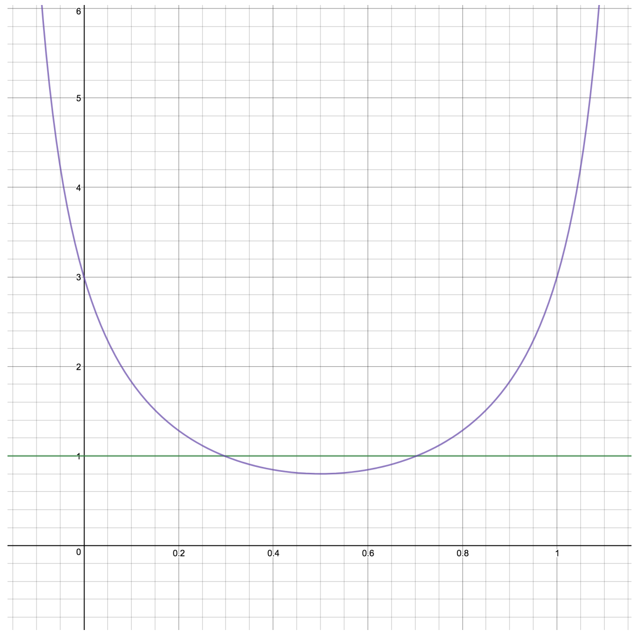

Figure 2: Graph of $\frac{-3(\alpha_i^{d-2}+\alpha_i^{d-3}(1-\alpha_i)\cdots+(1-\alpha_i)^{d-2})}{2-\alpha_i^{d-1}-(1-\alpha_i)^{d-1}}$ when $d = 5$.

We will now prove that as d increases, the value of the fractional term in (1.5) strictly decreases when α_i is in the interval $[0, 1]$.

$$\left(\frac{3(\alpha_i^{d-2}+\alpha_i^{d-3}(1-\alpha_i)\cdots+(1-\alpha_i)^{d-2})}{2-\alpha_i^{d-1}-(1-\alpha_i)^{d-1}}\right)$$

Showing that the entire fraction decreases as d increases is equivalent to demonstrating that the denominator increases and the numerator decreases as d increases. It is clear that the denominator increases as d increases, since $0 < \alpha_i < 1$. We will now show that the numerator $3\sum_{j=0}^{d-2}\alpha_i^j(1-\alpha_i)^{d-2-j}$ decreases as d increases. First, substitute d' for $d-2$. We will show then that $3\sum_{j=0}^{d'+1}\alpha_i^j(1-\alpha_i)^{d'+1-j} - 3\sum_{j=0}^{d'}\alpha_i^j(1-\alpha_i)^{d'-j} < 0$. Indeed,

$$3\sum_{j=0}^{d'+1}\alpha_i^j(1-\alpha_i)^{d'+1-j} - 3\sum_{j=0}^{d'}\alpha_i^j(1-\alpha_i)^{d'-j} = 3\sum_{j=0}^{d'}\alpha_i^j(1-\alpha_i)^{d'-j}((1-\alpha_i)-1) + 3\alpha_i^{d'+1} =$$

$$= -3\sum_{j=0}^{d'}\alpha_i^{j+1}(1-\alpha_i)^{d'-j} + 3\alpha_i^{d'+1} = -3\sum_{j=1}^{d'+1}\alpha_i^j(1-\alpha_i)^{d'+1-j} + 3\alpha_i^{d'+1} =$$

$$= -3\sum_{j=1}^{d'}\alpha_i^j(1-\alpha_i)^{d'+1-j} < 0,$$

whenever $d \geq 1$ and $0 < \alpha_i < 1$.

Therefore, the numerator decreases as d increases.

3 Mercedes-Benz decomposable tensors in higher dimensions

Definition 3.1. A *Mercedes-Benz frame* in $\mathbb{R}^n$ is a set of $n+1$ vectors $v_1, \ldots, v_{n+1} \in \mathbb{R}^n$ such that, for all $i \neq j$,

$$\langle v_i, v_j \rangle = -\frac{1}{n}.$$

In this section we show that if a tensor decomposes via such an ETF, then the elements of the ETF are set of robust eigenvectors for this tensor.

4 General ETF decomposable tensors

References

[1] A. Anandkumar, R. Ge, D. Hsu, S. Kakade, and M. Telegarsky: *Tensor Decompositions for Learning Latent Variable Models*

[2] A. Anandkumar, R. Ge, and M. Janzamin: *Learning Overcomplete Latent Variable Models through Tensor Methods*

[3] L. Oeding, E. Robeva, and B. Sturmfels: *Decomposing Tensors into Frames*, Advances in Applied Mathematics, 76 (2016), pp. 125-153

[4] E. Robeva. *Orthogonal Decomposition of Symmetric Tensors.* SIAM Journal on Matrix Analysis and Applications, 37 (2016), pp. 86-102