Tensor Ranks and Norms

by

Robert Cochrane

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Mathematics)
in The University of Michigan
2022

Doctoral Committee:

Professor Harm Derksen, Co-Chair
Professor Mattias Jonsson, Co-Chair
Professor Sarah Koch
Professor Kayvan Najarian

Robert Cochrane

robcoch@umich.edu

ORCID iD: 0000-0001-5809-8045

ACKNOWLEDGEMENTS

This thesis was only made possible by the support of many, many people. It is not possible to list them all, never mind all the contributions they made. However, I would like to give mention to just a few of them.

Thanks go to my friends and family from Scotland for their support and encouragement which helped me to make it to Michigan, and for their work in keeping me sane while I was there.

At Michigan, Alana, Francesca, Visu, and Zhi all inspired me to work with tensors, and provided useful insights. Many others there were also instrumental to my development as a mathematician and as a human, including but not limited to Alex, Christina, David, Mitul, Nawaz, and Rishi.

I'd also like to thank Pat and Paul for teaching alongside me, and in the process, helping me to find my footing in a profession that I love.

I am appreciative to all my thesis committee for their support, their advice, and for their flexibility when I didn't keep to schedule. Professor Jonsson in particular stepped up to act as my advisor when I was left without one at Michigan, and I am very thankful.

Most of all, I would like to thank my advisor, Harm Derksen. Throughout the challenges I have faced in graduate school, Harm was patient, kind, and empathetic. He was always willing to share his knowledge, his insight, and his time. I cannot envision a path through my PhD without the suport that he gave me. I am so grateful.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

ABSTRACT

Tensor decompositions have been studied for nearly a century, but the well-known notion of tensor rank does not capture all the properties that one may desire of a rank function on tensors. For that reason, in recent years many alternative notions of tensor rank have been developed and studied. Some have also studied related notions of norms on tensors, especially the nuclear norm, which can be viewed as a convex relaxation of tensor rank.

In this thesis, we calculate the value of the recently introduced G-stable rank for all weights on $2 \times 2 \times 2$ and $2 \times 2 \times 3$ complex-valued tensors, and introduce X-rank, which can be viewed as a refinement of G-stable rank.

We then investigate the nuclear norm, and how it and some other norms on tensor product spaces behave with respect to the vertical, or Kronecker, tensor product.

Finally, we introduce some notions of stable ranks on tensors, built from common norms on tensor products, and discuss how these stable ranks relate to other notions of tensor rank.

CHAPTER I

Introduction

1.1 History and Motivation

One of the most fundamental concepts in linear algebra is that of matrix rank. In some sense, rank measures the "amount of information" contained within a matrix. For this reason, we are often motivated to find the rank of a matrix, or to approximate a matrix by one of lower rank. This occurs in many subfields of mathematics - signal processing, numerical linear algebra, computer vision, and many others.

While data from these fields can often be classified using two indices (the rows and columns of a matrix), it is often desirable to have more indices for the data. For example, an image may be described using a matrix where each entry represents the color of a single pixel in a given row and column, but processing a video may require an extra index to represent the color of a pixel in a given row and column at a certain point in time. Data with multiple indices may be arranged into a mathematical object known as a tensor, which can be seen as a generalization of vectors and matrices.

As with matrices, mathematicians are motivated to measure the "amount of information" contained within tensors, and so are motivated to generalize the concept of matrix rank to higher orders. Such problems were studied as early as the 1920's (*Hitchcock* (1927)), and were of interest in the fields of psychometrics (*Tucker* (1966)) and chemometrics (*Appellof and Davidson* (1981)) before gaining more traction in the subfields of mathematics discussed above (e.g. *De Lathauwer and De Moor* (1998), *De Lathauwer et al.* (2000), *Vasilescu and Terzopoulos* (2002)).

Psychometricians were among the first to make use of CP (*Canonical Polyadic*) decompositions (which we will refer to as *rank decompositions*) - a decomposition of a rank r tensor into a sum of r rank one tensors. In chemometrics, these decompostions were studied as a tool in fluorescent spectroscopy - lights of different wavelengths are shined into n samples of a mixture of chemicals, and the intensity of different wavelengths of emitted light is measured. This data naturally forms a tensor, and each molecule which appears in the mixture has its own characteristic fluorescent spectrum, which corresponds to a rank one tensor. If there are r chemicals present in the mixture, then the resulting tensor should have rank r. The CP decomposition further gives information about the relative concentration of different compounds in the mixture. For a more detailed history, *Kolda and Bader* (2009) is an excellent resource.

Unfortunately, it soon became obvious that the most standard generalization of matrix rank to higher order tensors, known as tensor rank, is not as well-behaved as matrix rank. For example, a tensor may have different rank when its entries are considered to be elements of $\mathbb{R}$ as opposed to elements of $\mathbb{C}$ (see Example 1.4), some tensors can be arbitrarily well-approximated by tensors of lower rank (see Example 1.5), and even computing the rank of a tensor can be an NP-hard problem (*Håstad* (1990)). For these reasons, we are motivated to find alternative notions of rank for tensors which may not have all of the same issues.

In Chapter 2 we explore the recently introduced G-stable α-rank, and in particular, we calculate its values for all tensors of small dimension, in Proposition 2.20. We also offer an unweighted version of this rank, which we have called the X-rank, and discuss some of its basic properties.

As rank, even for matrices, is not a continuous function, the nuclear norm, which can be viewed as a convex relaxation of rank, is sometimes used a substitute (*Candès and Recht* (2012)). Much as when the l_1 norm is used a substitute for sparsity (the 'l_0 norm'), the nuclear norm is generally easier to work with, and produces good results with high probability, and so if one is interested in tensor rank then they are motivated to understand the nuclear norm for tensors too.

In Chapter 3, we explore some questions about the nuclear norm which have classically been posed about tensor rank, and in particular show it is multiplicative under the tensor Kronecker product (see Proposition 3.6). We also show that a similar property holds for a wider class of norms (see Proposition 3.14).

In the final Chapter, we explore the relationship between this nuclear norm, and the more commonly known spectral and Frobenius norms, and use these to develop some new notions of rank which unlike norms are invariant under multiplication by scalars, and which are more 'stable' than some other notions of rank in that the ranks in the small neighborhood of a tensor are not expected to vary wildly. We give some bounds on low-rank approximations to tensors depending on these stable ranks in Proposition 4.21.

1.2 Elementary Notions of Tensor Products and Notation

We will mostly follow the notation of Derksen (*Derksen* (2016)). In particular, a *d-th order tensor product space* is a pair $\mathbf{U} = (U, (U^{(1)}, \ldots, U^{(d)}))$ where each $U^{(i)}$ is a finite dimensional vector space and

$$U = U^{(1)} \otimes \cdots \otimes U^{(d)}.$$

For ease of notation, we will sometimes write $U^{(i)}$ as U_i.

A tensor $u \in \mathbf{U}$ with $u = u^{(1)} \otimes \ldots \otimes u^{(d)}$ where $u^{(i)} \in U^{(i)}$ for all i is said to be a *pure* or *simple* tensor.

We will often denote the standard basis vectors in K^n by $e_1, e_2, \ldots, e_n$ or sometimes as $[1], [2], \ldots, [n]$. Under this shorthand, we may abbreviate the tensor $[i_1]\otimes[i_2]\otimes\ldots[i_d]$ by $[i_1, i_2, \ldots, i_d]$.

For a tensor product space $\mathbf{U}$, and $S, T \in \mathbf{U}$, we will denote the Frobenius inner product of S and T by $\langle S, T\rangle$, and the associated Frobenius norm (sometimes also known as the Euclidean norm) of S by

$$\|S\| := \sqrt{|\langle S, S\rangle|}.$$

For a tensor S in a tensor product space $\mathbf{U}$, we may write S as a sum of pure (or simple) tensors

$$S = \sum_{i=1}^{r} v_i \text{ where } v_i = v_i^{(1)} \otimes \ldots \otimes v_i^{(d)} \text{ and } v_i^{(e)} \in U^{(e)}. \tag{1}$$

We define the *nuclear norm* of S to be the infimum of $\sum_{i=1}^{r} \|v_i\|$ over all such decompositions (1), and denote it by $\|S\|_\star$. It can be shown that this infimum can always be achieved (*Friedland and Lim* (2018)), and so we may take the minimum over such decompositions instead.

The *spectral norm* of S, denoted $\|S\|_\sigma$, is defined to be the maximum value of $|\langle S, u\rangle|$ where u ranges over all pure tensors of unit length. More generally, as in *Derksen* (2016), for an r-tuple of tensors, $\mathbf{S} = (S_1, \ldots S_r)$, and $1 \leq \alpha < \infty$, we will define $[\mathbf{S}]_\alpha$ as the maximum of

$$\left(\sum_{i=1}^{r} |\langle S_i, u\rangle|^\alpha\right)^{1/\alpha}$$

over all pure tensors u of unit length. For $\alpha = \infty$, we will define $[\mathbf{S}]_\alpha$ as the maximum over all pure tensors u of unit length of $\max_i |\langle S_i, u\rangle|$, or equivalently, as $\max_i \|S_i\|_\sigma$.

If $\mathbf{U} = (U, (U^{(1)}, \ldots, U^{(d)}))$ and $\mathbf{V} = (V, (V^{(1)}, \ldots, V^{(d)}))$ are tensor product spaces, then we define their *vertical tensor product* as

$$\mathbf{U} \boxtimes \mathbf{V} = \left(U \otimes V, (U^{(1)} \otimes V^{(1)}, \ldots, U^{(d)} \otimes V^{(d)})\right).$$

For $S \in \mathbf{U}$, $T \in \mathbf{V}$, we define $S \boxtimes T$ to be $S \otimes T$, viewed as an element of $\mathbf{U} \boxtimes \mathbf{V}$. Note that for order d tensors S and T, $S \boxtimes T$ is an order d tensor, where $S \otimes T$ has order $2d$. In particular, it is easy to see that for matrices A and B, $A \boxtimes B$ is the well-known *Kronecker product* of A and B. For this reason, the vertical tensor product is also sometimes known as the *tensor Kronecker product*.

For $S = (S_1, \ldots, S_r) \in \mathbf{U}^r$, $T = (T_1, \ldots, T_s) \in \mathbf{V}^s$, we define

$$S \boxtimes T = (S_i \boxtimes T_j | 1 \leq i \leq r, 1 \leq j \leq s).$$

If $\mathbf{U} = (U, (U^{(1)}, \ldots, U^{(d)}))$ and $\mathbf{V} = (V, (V^{(1)}, \ldots, V^{(d)}))$ are tensor product spaces, then we define their *direct sum* as

$$\mathbf{U} \oplus \mathbf{V} = \left(U \oplus V, (U^{(1)} \oplus V^{(1)}, \ldots, U^{(d)} \oplus V^{(d)})\right).$$

For $S \in \mathbf{U}$, $T \in \mathbf{V}$, we define $S \oplus T$ as an element of $\mathbf{U} \oplus \mathbf{V}$ in the obvious way.

1.3 Rank for Tensors

One of the most foundational notions in linear algebra is the rank of a matrix. There are many equivalent notions of rank for matrices. We will make use of one which involves defining, for finite dimensional vector spaces V, W over a field K, an isomorphism $V^\star \otimes W \cong \mathrm{Hom}(V, W)$:

Definition 1.1. Let V, W be finite dimensional vector spaces and denote by $V^\star$ the dual space to V. For $f \in V^\star$ and $w \in W$, we may define a linear map $f \otimes w : V \to W$

by $v \mapsto f(v)w$. Any linear map $A : V \to W$ may be described as

$$\sum_{i=1}^{r} f_i \otimes w_i$$

for some combination of f_i and w_i. The minimum value of r required is precisely the *rank* of the matrix A.

This definition may be extended from linear maps to *multilinear* maps to give us the concept of a tensor.

Definition 1.2. Let $V_1, \ldots, V_{d-1}, W$ be finite dimensional vector spaces. A function

$$S : V_1 \times \ldots \times V_{d-1} \to W$$

is said to be *multilinear* if it is linear in each factor V_i. The space of such multilinear functions is isomorphic to the tensor product space $V_1^\star \otimes \ldots \otimes V_{d-1}^\star \otimes W$.

There are many notions of rank for tensors. The most common is simply known as *tensor rank*.

Definition 1.3. Let $S \in V_1 \otimes \ldots \otimes V_d$. Then we may express

$$S = \sum_{j=1}^{r} v_{1,j} \otimes \ldots \otimes v_{d,j}$$

where we have $v_{i,k} \in V_i$ for all i, k. The minimum value of r required for such an expression of S is known as the *rank* of S, and is denoted $\mathrm{rk}(S)$. A tensor of rank 1 is said to be *simple*.

The rank for matrices has a number of useful algebraic and topological properties. A convincing generalization of rank to higher order tensors (let us call such a function Rank) would hope to satisfy some of these properties. Note that these properties all hold for matrices in the case that the tensors involved have order $d = 2$.

1. (Zero tensor) Rank(S)=0 if and only if S is the zero tensor.
2. (Scale invariance) For a nonzero scalar λ and a tensor S, Rank(λS)= Rank(S).
3. (Simple tensors) If v is a simple tensor then Rank(v)=1.
4. (Diagonal tensor) If $I_r = \sum_{i=1}^{r} e_i \otimes e_i \otimes \ldots \otimes e_i$ where e_i are unit basis vectors then Rank(I_r)=r.
5. (Whole numbers) For any tensor S, Rank(S) is a whole number.
6. (Field invariance) If $K \subseteq L$ are fields and S is a tensor over K, then its rank when considered as a tensor over K is the same as its rank when considered as a tensor over L.
7. (Basis invariance) For a tensor $S \in V_1 \otimes V_2 \otimes \ldots \otimes V_d$, and $g \in \mathrm{GL}(V_1) \times \mathrm{GL}(V_2) \times \ldots \times \mathrm{GL}(V_d)$, Rank($g \cdot S$)=Rank($S$).
8. (Rank Closure) The set $\{S : \mathrm{Rank}(S) \leq r\}$ is Zariski closed for all r.
9. (Triangle inequality) For matrices S and T, Rank($S + T$) $\leq$ Rank(S)+Rank(T).
10. (Direct sums) For matrices S and T, Rank($S \oplus T$)=Rank(S)+Rank(T).
11. (Kronecker multiplicativity) If $S \boxtimes T$ is the Kronecker product of S and T then Rank($S \boxtimes T$)=Rank(S)Rank(T).

Note that some of these properties are stronger than others - for example, property 4 is implied if properties 3 and 10 hold.

One additional property that we might hope holds for a rank function on tensors of higher order is:

12. (Tensor multiplicativity) If $S \otimes T$ is the usual tensor product of S and T then Rank($S \otimes T$)=Rank(S)Rank(T).

While there are many other criteria that we may desire a rank function to satisfy, even this list is not known to be fully satisfied by any notion of tensor rank. While our usual notion of tensor rank satisfies properties 1-5, 7, 9, and 10, it does not satisfy the others. Some well-known counter-examples for some of these properties are listed below:

Example 1.4. Let $U^{(1)} = U^{(2)} = U^{(3)} = \mathbb{R}^2$ and let $S \in U^{(1)} \otimes U^{(2)} \otimes U^{(3)}$ be given by

$$[112] + [121] + [211] - [222].$$

It is not hard to show that this S has rank 3. However, if instead $U^{(1)} = U^{(2)} = U^{(3)} = \mathbb{C}^2$ and $S \in U^{(1)} \otimes U^{(2)} \otimes U^{(3)}$ is defined in the same way as before, we may rewrite S as

$$\frac{1}{2i} \left((e_1 + ie_2) \otimes (e_1 + ie_2) \otimes (e_1 + ie_2) - (e_1 - ie_2) \otimes (e_1 - ie_2) \otimes (e_1 - ie_2) \right),$$

and hence the rank of S is 2. Thus tensor rank does not satisfy Field invariance property.

Example 1.5. Let the (real or complex) tensor S_t be defined by

$$S_t = [112] + [121] + [211] + t\left([122] + [212] + [221]\right) + t^2[222].$$

For $t \neq 0$, we may rewrite S_t as

$$\frac{1}{t} \left((e_1 + te_2) \otimes (e_1 + te_2) \otimes (e_1 + te_2) - e_1 \otimes e_1 \otimes e_1 \right),$$

which has rank 2. However, for $t = 0$, we have

$$S_0 = [112] + [121] + [211],$$

which has rank 3. Hence the tensor S_0 has rank 3, but lies in the Zariski closure of the tensors of rank less than or equal to 2.

As the usual tensor rank does not satisfy all the properties one might hope it would, many alternate notions of tensor rank have also been studied. The concept of *border rank* in particular was developed as a version of rank which does satisfy the Rank Closure property (originally in *Bini et al.* (1979), see *Bürgisser et al.* (1997), *Landsberg* (2012) for more recent developments).

Definition 1.6. The *border rank* of a tensor S, denoted by $\mathrm{brk}(S)$, is the smallest value of r such that there exists a sequence of tensors of rank r whose limit is S.

Equivalently, if $X(r) \subseteq V_1 \otimes \ldots \otimes V_d$ denotes the Zariski closure of the set of tensors of rank at most r, then $\mathrm{brk}(S)$ is the smallest value of r such that $S \in X(r)$.

As we saw in Example 1.5, the tensor $S = [112] + [121] + [211]$ has rank 3, but border rank 2. The tensor $S \oplus S$ has rank 6 and border rank 4, and in general $S^{\oplus n}$ has rank $3n$ and border rank $2n$, so the difference between rank and border rank can be arbitrarily large.

Another popular notion of rank, introduced recently (*Tao and Sawin* (2016)), is *slice rank*.

Definition 1.7. A nonzero tensor $S \in V_1 \otimes \ldots \otimes V_d$ has *slice rank* 1 if it is contained in

$$V_1 \otimes \cdots \otimes V_{i-1} \otimes v \otimes V_{i+1} \otimes \cdots \otimes V_d$$

for some i and $v \in V_i$. The slice rank of an arbitrary tensor $S \in V_1 \otimes \ldots \otimes V_d$, denoted by $\mathrm{srk}(S)$ is the smallest value of r such that S is the sum of r tensors of slice rank 1.

The motivation for slice rank was found in algebraic combinatorics, where the slice rank of a particular tensor was used to give an alternate proof of the best-known upper bounds in the capset problem, from *Ellenberg and Gijswijt* (2017). This motivated

Derksen to introduce a new notion of rank, the G-stable rank, which was used to further improve those upper bounds (*Derksen* (2020)).

In Chapter II, we investigate the G-stable rank in more detail, and calculate the value of the G-stable rank of all possible weights on tensors of small dimension. We also introduce a new notion of rank inspired by the G-stable rank, which we call "X-rank", and prove that it has some of our desired properties for a rank function.

As most versions of rank are difficult to calculate, authors often prefer to work with the nuclear norm of tensors. The nuclear norm can be seen as a convex relaxation of rank, in the sense that the convex hull of unit simple tensors forms the unit ball for the nuclear norm. In Chapter III, we investigate the nuclear, and related, norms on tensor product spaces, and prove that the nuclear norm has the Kronecker multiplicativity property mentioned above.

Extending this further, we introduce in Chapter IV some new notions of "stable ranks" which are relatively easy to calculate or approximate compared to more conventional notions of rank. The tradeoff is that these do not satisfy even some of the more basic notions of rank (not even the triangle inequality).

CHAPTER II

G-Stable Rank

Derksen (*Derksen* (2020)) introduced the notion of the G-stable α-rank, $\mathrm{rk}_\alpha^G(S)$, of a tensor S over a perfect field K. We reproduce this definition below, restricted to the case where K is algebraically closed. It makes use of the ring $K[[t]]$ of formal power series and its quotient field $K((t))$ of formal Laurent series, and *valuations* of its elements.

Definition 2.1. The *t-valuation* of a power series $a(t) \in K((t))$, denoted by $\mathrm{val}_t(a(t))$, is the largest integer d such that $a(t) = t^d b(t)$ for some $b(t) \in K[[t]]$. By convention, $\mathrm{val}_t(0) = \infty$.

If W is a K-vector space and $v(t) \in K((t)) \otimes W$, then we define

$$\mathrm{val}_t(v(t)) = \max\{d \mid v(t) = t^d w(t) \text{ and } w(t) \in K[[t]] \otimes W\}.$$

We have $\mathrm{val}(v(t)) \geq 0$ if and only if $w(t) \in K[[t]] \otimes W$, and in this case, we say $v(t)$ has no poles. We say that $\lim_{t\to 0} v(t)$ exists, and is equal to $v(0) \in W$.

We denote by $\mathrm{GL}(W, K((t)))$ the group of invertible $K((t))$-linear endomorphisms of the space $K((t)) \otimes W$. This group can be viewed as a subset of $K((t)) \otimes \mathrm{End}(W)$. For $W = K^n$, we have $K((t)) \otimes W \cong K((t))^n$ and we can identify $\mathrm{GL}(W, K((t)))$ with the set of $n \times n$ matrices with entries in the field $K((t))$. For any K-subalgebra

$R \subseteq K(\!(t)\!)$ of $K(\!(t)\!)$, then we denote by $\mathrm{GL}(W, R)$ the intersection of $\mathrm{GL}(W, K(\!(t)\!))$ with $R \otimes \mathrm{End}(W)$ in $K(\!(t)\!) \otimes \mathrm{End}(W)$. It is important to note that the elements of $\mathrm{GL}(W, R)$ must have an inverse in $\mathrm{GL}(W, K(\!(t)\!))$, but not necessarily in $\mathrm{GL}(W, R)$.

We will consider the action of $G = \mathrm{GL}(V_1) \times \mathrm{GL}(V_2) \times \ldots \times \mathrm{GL}(V_d)$ on the tensor product space $\mathbf{V} = V_1 \otimes V_2 \otimes \ldots \otimes V_d$. For any K-subalgebra $R \subseteq K(\!(t)\!)$, we may define

$$G(R) = \mathrm{GL}(V_1, R) \times \ldots \times \mathrm{GL}(V_d, R).$$

The group $G\left(K(\!(t)\!)\right)$ acts on $K(\!(t)\!) \otimes \mathbf{V}$. We will call any nonzero $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_d) \in \mathbb{R}^d_{\geq 0}$ a *weight*. Let $g(t) = (g_1(t), g_2(t), \ldots, g_d(t)) \in G\left(K[[t]]\right)$, $S \in \mathbf{V}$, and $\mathrm{val}_t(g(t) \cdot S) > 0$. Then we may consider the *slope*,

$$\mu_\alpha(g(t), S) = \frac{\sum_{i=1}^{d} \alpha_i \, \mathrm{val}_t(\det g_i(t))}{\mathrm{val}_t(g(t) \cdot S)}.$$

In general, a small slope means that S is unstable, in the sense that as $t \to 0$, $g(t) \cdot S$ goes to 0 quickly in comparison to the eigenvalues of $g_i(t)$. It is this property which leads to the notion of G-stable α-rank.

Definition 2.2. For a fixed nonzero weight $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_d) \in \mathbb{R}^d_{\geq 0}$ and $S \in \mathbf{V}$, the *G-stable α-rank* of S, denoted by $\mathrm{rk}^G_\alpha(S)$, is the infimum of all $\mu_\alpha(g(t), S)$ where $g(t) \in G\left(K[[t]]\right)$ and $\mathrm{val}_t(g(t) \cdot S) > 0$.

The case where $\alpha = (1, 1, \ldots, 1)$, is sometimes referred to as the *G-stable rank* of S, and is denoted by $\mathrm{rk}^G(S)$.

Derksen furthermore showed that it is sufficient to consider $g(t)$ which are 1-parameter subgroups of G without poles. Here, $g(t) = (g_1(t), g_2(t), \ldots, g_d(t)) \in G\left(K[[t]]\right)$ is a 1-parameter subgroup if for every i we can choose a basis of V_i such that the matrix of $g_i(t)$ is diagonal and each diagonal entry of that matrix is a nonnegative power of t.

This notion of rank satisfies many nice properties. However, one important difference to many other notions of rank is that the G-stable rank does not have to be a whole number:

Example 2.3. As we will show, the tensor $S = e_2 \otimes e_1 \otimes e_1 + e_1 \otimes e_2 \otimes e_1 + e_1 \otimes e_1 \otimes e_2$ has $\mathrm{rk}_\alpha^G(S) = \frac{3}{2}$.

Derksen also gave an alternate characterization of the G-stable α-rank, specifically over $\mathbb{C}$, which we will use later.

Definition 2.4. For $S \in \mathbf{V} = V_1 \otimes \ldots \otimes V_d$, let $\Phi_i(S) : (V_1 \otimes \ldots \otimes \hat{V_i} \otimes \ldots \otimes V_d)^\star \to V_i$ be the i-th flattening. For a weight $\alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{R}^d_{\geq 0}$, the *$G$-stable α-rank of S* is

$$\mathrm{rk}_\alpha^G(S) := \sup_{g \in G} \min_i \frac{\alpha_i \|g \cdot S\|^2}{\|\Phi_i(g \cdot S)\|^2_\sigma}.$$

We will aim to investigate the G-stable α-rank for tensors of small dimensions. To do this, we will start by calculating a related notion of rank for these tensors.

Recall that to find the G-stable α-rank, we look for the infimum of $\mu_\alpha(g(t), S)$ over all 1-parameter subgroups $g(t)$ with $\mathrm{val}_t(g(t) \cdot S) > 0$. Every 1-parameter subgroup is contained in some maximal torus T (which itself is contained in some Borel subgroup B of G). For that reason, we are motivated to fix a maximal torus T and consider all 1-parameter subgroups contained within it. Choosing a maximal torus of G is equivalent to choosing a basis in each vector space V_i. Let us choose bases in each V_i such that $\mathrm{GL}(V_i)$ can be identified with GL_{n_i}. Denote by T_{n_i} the subgroup of GL_{n_i} of invertible $n_i \times n_i$ diagonal matrices, and let $T = T_{n_1} \times T_{n_2} \times \cdots \times T_{n_d} \subseteq G$. Then T is a maximal torus of G. This naturally leads to Derksen's concept of T-stable α-rank:

Definition 2.5. The *T-stable α-rank* of S, denoted by $\mathrm{rk}_\alpha^T(S)$ is the infimum of $\mu_\alpha(g(t), S)$ over all 1-parameter subgroups $g(t) \in T(K[t])$ of T with $\mathrm{val}_t(g(t) \cdot S) > 0$.

Using the fact that every 1-parameter subgroup is conjugate to one in the maximal torus T, we immediately obtain:

Lemma 2.6. We have

$$\mathrm{rk}_\alpha^G(S) = \inf_{g\in G} \mathrm{rk}_\alpha^T(g \cdot S).$$

The form of T allows us to calculate the T-stable α-rank using a particular linear program. We will make use first of the notion of the T-*support* of a tensor in a fixed basis.

Definition 2.7. For a field K and a tensor $S = (s_{i_1,i_2,\ldots,i_d}) \in K^{n_1} \otimes K^{n_2} \otimes \ldots \otimes K^{n_d}$, we define its T-*support* by

$$\mathrm{supp}^T(S) = \{(i_1, i_2, \ldots, i_d) \mid s_{i_1,i_2,\ldots,i_d} \neq 0\}.$$

Definition 2.8. Let $S \in K^{n_1} \otimes K^{n_2} \otimes \ldots \otimes K^{n_d}$ be a tensor. Then the linear program $\mathbf{LP}_\alpha(\mathrm{supp}^T(S))$ is to find the minimum of $\displaystyle\sum_{i=1}^{d} \alpha_i \sum_{j=1}^{n_i} x(i,j)$ where the $x(i,j)$ are non-negative real variables with the constraints that $\displaystyle\sum_{i=1}^{d} x(i,t_i) \geq 1$ for all $t = (t_1, t_2, \ldots, t_d) \in \mathrm{supp}^T(S)$.

Derksen showed in his paper that the value of $\mathbf{LP}_\alpha(\mathrm{supp}^T(S))$ is precisely $\mathrm{rk}_\alpha^T(S)$. Since this is a linear program, it is often straight-forward to calculate the T-stable α-rank for a fixed T-support. The G-stable α-rank can then be found by applying Lemma 2.6.

This motivates us to consider the T-stable α-rank on different G-orbits. We will begin by considering $2 \times 2 \times 2$ tensors.

2.1 The G-stable rank for small tensors

For $2 \times 2 \times 2$ tensors over an algebraically closed field, there are seven $G = \mathrm{GL}(V_1) \times \mathrm{GL}(V_2) \times \mathrm{GL}(V_3)$ orbits (e.g. see *Landsberg* (2012)). A representative for each orbit is given below:

1. 0
2. $e_1 \otimes e_1 \otimes e_1$
3. $e_1 \otimes e_1 \otimes e_1 + e_1 \otimes e_2 \otimes e_2$
4. $e_1 \otimes e_1 \otimes e_1 + e_2 \otimes e_1 \otimes e_2$
5. $e_1 \otimes e_1 \otimes e_1 + e_2 \otimes e_2 \otimes e_1$
6. $e_2 \otimes e_1 \otimes e_1 + e_1 \otimes e_2 \otimes e_1 + e_1 \otimes e_1 \otimes e_2$
7. $e_1 \otimes e_1 \otimes e_1 + e_2 \otimes e_2 \otimes e_2$

The T-supports of these representatives, respectively, are

1. 0
2. $\{(1,1,1)\}$
3. $\{(1,1,1),(1,2,2)\}$
4. $\{(1,1,1),(2,1,2)\}$
5. $\{(1,1,1),(2,2,1)\}$
6. $\{(2,1,1),(1,2,1),(1,1,2)\}$
7. $\{(1,1,1),(2,2,2)\}$

For the first representative, the linear program detailed in Definition 2.8 has no constraints, and so it is clear that $\mathrm{rk}_\alpha^T(0) = 0$ for all α. For the other representatives, let us consider how the T-stable α-rank varies over $\alpha = (\alpha_1, \alpha_2, \alpha_3)$ with $\alpha_1 + \alpha_2 + \alpha_3 = 1$.

Example 2.9. Suppose S has T-support $\{(1,1,1)\}$. Our linear program is to minimize

$$\alpha_1 \left(x(1,1) + x(1,2)\right) + \alpha_2 \left(x(2,1) + x(2,2)\right) + \alpha_3 \left(x(3,1) + x(3,2)\right)$$

subject to

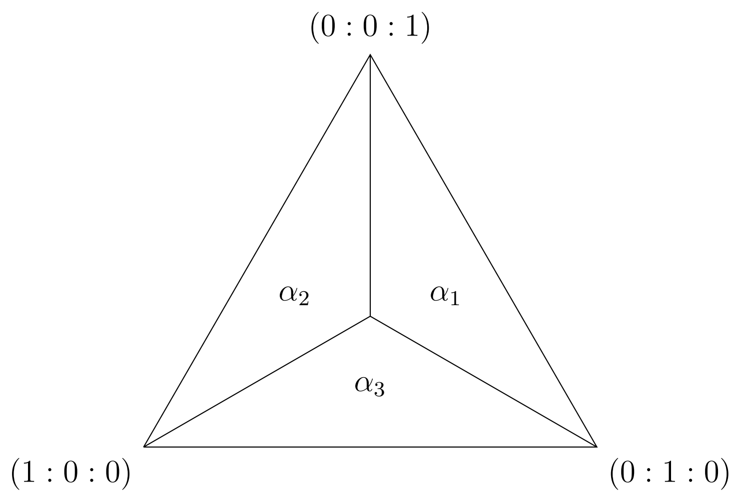
$$x(1,1) + x(2,1) + x(3,1) \geq 1$$

and all the $x(i,j)$ being nonnegative.

Suppose that $\min_j a_j = a_i$. Then this minimum is achieved when $x(i,1) = 1$ and all other variables are 0. In that case, we have $\mathrm{rk}_\alpha^T(S) = \alpha_i$. Hence $\mathrm{rk}_\alpha^T(S) = \min_i\{\alpha_i\}$.

We represent the possible values of $\mathrm{rk}_\alpha^T(S)$ with $\alpha_1 + \alpha_2 + \alpha_3 = 1$ in Figure 2.1 using a bifurcation diagram with 3 vertices, each representing a point where one of the α_i is 1 and the others are 0. Points close to vertex i have a large value of α_i, and points further from it have a small α_i.

Figure 2.1: T-stable rank for T-support $\{(1,1,1)\}$



Example 2.10. Suppose S has T-support $\{(1,1,1),(1,2,2)\}$. Our linear program is to minimize

$$\alpha_1 \left(x(1,1) + x(1,2)\right) + \alpha_2 \left(x(2,1) + x(2,2)\right) + \alpha_3 \left(x(3,1) + x(3,2)\right)$$

subject to

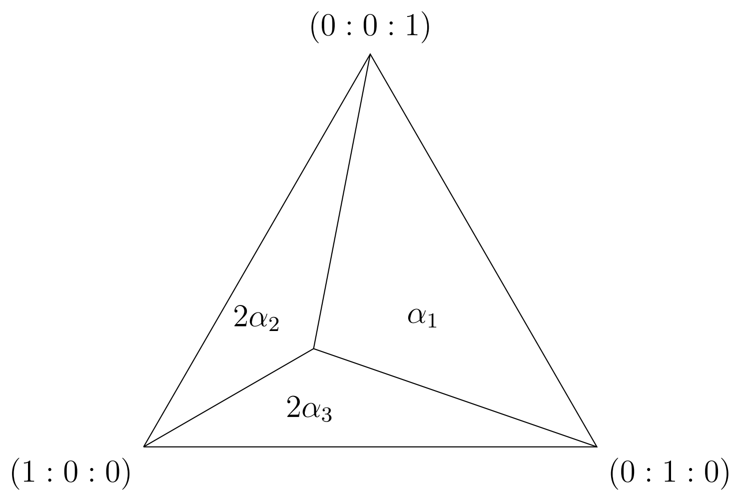
$$x(1,1) + x(2,1) + x(3,1) \geq 1$$

$$x(1,1) + x(2,2) + x(3,2) \geq 1$$

and all the $x(i,j)$ being nonnegative.

Suppose that α_1 is small relative to α_2 and α_3. Then this minimum is achieved when $x(1,1) = 1$ and all other variables are 0. In that case, we have $\mathrm{rk}^T_\alpha(S) = \alpha_1$. Otherwise, if α_2 is small relative to α_1 and α_3 then the minimum is achieved when $x(2,1) = x(2,2) = 1$ and all other variables are 0. In that case, we have $\mathrm{rk}^T_\alpha(S) = 2\alpha_2$. The situation for small α_3 is similar. Hence $\mathrm{rk}^T_\alpha(S) = \min\{\alpha_1, 2\alpha_2, 2\alpha_3\}$. This situation is represented by Figure 2.2.

Figure 2.2: T-stable rank for T-support $\{(1,1,1),(1,2,2)\}$



Example 2.11. The cases for the fourth and fifth T-supports follow from the above example by symmetry. Suppose S has support $\{(1,1,1),(2,1,2)\}$. Then $\mathrm{rk}^T_\alpha(S) = \min\{2\alpha_1, \alpha_2, 2\alpha_3\}$. This situation is represented by Figure 2.3. Suppose S has T-support $\{(1,1,1),(2,2,1)\}$. Then $\mathrm{rk}^T_\alpha(S) = \min\{2\alpha_1, 2\alpha_2, \alpha_3\}$. This situation is represented by Figure 2.4.

Example 2.12. Suppose S has T-support $\{(2,1,1),(1,2,1),(1,1,2)\}$. Our linear program is to minimize

$$\alpha_1\,(x(1,1)+x(1,2)) + \alpha_2\,(x(2,1)+x(2,2)) + \alpha_3\,(x(3,1)+x(3,2))$$

Figure 2.3: T-stable rank for T-support $\{(1,1,1),(2,1,2)\}$

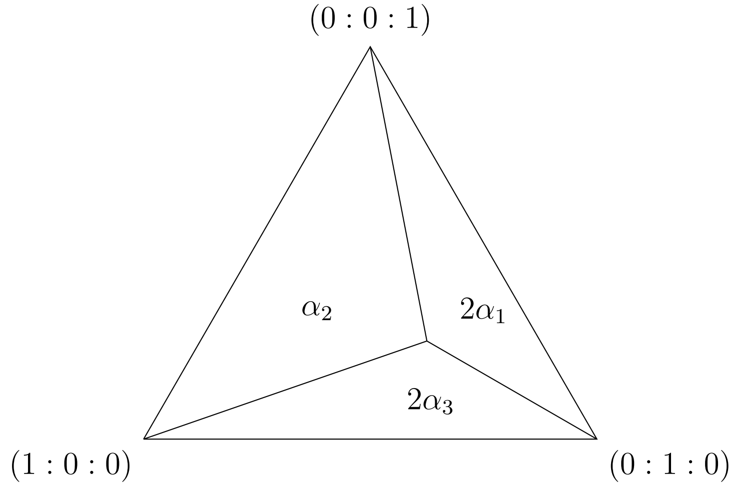
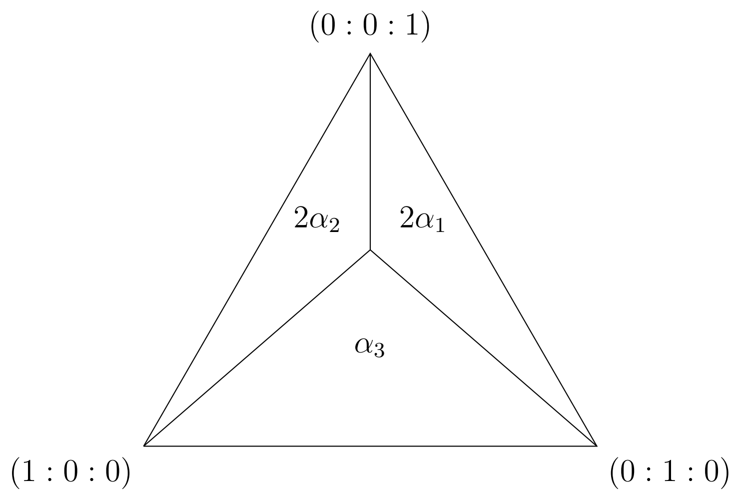


Figure 2.4: T-stable rank for T-support $\{(1,1,1),(2,2,1)\}$



subject to

$$x(1,2) + x(2,1) + x(3,1) \geq 1$$

$$x(1,1) + x(2,2) + x(3,1) \geq 1$$

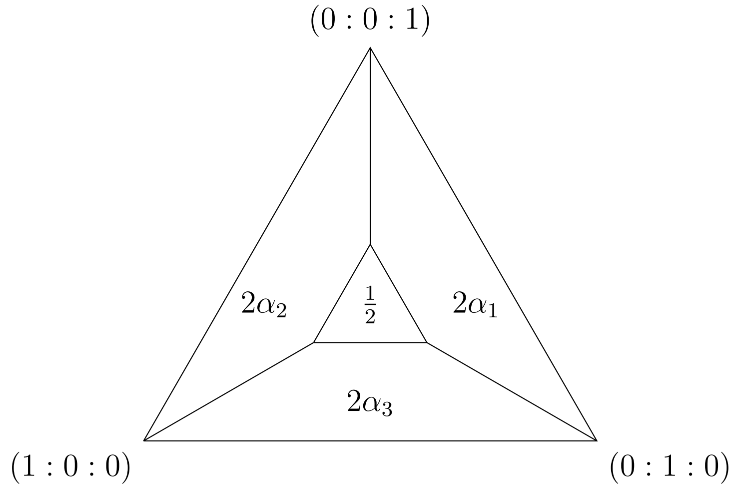
$$x(1,1) + x(2,1) + x(3,2) \geq 1$$

and all the $x(i,j)$ being nonnegative.

Suppose that α_1 is small relative to α_2 and α_3. Then this minimum is achieved when $x(1,1) = x(1,2) = 1$ and all other variables are 0. In that case, we have

$\mathrm{rk}^T_\alpha(S) = 2\alpha_1$. The cases for small α_2 and α_3 are similar. If instead the α_i are close to each other then there is a another possibility - the minimum can be achieved when $x(1,1) = x(2,1) = x(3,1) = \frac{1}{2}$, in which case we have $\mathrm{rk}^T_\alpha(S) = \frac{1}{2}(\alpha_1 + \alpha_2 + \alpha_3) = \frac{1}{2}$. Hence $\mathrm{rk}^T_\alpha(S) = \min\{2\alpha_1, 2\alpha_2, 2\alpha_3, \frac{1}{2}\}$. This situation is represented by Figure 2.5.

Figure 2.5: T-stable rank for T-support $\{(2,1,1),(1,2,1),(1,1,2)\}$



Example 2.13. Suppose S has T-support $\{(1,1,1),(2,2,2)\}$. Our linear program is to minimize

$$\alpha_1\left(x(1,1) + x(1,2)\right) + \alpha_2\left(x(2,1) + x(2,2)\right) + \alpha_3\left(x(3,1) + x(3,2)\right)$$

subject to

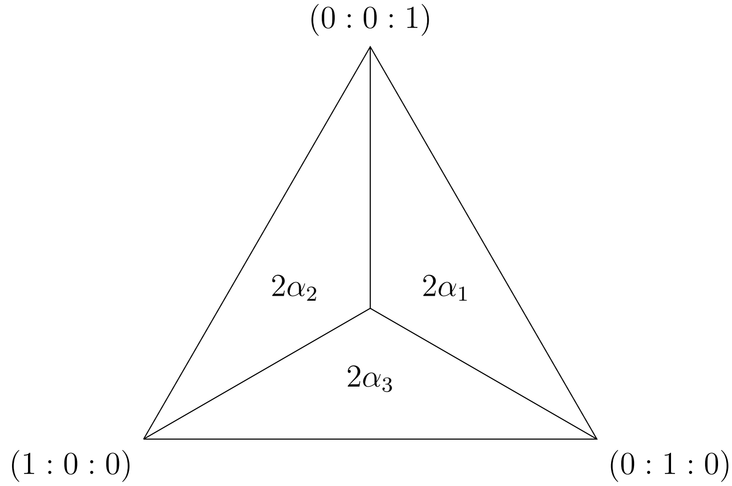
$$x(1,1) + x(2,1) + x(3,1) \geq 1$$

$$x(1,2) + x(2,2) + x(3,2) \geq 1$$

and all the $x(i,j)$ being nonnegative.

Suppose that α_1 is small relative to α_2 and α_3. Then this minimum is achieved when $x(1,1) = x(1,2) = 1$ and all other variables are 0. In that case, we have $\mathrm{rk}^T_\alpha(S) = 2\alpha_1$. The cases for small α_2 and α_3 are similar. Hence $\mathrm{rk}^T_\alpha(S) = \min\{2\alpha_1, 2\alpha_2, 2\alpha_3\}$. This situation is represented by Figure 2.6.

Figure 2.6: T-stable rank for T-support $\{(1,1,1),(2,2,2)\}$



We are interested in proving that these T-stable α ranks are enough to classify the G-stable α ranks of the same tensors. To do this, we will introduce an intermediary concept - instead of working with the maximal torus T, we will work with the maximal Borel subgroup B where B is defined as follows: For each V_i, denote by B_{n_i} the subgroup of GL_{n_i} consisting of invertible $n_i \times n_i$ upper triangular matrices, and let $B = B_{n_1} \times B_{n_2} \times \ldots B_{n_d}$. The B is a maximal Borel subgroup of G, and T is contained in B.

Definition 2.14. For a field K and a tensor $S = (s_{i_1,i_2,\ldots,i_d}) \in K^{n_1} \otimes K^{n_2} \otimes \ldots \otimes K^{n_d}$, we define its *Borel support*, or *B-support* by

$$\operatorname{supp}^B(S) = \{(i_1, i_2, \ldots, i_d) \mid \exists (j_1, j_2, \ldots, j_d) \text{ with } i_k \leq j_k \text{ for all } k \text{ and } s_{j_1,j_2,\ldots,j_d} \neq 0\}.$$

In other words, we have $(i_1, i_2, \ldots, i_d) \in \operatorname{supp}^B(S)$ if and only if there is $(j_1, j_2, \ldots, j_d) \in \operatorname{supp}(S)$ with $i_k \leq j_k$ for all k.

Definition 2.15. Let $S \in K^{n_1} \otimes K^{n_2} \otimes \ldots \otimes K^{n_d}$ be a tensor. Then the linear program $\mathbf{LP'}_\alpha(\operatorname{supp}^B(S))$ is to find the minimum of $\displaystyle\sum_{i=1}^{d} \alpha_i \sum_{j=1}^{n_i} x(i,j)$ where the $x(i,j)$ are nonnegative real variables with the constraints that $\sum_{i=1}^{d} x(i, t_i) \geq 1$ for all $t \in \mathrm{supp}^B(S)$, and whenever $j_1 \leq j_2$, we have $x(i, j_1) \geq x(i, j_2)$.

We define the B-stable α-rank of S to be

$$\mathrm{rk}_\alpha^B(S) := \mathbf{LP'}_\alpha(\mathrm{supp}^B(S)).$$

Remark 2.16. Let S_1, S_2 be tensors such that $\mathrm{supp}^T(S_1) \subseteq \mathrm{supp}^T(S_2)$. Then

$$\mathrm{rk}_\alpha^T(S_1) \leq \mathrm{rk}_\alpha^T(S_2).$$

This is clear when we consider that the linear program for S_2 has more constraints than the one for S_1 but the objective functions are the same. This means the minimum obtained in the linear program for S_1 is smaller than the one for S_2. Similarly, we always have

$$\mathbf{LP}_\alpha(\mathrm{supp}^B(S)) \leq \mathbf{LP'}_\alpha(\mathrm{supp}^B(S)),$$

and hence

$$\mathrm{rk}_\alpha^T(S) \leq \mathrm{rk}_\alpha^B(S)$$

whenever $T \subseteq B$.

Lemma 2.17. Let S be a tensor. Then

$$\inf_{g \in G} \mathrm{rk}_\alpha^T(g \cdot S) = \inf_{g \in G} \mathrm{rk}_\alpha^B(g \cdot S).$$

Proof. From Remark 2.16, we see that

$$\inf_{g \in G} \mathrm{rk}_\alpha^T(g \cdot S) \leq \inf_{g \in G} \mathrm{rk}_\alpha^B(g \cdot S).$$

Additionally, the value of the linear program $\mathbf{LP}_\alpha(\mathrm{supp}^T(S))$ is invariant under permutations of the coordinates in each mode, so we may assume that the variables $x(i,j)$ are weakly decreasing. Hence we may assume the non-zero variables $x(i,j)$ come from a B-support, which solves the linear program $\mathbf{LP'}_\alpha(\mathrm{supp}^B(\sigma \cdot S))$ for some $\sigma \in G$, i.e. $\mathrm{rk}_\alpha^T(S) = \mathrm{rk}_\alpha^B(\sigma \cdot S)$ for some $\sigma \in G$. Acting by $g \in G$ on both sides, and taking the infimum, gives the desired result. □

Combining this with Lemma 2.6, we see that when we are trying to calculate the G-stable α-rank of a tensor S, we may restrict our attention to the B-stable α-ranks of minimal B-supports of S.

Lemma 2.18. Suppose that $S \in K^{n_1} \otimes K^{n_2} \otimes \ldots \otimes K^{n_d}$ is a tensor and that T is a maximal torus such that for every minimal B-support, $\mathrm{supp}^{B_i}(S)$, of S, there exist permutations $\sigma_1, \sigma_2, \ldots, \sigma_d$ with $\sigma_j \in \mathfrak{S}_{n_j}$ for all j such that

$$(\sigma_1, \sigma_2, \ldots, \sigma_d) \cdot \mathrm{supp}^T(S) \subseteq \mathrm{supp}^{B_i}(S).$$

Then we have

$$\mathrm{rk}_\alpha^G(S) = \mathrm{rk}_\alpha^T(S).$$

Proof. Using Lemmas 2.6 and 2.17, we see that for some B_i, we have $\mathrm{rk}_\alpha^G(S) = \mathrm{rk}_\alpha^{B_i}(S)$. Since there exist $\sigma_1, \sigma_2, \ldots, \sigma_d$ such that $(\sigma_1, \sigma_2, \ldots, \sigma_d) \cdot \mathrm{supp}^T(S) \subseteq \mathrm{supp}^{B_i}(S)$, we must have $\mathrm{rk}_\alpha^T(S) \leq \mathrm{rk}_\alpha^{B_i}(S) = \mathrm{rk}_\alpha^G(S)$. But from Lemma 2.6, we must have $\mathrm{rk}_\alpha^G(S) \leq \mathrm{rk}_\alpha^T(S)$, and so we have equality, as claimed. □

For $2 \times 2 \times 2$ tensors, it is easy to see that most of the seven orbits described above have a unique minimal B-support. The first two are obvious:

1. The orbit of 0 has minimal B-support $\emptyset$.

2. The orbit of $e_1 \otimes e_1 \otimes e_1$ has minimal B-support $\{(1,1,1)\}$.

For the next orbit, note that it has slice rank 1, and so any minimal B-support must contain only elements of the form $(1, a, b)$. It must have $(1,1,1)$, an if it had only one more element (either $(1,1,2)$ or $(1,2,1)$), then the resulting tensor would have only rank one. For that reason, it must have both these elements. Permuting the basis elements in the 3rd mode for $e_1 \otimes e_1 \otimes e_1 + e_1 \otimes e_2 \otimes e_2$ gives $e_1 \otimes e_1 \otimes e_2 + e_1 \otimes e_2 \otimes e_1$ which does in fact have B-support $\{(1,1,1),(1,2,1),(1,1,2)\}$. Hence:

3. The orbit of $e_1 \otimes e_1 \otimes e_1 + e_1 \otimes e_2 \otimes e_2$ has minimal B-support $\{(1,1,1),(1,2,1),(1,1,2)\}$.

Similarly,

4. The orbit of $e_1 \otimes e_1 \otimes e_1 + e_2 \otimes e_1 \otimes e_2$ has minimal B-support $\{(1,1,1),(2,1,1),(1,1,2)\}$.

5. The orbit of $e_1 \otimes e_1 \otimes e_1 + e_2 \otimes e_2 \otimes e_1$ has minimal B-support $\{(1,1,1),(1,2,1),(2,1,1)\}$.

For the next orbit, note that the flattenings of $e_2 \otimes e_1 \otimes e_1 + e_1 \otimes e_2 \otimes e_1 + e_1 \otimes e_1 \otimes e_2$ each have rank 2, and so any B-support must contain at least the elements $\{(1,1,1),(2,1,1),(1,2,1),(1,1,2)\}$. This is a B-support for our representative and so we have:

6. The orbit of $e_2 \otimes e_1 \otimes e_1 + e_1 \otimes e_2 \otimes e_1 + e_1 \otimes e_1 \otimes e_2$ has minimal B-support $\{(1,1,1),(2,1,1),(1,2,1),(1,1,2)\}$.

The only orbit with more than one minimal B-support is that of $e_1 \otimes e_1 \otimes e_1 + e_2 \otimes e_2 \otimes e_2$ which has three minimal B-supports - each of

- $\{(1,1,1),(2,1,1)(1,2,1),(1,1,2),(1,2,2)\}$,
- $\{(1,1,1),(2,1,1)(1,2,1),(1,1,2),(2,1,2)\}$, and

- $\{(1,1,1),(2,1,1)(1,2,1),(1,1,2),(2,2,1)\}$.

are minimal B-supports. To see this, note that the closure of this orbit contains the closure of orbit 6, and so any minimal B-support of orbit 7 must contain the minimal B-support, $\{(1,1,1),(2,1,1),(1,2,1),(1,1,2)\}$, of orbit 6. Conversely, any tensor with the B-support $\{(1,1,1),(2,1,1),(1,2,1),(1,1,2)\}$ contains $e_2 \otimes e_1 \otimes e_1 + e_1 \otimes e_2 \otimes e_1 + e_1 \otimes e_1 \otimes e_2$ in its G-orbit, and so $\{(1,1,1),(2,1,1),(1,2,1),(1,1,2)\}$ must be strictly contained in any B-support of $e_1 \otimes e_1 \otimes e_1 + e_2 \otimes e_2 \otimes e_2$. It is easy to verify that permuting the basis elements in each mode can result in the three B-supports above.

As there exist permutations of basis vectors in each mode though so that our representative lies in each B-support, Lemma 2.18 tells us that the T-support of this representative is sufficient to calculate the G-stable rank.

This allows us to conclude our analysis of G-stable α-ranks of $2 \times 2 \times 2$ tensors over $\mathbb{C}$:

Proposition 2.19. The G-stable α-ranks of $2 \times 2 \times 2$ tensors over $\mathbb{C}$ are classified by orbit as follows:

1. $\mathrm{rk}_\alpha^G(0) = 0$.
2. $\mathrm{rk}_\alpha^G(e_1 \otimes e_1 \otimes e_1) = \min_i\{\alpha_i\}$.
3. $\mathrm{rk}_\alpha^G(e_1 \otimes e_1 \otimes e_1 + e_1 \otimes e_2 \otimes e_2) = \min\{\alpha_1, 2\alpha_2, 2\alpha_3\}$.
4. $\mathrm{rk}_\alpha^G(e_1 \otimes e_1 \otimes e_1 + e_2 \otimes e_1 \otimes e_2) = \min\{2\alpha_1, \alpha_2, 2\alpha_3\}$.
5. $\mathrm{rk}_\alpha^G(e_1 \otimes e_1 \otimes e_1 + e_2 \otimes e_2 \otimes e_1) = \min\{2\alpha_1, 2\alpha_2, \alpha_3\}$.
6. $\mathrm{rk}_\alpha^G(e_2 \otimes e_1 \otimes e_1 + e_1 \otimes e_2 \otimes e_1 + e_1 \otimes e_1 \otimes e_2) = \min\{2\alpha_1, 2\alpha_2, 2\alpha_3, \frac{1}{2}(\alpha_1 + \alpha_2 + \alpha_3)\}$.
7. $\mathrm{rk}_\alpha^G(e_1 \otimes e_1 \otimes e_1 + e_2 \otimes e_2 \otimes e_2) = \min\{2\alpha_1, 2\alpha_2, 2\alpha_3\}$.

Proof. This is a result of applying Lemma 2.18 to Examples 2.9-2.13, together with the determination of the minimal B-supports given above. $\square$

For $2 \times 2 \times 3$ tensors over an algebraically closed field, there are a total of 9 $G = \mathrm{GL}(V_1) \times \mathrm{GL}(V_2) \times \mathrm{GL}(V_3)$ orbits - each of the orbits of $2 \times 2 \times 3$ tensors embeds into one in the space of $2 \times 2 \times 3$ tensors, and there are two additional orbits, with representatives given by:

8. $e_2 \otimes e_1 \otimes e_1 + e_1 \otimes e_2 \otimes e_2 + e_1 \otimes e_1 \otimes e_3$

9. $e_1 \otimes e_2 \otimes e_2 + e_2 \otimes e_1 \otimes e_2 + e_2 \otimes e_2 \otimes e_1 + e_1 \otimes e_1 \otimes e_3$

Additionally, there are 3 more minimal B-supports for $2 \times 2 \times 3$ tensors. The tensor $e_2 \otimes e_1 \otimes e_1 + e_1 \otimes e_2 \otimes e_2 + e_1 \otimes e_1 \otimes e_3$ has B-support

$$\{(1,1,1),(2,1,1)(1,2,1),(1,1,2),(1,2,2),(1,1,3)\},$$

and by permuting the first two basis elements in the third mode, one obtains the tensor $e_2 \otimes e_1 \otimes e_2 + e_1 \otimes e_2 \otimes e_1 + e_1 \otimes e_1 \otimes e_3$ which has B-support

$$\{(1,1,1),(2,1,1)(1,2,1),(1,1,2),(2,1,2),(1,1,3)\}.$$

To see that these are the only minimal B-supports for this orbit, note first that this orbit must contain some element of the form $(a, b, 3)$, and that any B-support which contains an element of this form must contain $(1, 1, 3)$. It is also easy to see that orbit 7 is contained in the closure of orbit 8 (by replacing e_3 with te_3 and taking the limit as $t \to 0$), and so any minimal B-support for orbit 8 should contain one of the B-supports for orbit 7. Appending $(1, 1, 3)$ to the first two B-supports of orbit 7 gives the claimed B-supports above. The only other possibility for orbit 8 would be the B-support

$$\{(1,1,1),(2,1,1)(1,2,1),(1,1,2),(2,2,1),(1,1,3)\},$$

but this is not minimal - one could make a change of basis in the third mode to result in the B-support of

$$\{(1,1,1),(2,1,1)(1,2,1),(1,1,2),(2,2,1)\}.$$

The other minimal B-support is

$$\{(1,1,1),(2,1,1)(1,2,1),(1,1,2),(1,2,2),(2,1,2),(2,2,1),(1,1,3)\},$$

which is precisely the B-support for $e_1 \otimes e_2 \otimes e_2 + e_2 \otimes e_1 \otimes e_2 + e_2 \otimes e_2 \otimes e_1 + e_1 \otimes e_1 \otimes e_3$.

To see that this is minimal for this orbit, first note that the closure of orbit 8 is contained in that of 9, and so the minimal support must contain at least one of

$$\{(1,1,1),(2,1,1)(1,2,1),(1,1,2),(1,2,2),(1,1,3)\}$$

or

$$\{(1,1,1),(2,1,1)(1,2,1),(1,1,2),(2,1,2),(1,1,3)\}.$$

But it is easy to verify that any tensor whose B-support is strictly between one of these and

$$\{(1,1,1),(2,1,1)(1,2,1),(1,1,2),(1,2,2),(2,1,2),(2,2,1),(1,1,3)\}$$

is contained in orbit 8. Thus, this B-support is the minimal one for orbit 9. Armed with the knowledge that these representatives have minimal B-supports for their orbits, we may determine the possible G-stable α-ranks for $2 \times 2 \times 3$ tensors.

Proposition 2.20. The G-stable α-ranks of $2 \times 2 \times 3$ tensors over $\mathbb{C}$ are classified by orbit as follows:

1. $\mathrm{rk}_\alpha^G(0) = 0$.

2. $\mathrm{rk}_\alpha^G(e_1 \otimes e_1 \otimes e_1) = \min_i\{\alpha_i\}$.

3. $\mathrm{rk}_\alpha^G(e_1 \otimes e_1 \otimes e_1 + e_1 \otimes e_2 \otimes e_2) = \min\{\alpha_1, 2\alpha_2, 2\alpha_3\}$.

4. $\mathrm{rk}_\alpha^G(e_1 \otimes e_1 \otimes e_1 + e_2 \otimes e_1 \otimes e_2) = \min\{2\alpha_1, \alpha_2, 2\alpha_3\}$.

5. $\mathrm{rk}_\alpha^G(e_1 \otimes e_1 \otimes e_1 + e_2 \otimes e_2 \otimes e_1) = \min\{2\alpha_1, 2\alpha_2, \alpha_3\}$.

6. $\mathrm{rk}_\alpha^G(e_2 \otimes e_1 \otimes e_1 + e_1 \otimes e_2 \otimes e_1 + e_1 \otimes e_1 \otimes e_2) = \min\{2\alpha_1, 2\alpha_2, 2\alpha_3, \frac{1}{2}(\alpha_1 + \alpha_2 + \alpha_3)\}$.

7. $\mathrm{rk}_\alpha^G(e_1 \otimes e_1 \otimes e_1 + e_2 \otimes e_2 \otimes e_2) = \min\{2\alpha_1, 2\alpha_2, 2\alpha_3\}$.

8. $\mathrm{rk}_\alpha^G(e_2 \otimes e_1 \otimes e_1 + e_1 \otimes e_2 \otimes e_2 + e_1 \otimes e_1 \otimes e_3) = \min\{2\alpha_1, 2\alpha_2, 3\alpha_3, \frac{1}{2}\alpha_1 + \frac{1}{2}\alpha_2 + \alpha_3\}$.

9. $\mathrm{rk}_\alpha^G(e_1 \otimes e_2 \otimes e_2 + e_2 \otimes e_1 \otimes e_2 + e_2 \otimes e_2 \otimes e_1 + e_1 \otimes e_1 \otimes e_3) = \min\{2\alpha_1, 2\alpha_2, 3\alpha_3\}$.

Proof. The first seven orbits behave in the same way as for $2 \times 2 \times 2$ tensors, and so the results are the same as in Proposition 2.19.

As shown above, the representatives we have chosen for the other two orbits have minimal B-supports, up to permutation of basis elements. Using Lemma 2.18, we can see that these representatives we have chosen are thus sufficient to determine the G-stable ranks on each orbit - we need only find the T-stable α-ranks for the representatives of orbits 8 and 9.

For orbit 8, suppose S has T-support $\{(2,1,1), (1,2,2), (1,1,3)\}$. Our linear program is to minimize

$$\alpha_1 (x(1,1) + x(1,2)) + \alpha_2 (x(2,1) + x(2,2)) + \alpha_3 (x(3,1) + x(3,2))$$

subject to

$$x(1,2) + x(2,1) + x(3,1) \geq 1$$

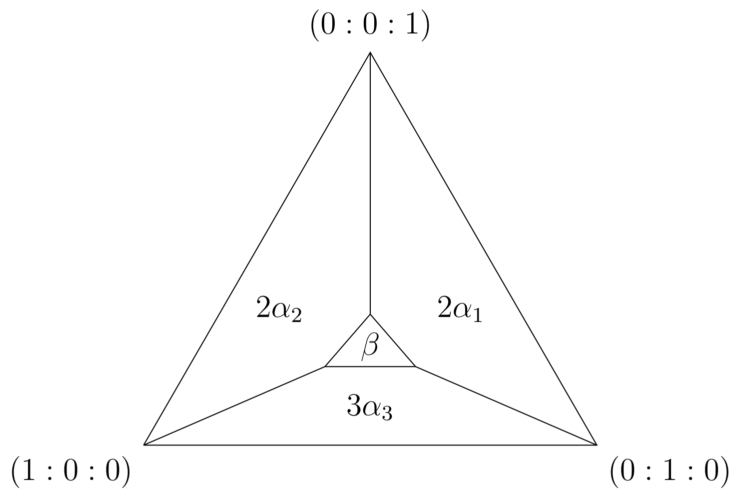
$$x(1,1) + x(2,2) + x(3,2) \geq 1$$

$$x(1,1) + x(2,1) + x(3,3) \geq 1$$

and all the $x(i,j)$ being nonnegative.

Suppose that α_1 is small relative to α_2 and α_3. Then this minimum is achieved when $x(1,1) = x(1,2) = 1$ and all other variables are 0. In that case, we have $\mathrm{rk}^T_\alpha(S) = 2\alpha_1$. The case for small α_2 is similar. When α_3 is small, the minimum is achieved when $x(3,1) = x(3,2) = x(3,3) = 1$, and in that case we have $\mathrm{rk}^T_\alpha(S) = 3\alpha_3$. If instead the α_i are close to each other then there is a another possibility - the minimum can be achieved when $x(1,1) = x(2,1) = x(3,1) = x(3,2) = \frac{1}{2}$, in which case we have $\mathrm{rk}^T_\alpha(S) = \frac{1}{2}\alpha_1 + \frac{1}{2}\alpha_2 + \alpha_3$. Hence $\mathrm{rk}^T_\alpha(S) = \min\{2\alpha_1, 2\alpha_2, 3\alpha_3, \frac{1}{2}\alpha_1 + \frac{1}{2}\alpha_2 + \alpha_3\}$. This situation is represented by Figure 2.7, where $\beta = \frac{1}{2}\alpha_1 + \frac{1}{2}\alpha_2 + \alpha_3$.

Figure 2.7: T-stable rank for T-support $\{(2,1,1),(1,2,2),(1,1,3)\}$



For orbit 9, suppose S has T-support $\{(2,1,1),(1,2,1),(1,1,2),(1,1,3)\}$. Our linear program is to minimize

$$\alpha_1\left(x(1,1)+x(1,2)\right)+\alpha_2\left(x(2,1)+x(2,2)\right)+\alpha_3\left(x(3,1)+x(3,2)+x(3,3)\right)$$

subject to

$$x(1,2)+x(2,1)+x(3,1) \geq 1$$

$$x(1,1)+x(2,2)+x(3,1) \geq 1$$

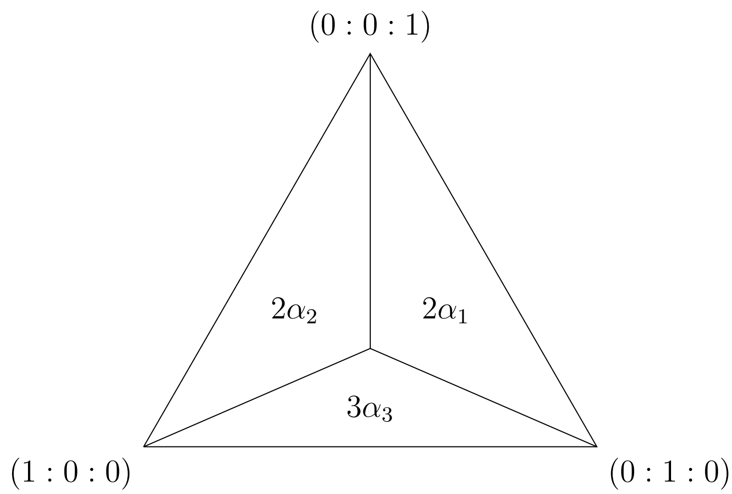
$$x(1,1) + x(2,1) + x(3,2) \geq 1$$

$$x(1,1) + x(2,1) + x(3,3) \geq 1$$

and all the $x(i,j)$ being nonnegative.

Suppose that α_1 is small relative to α_2 and α_3. Then this minimum is achieved when $x(1,1) = x(1,2) = 1$ and all other variables are 0. In that case, we have $\mathrm{rk}_\alpha^T(S) = 2\alpha_1$. The case for small α_2 is similar. When α_3 is small, the minimum is achieved when $x(3,1) = x(3,2) = x(3,3) = 1$, and in that case we have $\mathrm{rk}_\alpha^T(S) = 3\alpha_3$. There are no other possibilities - when $(\alpha_1, \alpha_2, \alpha_3) = (\frac{3}{8}, \frac{3}{8}, \frac{1}{4})$, we have $2\alpha_1 = 2\alpha_2 = 3\alpha_3$. Hence $\mathrm{rk}_\alpha^T(S) = \min\{2\alpha_1, 2\alpha_2, 3\alpha_3\}$. This situation is represented by Figure 2.8.

Figure 2.8: T-stable rank for T-support $\{(2,1,1), (1,2,1), (1,1,2), (1,1,3)\}$



□

Remark 2.21. From this analysis, we can see that for any pair S_1, S_2, of $2 \times 2 \times 2$ or $2 \times 2 \times 3$ tensors which do not belong to the same G-orbit, there exists a weight α such that $\mathrm{rk}_\alpha^G(S_1) \neq \mathrm{rk}_\alpha^G(S_2)$. One might hope that this would be true for tensors S_1, S_2, of arbitrary dimension. However, for sufficiently large dimensions, there are infinitely many G-orbits, but still only finitely many possible T- (and B-) supports, which implies that some orbits must share the same linear programs, and hence share the same G-stable α ranks for every value of α.

2.2 X-Rank

Recall that Derksen (*Derksen* (2020)) introduced the notion of the G-stable α-rank, $\mathrm{rk}_\alpha^G(S)$, of a tensor S, as seen in Definition 2.4. The definition, over $\mathbb{C}$, is restated below:

Definition 2.22. For $S \in V_1 \otimes \ldots \otimes V_d$, let $\Phi_i(S) : (V_1 \otimes \ldots \otimes \hat{V}_i \otimes \ldots V_d)^\star \to V_i$ be the i-th flattening. For a weight $\alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{R}^d_{>0}$, the *$G$-stable α-rank of S* is

$$\mathrm{rk}_\alpha^G(S) := \sup_{g \in G} \min_i \frac{\alpha_i \|g \cdot S\|^2}{\|\Phi_i(g \cdot S)\|_\sigma^2}.$$

In the above paper, the author commonly uses the weight $\alpha = (1, \ldots, 1)$. While this is useful in many applications, it is not necessarily the case that weighting each mode equally will be the best in every circumstance. Particularly when the dimensions of the spaces V_i are not all equal, we are motivated to find weights in each mode which are in some way optimal. This leads us to introduce a new notion of rank, which we will call the *X-rank*.

Definition 2.23. The G-stable X-rank of a nonzero tensor S is given by

$$\mathrm{Xrk}^G(S) = \max_\alpha \mathrm{rk}_\alpha^G(S)$$

where the maximum is taken over all $\alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{R}^d_{\geq 0}$ such that $\sum_{i=1}^d \alpha_i = d$. By convention, we say that $\mathrm{Xrk}^G(0) = 0$.

Unless otherwise stated, we will assume $G = \mathrm{GL}(V_1) \times \mathrm{GL}(V_2) \times \cdots \times \mathrm{GL}(V_d)$.

Proposition 2.24. For complex $2 \times 2 \times 3$ tensors (and $2 \times 2 \times 2$ tensors where appropriate), we have:

1. $\mathrm{Xrk}^G(0) = 0$.

2. $\mathrm{Xrk}^G(e_1 \otimes e_1 \otimes e_1) = 1$.

3. $\mathrm{Xrk}^G(e_1 \otimes e_1 \otimes e_1 + e_1 \otimes e_2 \otimes e_2) = \frac{3}{2}$.

4. $\mathrm{Xrk}^G(e_1 \otimes e_1 \otimes e_1 + e_2 \otimes e_1 \otimes e_2) = \frac{3}{2}$.

5. $\mathrm{Xrk}^G(e_1 \otimes e_1 \otimes e_1 + e_2 \otimes e_2 \otimes e_1) = \frac{3}{2}$.

6. $\mathrm{Xrk}^G(e_2 \otimes e_1 \otimes e_1 + e_1 \otimes e_2 \otimes e_1 + e_1 \otimes e_1 \otimes e_2) = \frac{3}{2}$.

7. $\mathrm{Xrk}^G(e_1 \otimes e_1 \otimes e_1 + e_2 \otimes e_2 \otimes e_2) = 2$.

8. $\mathrm{Xrk}^G(e_2 \otimes e_1 \otimes e_1 + e_2 \otimes e_2 \otimes e_1 + e_1 \otimes e_1 \otimes e_3) = 2$.

9. $\mathrm{Xrk}^G(e_1 \otimes e_2 \otimes e_2 + e_2 \otimes e_1 \otimes e_2 + e_2 \otimes e_2 \otimes e_1 + e_1 \otimes e_1 \otimes e_3) = 3$.

Proof. We use the results from Proposition 2.20. For most of the orbits, we may set all of the minima equal, and solve for when $\alpha_1 + \alpha_2 + \alpha_3 = 3$. The only orbits for which this is not appropriate are orbits 6 and 8.

For orbit 6, recall Figure 2.5. The maximum is obtained when α_1, α_2, and α_3 are relatively close, and in that case, the maximum is $\frac{1}{2}(\alpha_1 + \alpha_2 + \alpha_3) = \frac{3}{2}$.

For orbit 8, recall Figure 2.7. The maximum is obtained when $\beta = \frac{1}{2}\alpha_1 + \frac{1}{2}\alpha_2 + \alpha_3$ is as large as possible, while still being less than or equal to each of $2\alpha_1$, $2\alpha_2$, and $3\alpha_3$. This occurs when $\alpha_1 = \alpha_2 = \alpha_3 = 1$, which results in a maximum of 2. $\square$

The X-rank has many properties that we would desire a rank function to have.

Proposition 2.25. The X-rank satisfies the following properties:

1. (Zero tensor) $\mathrm{Xrk}^G(S) = 0$ if and only if S is the zero tensor.

2. (Scale Invariance) For a nonzero scalar λ and a tensor S, $\mathrm{Xrk}^G(\lambda S) = \mathrm{Xrk}^G(S)$.

3. (Basis Invariance) The X-rank is constant on G-orbits.

4. (Triangle Inequality) For all $S, T \in \mathbf{V}$, $\mathrm{Xrk}^G(S+T) \leq \mathrm{Xrk}^G(S) + \mathrm{Xrk}^G(T)$.

5. (Simple tensors) If v is a simple tensor, then $\mathrm{Xrk}^G(v)$=1.

Proof. The first three properties are clear from the definition, and Property 4 is a consequence of the triangle inequality for the G-stable rank.

For Property 5, recall that for a simple tensor v, we have for all i, $\|v\| = \|\Phi_i(v)\| = \|\Phi_i(v)\|_\sigma$, and also that $g \cdot v$ is simple for all g. Hence $\mathrm{Xrk}^G(v) = \max_\alpha \min_i \alpha_i$. This maximum is clearly achieved when $\alpha_1 = \alpha_2 = \cdots = \alpha_d = 1$, and results in $\mathrm{Xrk}^G(v) = 1$. □

An alternate description of the X-rank may be useful for computations:

Lemma 2.26. $S \in V_1 \otimes \ldots \otimes V_d$, we have:

$$\mathrm{Xrk}^G(S) = \sup_{g \in G} \frac{d\|g \cdot S\|^2}{\sum\limits_i \|\Phi_i(g \cdot S)\|_\sigma^2}.$$

Proof. We have

$$\mathrm{Xrk}^G(S) := \sup_{g \in G} \max_\alpha \min_i \frac{\alpha_i \|g \cdot S\|^2}{\|\Phi_i(g \cdot S)\|_\sigma^2}.$$

For a fixed $g \in G$,

$$\max_\alpha \min_i \frac{\alpha_i \|g \cdot S\|^2}{\|\Phi_i(g \cdot S)\|_\sigma^2}$$

is obtained when for all i, j, we have

$$\frac{\alpha_i \|g \cdot S\|^2}{\|\Phi_i(g \cdot S)\|_\sigma^2} = \frac{\alpha_j \|g \cdot S\|^2}{\|\Phi_j(g \cdot S)\|_\sigma^2}.$$

Therefore, for all i, we have

$$\alpha_i = \frac{\alpha_1 \|\Phi_i(g \cdot S)\|_\sigma^2}{\|\Phi_1(g \cdot S)\|_\sigma^2}.$$

Since $\sum \alpha_i = d$, we have

$$\alpha_1 = \frac{d\|\Phi_1(g \cdot S)\|_\sigma^2}{\sum\limits_i \|\Phi_i(g \cdot S)\|_\sigma^2}.$$

Thus we have

$$\mathrm{Xrk}^G(S) := \sup_{g \in G} \frac{\alpha_1 \|g \cdot S\|^2}{\|\Phi_1(g \cdot S)\|_\sigma^2} = \sup_{g \in G} \frac{d\|g \cdot S\|^2}{\sum\limits_i \|\Phi_i(g \cdot S)\|_\sigma^2}.$$

□

Proposition 2.27. For the tensor $S = \sum\limits_{i=1}^{r} e_i \otimes e_i \otimes \ldots \otimes e_i \in V_1 \otimes V_2 \otimes \ldots \otimes V_d$, we have $\mathrm{Xrk}^G(S) = r$.

Proof. We have $\|S\|^2 = r$ and for all i, $\|\Phi_i(S)\|_\sigma = 1$. Using Lemma 2.26, we see that $\mathrm{Xrk}^G(S) \geq \frac{dr}{d} = r$.

By Proposition 2.25, we know that $\mathrm{Xrk}^G(S) \leq \sum_{i=1}^{r} \mathrm{Xrk}^G(e_i \otimes e_i \otimes \ldots \otimes e_i) = r$.

Therefore, $\mathrm{Xrk}^G(S) = r$. □

CHAPTER III

Nuclear Norm Under Tensor Kronecker Products

3.1 Introduction

Much work in algebraic complexity theory concerns the *vertical tensor product*, alternatively known as the *tensor Kronecker product*. It is natural to ask which properties of tensors are preserved under this product. It has been known since the work of Strassen that properties such as the tensor rank and border rank are not preserved when taking the vertical tensor product (or even the usual tensor product) of tensors of order at least 3 (*Christandl et al.* (2018), *Christandl et al.* (2019)). Meanwhile, it is clear that the Frobenius norm of tensors is preserved, and it has been shown that the spectral norm of tensors is also multiplicative under the tensor Kronecker product (*Derksen* (2016)), over both $\mathbb{R}$ and $\mathbb{C}$.

The notion of *cross norms* - norms on tensor product spaces which are multiplicative with respect to the usual tensor product - was originally introduced by Schatten *Schatten* (1950) in the study of Banach spaces, and has seen much interest in the decades since (e.g., *Defant and Floret* (1993), *Diestel et al.* (2008)). If such a norm has a dual norm which is also a cross norm, then it is said to be *reasonable*. We may say analogously that a norm on a tensor product space is a *Kronecker-cross norm* if it is multiplicative with respect to the tensor Kronecker product, and that such a norm is a *reasonable Kronecker-cross norm* if its dual is also Kronecker-cross. We will show that the nuclear norm is Kronecker-cross, and hence that it and the spectral norm of tensors are reasonable Kronecker-cross norms.

This result can be generalized in a number of ways. We present two such generalizations. The first shows that a nuclear norm on tuples of tensors is also Kronecker-cross, and the second extends the multiplicativity of the nuclear norm to a result about the multiplicativity of injective norms on tensor products more generally.

3.2 Notation

In addition to our notation from Chapter I, we will require some more definitions relating to tensor norms and tensor Kronecker products.

We say two norms $\|\cdot\|_X$ and $\|\cdot\|_Y$ on a tensor product space $\mathbf{U}$ are *dual* if for all $S, S' \in \mathbf{U}$, we have $|\langle S, S'\rangle| \leq \|S\|_X \|S'\|_Y$, and for every $S \in \mathbf{U}$, there exists some nonzero $S' \in \mathbf{U}$ such that the above inequality becomes an equality. (It is easy to see that the last condition is equivalent to: for every $S' \in \mathbf{U}$ there exists a nonzero $S \in \mathbf{U}$ for which the inequality becomes an equality.)

Following the work of Schatten (*Schatten* (1950)), we will say that a norm $\|\cdot\|_X$ on a tensor product space $\mathbf{U} \otimes \mathbf{V}$ is a *cross norm* if for all $S \in \mathbf{U}$ and $T \in \mathbf{V}$, we have $\|S \otimes T\|_X = \|S\|_X \|T\|_X$. Similarly, we will say that a norm $\|\cdot\|_X$ on a tensor product space $\mathbf{U} \boxtimes \mathbf{V}$ is a *Kronecker cross norm* if for all $S \in \mathbf{U}$ and $T \in \mathbf{V}$, we have $\|S \boxtimes T\|_X = \|S\|_X \|T\|_X$.

3.3 Norms under Tensor Kronecker Products

In this section, $\mathbf{U} = (U, (U^{(1)}, \ldots, U^{(d)}))$ and $\mathbf{V} = (V, (V^{(1)}, \ldots, V^{(d)}))$ will denote d-th order tensor product spaces where for each i and j, $U^{(i)} = K^{n_i}$ and $V^{(j)} = K^{n_j}$ where $K = \mathbb{R}$ or $K = \mathbb{C}$ and the $n_i, n_j \in \mathbb{N}$ are the dimensions of each space.

Many of the norms we will be interested in can be defined in similar ways on different tensor spaces, and so in an abuse of notation, we will often use $\|\cdot\|_X$ to simultaneously denote the norms $\|\cdot\|_{X,\mathbf{U}}$, $\|\cdot\|_{X,\mathbf{V}}$, and $\|\cdot\|_{X,\mathbf{U}\boxtimes\mathbf{V}}$ defined on the spaces $\mathbf{U}$, $\mathbf{V}$, and $\mathbf{U}\boxtimes\mathbf{V}$ respectively.

The following is clear by direct calculation:

Proposition 3.1. The Frobenius norm is a cross norm. That is to say, if S and T are tensors then $\|S\otimes T\| = \|S\|\|T\|$. Moreover, it is a Kronecker cross norm i.e. if $S\in\mathbf{U}$ and $T\in\mathbf{V}$ are d-th order tensors, then $\|S\boxtimes T\| = \|S\|\|T\|$.

We will also make use of the following propositions. The first is well known (e.g. see *Lim and Comon* (2014)).

Proposition 3.2. On any tensor product space, $\|\cdot\|_\star$ and $\|\cdot\|_\sigma$ are dual.

The second appears as Proposition 3.3 in *Derksen* (2016). Though the proof there is performed over $\mathbb{C}$, it also works over $\mathbb{R}$.

Proposition 3.3. If $S\in\mathbf{U}^r$ and $T\in\mathbf{V}^s$, then we have

$$[S\boxtimes T]_\alpha = [S]_\alpha[T]_\alpha.$$

In particular, taking $r = s = 1$, we see that for $S\in\mathbf{U}$, and $T\in\mathbf{V}$, we have $\|S\boxtimes T\|_\sigma = \|S\|_\sigma\|T\|_\sigma$, i.e. the spectral norm is a Kronecker cross norm.

We aim to use the duality of the spectral and nuclear norms, together with the above result on the spectral norm, to make a statement about the nuclear norm of the vertical tensor product of tensors. We require the following Lemma:

Lemma 3.4. Let $\|\cdot\|_X$ and $\|\cdot\|_Y$ be dual on tensor product spaces $\mathbf{U}$, $\mathbf{V}$, and $\mathbf{U}\boxtimes\mathbf{V}$, and suppose that for all $S\in\mathbf{U}$ and $T\in\mathbf{V}$, we have $\|S\boxtimes T\|_Y \le \|S\|_Y\|T\|_Y$. Then for all $S\in\mathbf{U}$, $T\in\mathbf{V}$, we have $\|S\boxtimes T\|_X \ge \|S\|_X\|T\|_X$.

Proof. Pick nonzero $S' \in \mathbf{U}$ and $T' \in \mathbf{V}$ such that $|\langle S, S'\rangle| = \|S\|_X \|S'\|_Y$ and $|\langle T, T'\rangle| = \|T\|_X \|T'\|_Y$. Then,

$$\begin{aligned}
\|S \boxtimes T\|_X \|S'\|_Y \|T'\|_Y &\geq \|S \boxtimes T\|_X \|S' \boxtimes T'\|_Y \\
&\geq |\langle S \boxtimes T, S' \boxtimes T'\rangle| \\
&= |\langle S, S'\rangle| |\langle T, T'\rangle| \\
&= \|S\|_X \|T\|_X \|S'\|_Y \|T'\|_Y.
\end{aligned}$$

Since S' and T' are nonzero, we conclude that $\|S \boxtimes T\|_X \geq \|S\|_X \|T\|_X$. □

Remark 3.5. Note that $\|\cdot\|_X$ and $\|\cdot\|_Y$ being dual and $\|\cdot\|_Y$ being a Kronecker cross norm is not sufficient to conclude that $\|\cdot\|_X$ is a Kronecker cross norm. A (Kronecker) cross norm whose dual is also a (Kronecker) cross norm is sometimes said to be *reasonable* (*Schatten* (1950)).

Applying the above lemma to Proposition 3.3 allows us to deduce the following:

Proposition 3.6. The nuclear norm is a Kronecker cross norm, i.e. if $S \in \mathbf{U}$, and $T \in \mathbf{V}$ are d-th order tensors, then $\|S \boxtimes T\|_\star = \|S\|_\star \|T\|_\star$.

Proof. Combining Propositions 3.2 and 3.3 with Lemma 3.4, we see that $\|S \boxtimes T\|_\star \geq \|S\|_\star \|T\|_\star$. So it remains to show that $\|S \boxtimes T\|_\star \leq \|S\|_\star \|T\|_\star$.

Let

$$S = \sum_{i=1}^{r_S} u_i \text{ with } u_i = u_i^{(1)} \otimes \ldots \otimes u_i^{(d)} \text{ and } u_i^{(e)} \in U^{(e)},$$

and similarly,

$$T = \sum_{j=1}^{r_T} v_j \text{ with } v_j = v_j^{(1)} \otimes \ldots \otimes v_j^{(d)} \text{ and } v_j^{(e)} \in V^{(e)}.$$

Then

$$S \boxtimes T = \sum_{i=1}^{r_S} \sum_{j=1}^{r_T} (u_i^{(1)} \otimes v_j^{(1)}) \otimes \ldots \otimes (u_i^{(d)} \otimes v_j^{(d)})$$

and so, applying Proposition 3.1,

$$\begin{aligned}
\|S \boxtimes T\|_\star &\leq \sum_{i=1}^{r_S} \sum_{j=1}^{r_T} \left\| (u_i^{(1)} \otimes v_j^{(1)}) \otimes \ldots \otimes (u_i^{(d)} \otimes v_j^{(d)}) \right\| \\
&= \sum_{i=1}^{r_S} \sum_{j=1}^{r_T} \left\| u_i^{(1)} \otimes \ldots \otimes u_i^{(d)} \right\| \left\| v_j^{(1)} \otimes \ldots \otimes v_j^{(d)} \right\| \\
&= \sum_{i=1}^{r_S} \left\| u_i^{(1)} \otimes \ldots \otimes u_i^{(d)} \right\| \sum_{j=1}^{r_T} \left\| v_j^{(1)} \otimes \ldots \otimes v_j^{(d)} \right\|.
\end{aligned}$$

Taking the minima of $\sum\limits_{i=1}^{r_S} \|u_i\|$ and $\sum\limits_{j=1}^{r_S} \|v_j\|$ over all decompositions $S = \sum\limits_{i=1}^{r_S} u_i$ and $T = \sum\limits_{j=1}^{r_T} v_j$, we see that $\|S \boxtimes T\|_\star \leq \|S\|_\star \|T\|_\star$. □

In fact, we may further generalize the above result by considering tuples of tensors.

Definition 3.7. For $\mathbf{S} \in \mathbf{U}^r$ an r-tuple of tensors with $\mathbf{S} = (S_1, \ldots, S_r)$ and $1 \leq \beta < \infty$ we define $[\mathbf{S}]_\beta^\star$ as the minimum of

$$\sum_{j=1}^{m} \left(\sum_{i=1}^{r} |\lambda_{i,j}|^\beta \right)^{1/\beta} \tag{3.1}$$

over all m and all $\{\lambda_{i,j}\}$ for which there exist unit simple tensors $v_1, v_2, \ldots, v_m$ and decompositions $S_i = \sum_{j=1}^m \lambda_{i,j} v_j$, $i = 1, 2, \ldots, r$. For $\beta = \infty$, we define $[\mathbf{S}]_\infty^\star$ by replacing (3.1) by

$$\sum_{j=1}^{m} \max_{1 \leq i \leq r} |\lambda_{i,j}|.$$

Remark 3.8. The minimum in the previous definition is well-defined. To see this, consider the compact set

$$\left\{ (k_1 v, k_2 v, \dots k_r v) \middle| \sum_{i=1}^{r} |k_i|^\beta = 1, v \text{ is a unit simple tensor} \right\}.$$

and let B be its convex hull. For an r-tuple $\mathbf{S}$ of tensors, it is easy to see that $[\mathbf{S}]^\star_\beta$ is the infimum of all t such that $\mathbf{S} \in tB$. But if $\mathbf{S} \in tB$, then by Carathéodory's Convexity Theorem (see (*Barvinok*, 2002, Theorem 2.3)), we can find decompositions $S_i = \sum_{j=1}^{m} \lambda_{i,j} v_j$ with $\sum_{j=1}^{m} \left(\sum_{i=1}^{r} |\lambda_{i,j}|^\beta \right)^{1/\beta} \leq t$, where $m \leq \dim \mathbf{U}^r + 1$. So in Definition 3.7 we may take $m = \dim_{\mathbb{R}} \mathbf{U}^r + 1$. The set of all $\lambda_{i,j}$ and v_j for which $S_i = \sum_{j=1}^{m} \lambda_{i,j} v_j$ $(1 \leq i \leq r)$ is closed, so the function (3.1) has a minimum on this set.

We make the following key observation:

Proposition 3.9. If $1 \leq \alpha, \beta \leq \infty$ are Hölder conjugates (i.e $\frac{1}{\alpha} + \frac{1}{\beta} = 1$ or $\{\alpha, \beta\} = \{1, \infty\}$), then $[\cdot]_\alpha$ and $[\cdot]^\star_\beta$ are dual.

Proof. Let $\mathbf{S} = (S_1, \dots S_r)$ and $\mathbf{T} = (T_1, \dots T_r)$ be r-tuples of tensors, and write $T_i = \sum_j \lambda_{i,j} v_j$ with v_j simple unit tensors. Then, for $1 < \alpha, \beta < \infty$, using the Hölder inequality,

$$\begin{aligned}
|\langle \mathbf{S}, \mathbf{T} \rangle| &= \left| \sum_i \left\langle S_i, \sum_j \lambda_{i,j} v_j \right\rangle \right| \\
&\leq \sum_i \sum_j |\lambda_{i,j}| |\langle S_i, v_j \rangle| \\
&\leq \sum_j \left(\sum_i |\langle S_i, v_j \rangle|^\alpha \right)^{1/\alpha} \left(\sum_i |\lambda_{i,j}|^\beta \right)^{1/\beta} \\
&\leq [\mathbf{S}]_\alpha \sum_j \left(\sum_i |\lambda_{i,j}|^\beta \right)^{1/\beta}.
\end{aligned}$$

Taking the minimum over all decompositions $T_i = \sum_j \lambda_{i,j} v_j$ with v_j simple unit tensors gives $|\langle \mathbf{S}, \mathbf{T}\rangle| \le [\mathbf{S}]_\alpha [\mathbf{T}]^\star_\beta$. The same inequality holds for $\{\alpha, \beta\} = \{1, \infty\}$ using the same reasoning.

Again suppose that $1 < \alpha, \beta < \infty$. Let u be a simple unit tensor such that

$$[\mathbf{S}]_\alpha = \left(\sum_{i=1}^r |\langle S_i, u\rangle|^\alpha\right)^{1/\alpha}$$

and take $\mathbf{T} = (T_1, \dots, T_r)$ with $T_i = |\langle S_i, u\rangle|^{\alpha/\beta - 1}\overline{\langle S_i, u\rangle} u$, where $\overline{\langle S_i, u\rangle}$ denotes the complex conjugate of $\langle S_i, u\rangle$ (or just denotes $\langle S_i, u\rangle$ if our ground field is $\mathbb{R}$). Then by definition,

$$[\mathbf{T}]^\star_\beta \le \left(\sum_i \left(|\langle S_i, u\rangle|^{\alpha/\beta}\right)^\beta\right)^{1/\beta} = \left(\sum_i |\langle S_i, u\rangle|^\alpha\right)^{1/\beta}.$$

We also have

$$|\langle S_i, T_i\rangle| = |\langle S_i, u\rangle|^{\alpha/\beta+1} = |\langle S_i, u\rangle|^{\alpha/\beta+\alpha/\alpha} = |\langle S_i, u\rangle|^\alpha$$

and hence, since $\frac{1}{\alpha} + \frac{1}{\beta} = 1$,

$$\begin{aligned}
|\langle \mathbf{S}, \mathbf{T}\rangle| &= \sum_i |\langle S_i, u\rangle|^\alpha \\
&= \left(\sum_i |\langle S_i, u\rangle|^\alpha\right)^{1/\alpha} \left(\sum_i |\langle S_i, u\rangle|^\alpha\right)^{1/\beta} \\
&\ge [\mathbf{S}]_\alpha [\mathbf{T}]^\star_\beta.
\end{aligned}$$

Thus, given any r-tuple of tensors $\mathbf{S}$ we can construct a nonzero $\mathbf{T}$ such that $|\langle \mathbf{S}, \mathbf{T}\rangle| = [\mathbf{S}]_\alpha [\mathbf{T}]^\star_\beta$, and so the norms are dual, as claimed.

The proof for $\alpha = 1$ and $\beta = \infty$ is similar - in particular, if u is a unit simple tensor such that $[\mathbf{S}]_1 = \sum_{i=1}^r |\langle S_i, u\rangle|$ and we take $T_i = u$ for all i, then $|\langle \mathbf{S}, \mathbf{T}\rangle| = [\mathbf{S}]_1 = [\mathbf{S}]_1 [\mathbf{T}]^\star_\infty$.

Similarly, for $\alpha = \infty$ and $\beta = 1$, if u is a unit simple tensor such that $[\mathbf{S}]_\infty =$

$\max_i |\langle S_i, u\rangle| = |\langle S_k, u\rangle|$ for some k, then let $T_k = |\langle S_k, u\rangle| u$ and $T_i = 0$ otherwise. Then $|\langle \mathbf{S}, \mathbf{T}\rangle| = |\langle S_k, u\rangle|^2 = [\mathbf{S}]_\infty [\mathbf{T}]_1^\star$. $\square$

Proposition 3.10. If $S \in \mathbf{U}^r$, $T \in \mathbf{V}^s$, and $1 \le \beta \le \infty$, then $[S \boxtimes T]_\beta^\star = [S]_\beta^\star [T]_\beta^\star$.

Proof. Again combining Lemma 3.4 and Proposition 3.3, we see that $[S \boxtimes T]_\beta^\star \ge [S]_\beta^\star [T]_\beta^\star$.

Suppose that $1 \le \beta < \infty$. Let $u_1, \ldots, u_n$ be unit simple tensors such that $S_i = \sum_{j=1}^{n} \lambda_{i,j} u_j$ and

$$[S]_\beta^\star = \sum_{j=1}^{n} \left(\sum_{i=1}^{r} |\lambda_{i,j}|^\beta \right)^{1/\beta}$$

and similarly, let $v_1, \ldots, v_m$ be unit simple tensors such that $T_k = \sum_{l=1}^{m} \mu_{k,l} v_l$ and

$$[T]_\beta^\star = \sum_{l=1}^{m} \left(\sum_{k=1}^{s} |\mu_{k,l}|^\beta \right)^{1/\beta}.$$

Then $S_i \boxtimes T_k = \sum_{j=1}^{n} \sum_{l=1}^{m} \lambda_{i,j} \mu_{k,l} u_j \boxtimes v_l$ and so

$$\begin{aligned}
[S \boxtimes T]_\beta^\star &\le \sum_{j=1}^{n} \sum_{l=1}^{m} \left(\sum_{i=1}^{r} \sum_{k=1}^{s} |\lambda_{i,j}|^\beta |\mu_{k,l}|^\beta \right)^{1/\beta} \\
&= \sum_{j=1}^{n} \sum_{l=1}^{m} \left(\left(\sum_{i=1}^{r} |\lambda_{i,j}|^\beta \right)^{1/\beta} \left(\sum_{k=1}^{s} |\mu_{k,l}|^\beta \right)^{1/\beta} \right) \\
&= \sum_{j=1}^{n} \left(\sum_{i=1}^{r} |\lambda_{i,j}|^\beta \right)^{1/\beta} \sum_{l=1}^{m} \left(\sum_{k=1}^{s} |\mu_{k,l}|^\beta \right)^{1/\beta} \\
&= [S]_\beta^\star [T]_\beta^\star.
\end{aligned}$$

By the same method, we can show that $[S \boxtimes T]_\infty^\star \le [S]_\infty^\star [T]_\infty^\star$

Hence for all $1 \le \beta \le \infty$, we have $[S \boxtimes T]_\beta^\star = [S]_\beta^\star [T]_\beta^\star$. $\square$

This is not the only way in which we can extend the multiplicativity of the nuclear norm to other norms. We may consider how a wider class of norms - *projective* norms - behaves under the Kronecker product.

Definition 3.11. Given norms $\|\cdot\|_X$ and $\|\cdot\|_Y$ on tensor product spaces $\mathbf{U}$ and $\mathbf{V}$, we may form their *injective cross norm*, $\|\cdot\|_{X\vee Y}$, defined for $S \in \mathbf{U} \otimes \mathbf{V}$ by

$$\|S\|_{X\vee Y} := \sup\{|\langle S, u \otimes v\rangle|\}$$

where the supremum is taken over all $u \in \mathbf{U}$ and $v \in \mathbf{V}$ with $\|u\|_X \leq 1$ and $\|v\|_Y \leq 1$.

Similarly, the *projective cross norm* for these norms is $\|\cdot\|_{X\wedge Y}$, defined for $S \in \mathbf{U}\otimes\mathbf{V}$ by

$$\|S\|_{X\wedge Y} := \inf\left\{\sum_{i=1}^{k} \|u_i\|_X \|v_i\|_Y\right\}$$

where the infimum is taken over all decompositions $S = \sum_{i=1}^{k} u_i \otimes v_i$.

An important fact is shown in *Diestel et al.* (2008):

Proposition 3.12. Given norms $\|\cdot\|_X$ and $\|\cdot\|_Y$ on tensor product spaces $\mathbf{U}$ and $\mathbf{V}$, the injective cross norm is the smallest reasonable cross norm over them, and the projective cross norm is the largest.

Remark 3.13. Note that when our base norms are both the Frobenius norm, the injective cross norm is the spectral norm, and the projective cross norm is the nuclear norm. Proposition 3.12 tells us that these are the smallest and largest reasonable cross norms, respectively, over the Frobenius norm.

Proposition 3.14. Let $\|\cdot\|_{X_1}$, $\|\cdot\|_{X_2}$, $\|\cdot\|_{Y_1}$, and $\|\cdot\|_{Y_2}$ be norms on tensor product spaces $\mathbf{U}_1$, $\mathbf{U}_2$, $\mathbf{V}_1$, and $\mathbf{V}_2$ respectively. Let $S_1 \in \mathbf{U}_1 \otimes \mathbf{V}_1$ and $S_2 \in \mathbf{U}_2 \otimes \mathbf{V}_2$. Then

$$\|S_1 \boxtimes S_2\|_{(X_1\wedge X_2)\vee(Y_1\wedge Y_2)} = \|S_1\|_{X_1\vee Y_1}\|S_2\|_{X_2\vee Y_2}.$$

Proof. First, let $u_1 \in \mathbf{U}_1$, $u_2 \in \mathbf{U}_2$, $v_1 \in \mathbf{V}_1$, and $v_2 \in \mathbf{V}_2$ be such that $\|u_1\|_{X_1} = \|u_2\|_{X_2} = \|v_1\|_{Y_1} = \|v_2\|_{Y_2} = 1$ and

$$\|S_1\|_{X_1 \vee Y_1} = |\langle S_1, u_1 \otimes v_1 \rangle| \qquad \text{and} \qquad \|S_2\|_{X_2 \vee Y_2} = |\langle S_2, u_2 \otimes v_2 \rangle|.$$

Then for $B := u_1 \boxtimes u_2$ and $C := v_1 \boxtimes v_2$, since the injective cross norm is a cross norm, we have $\|B\|_{X_1 \wedge X_2} = \|C\|_{Y_1 \wedge Y_2} = 1$, and

$$|\langle S_1 \boxtimes S_2, B \otimes C \rangle| = \|S_1\|_{X_1 \vee Y_1} \|S_2\|_{X_2 \vee Y_2},$$

and hence

$$\|S_1 \boxtimes S_2\|_{(X_1 \wedge X_2) \vee (Y_1 \wedge Y_2)} \geq \|S_1\|_{X_1 \vee Y_1} \|S_2\|_{X_2 \vee Y_2}.$$

Conversely, suppose that $B \in \mathbf{U}_1 \boxtimes \mathbf{U}_2$ and $C \in \mathbf{V}_1 \boxtimes \mathbf{V}_2$ are such that $\|B\|_{X_1 \wedge X_2} \leq 1$ and $\|C\|_{Y_1 \wedge Y_2} \leq 1$. Pick decompositions

$$B = \sum_j b_{1,j} \boxtimes b_{2,j} \qquad \text{and} \qquad C = \sum_k c_{1,k} \boxtimes c_{2,k}$$

with $b_{1,j} \in \mathbf{U}_1$, $b_{2,j} \in \mathbf{U}_2$, $c_{1,k} \in \mathbf{V}_1$, and $c_{2,k} \in \mathbf{V}_2$. Then

$$\begin{aligned}
|\langle S_1 \boxtimes S_2, B \otimes C \rangle| &= \left| \sum_{j,k} \langle S_1, b_{1,j} \otimes c_{1,k} \rangle \langle S_2, b_{2,j} \otimes c_{2,k} \rangle \right| \\
&\leq \sum_{j,k} |\langle S_1, b_{1,j} \otimes c_{1,k} \rangle| \, |\langle S_2, b_{2,j} \otimes c_{2,k} \rangle| \\
&\leq \sum_{j,k} \|S_1\|_{X_1 \vee Y_1} \|b_{1,j}\|_{X_1} \|c_{1,k}\|_{Y_1} \|S_2\|_{X_2 \vee Y_2} \|b_{2,j}\|_{X_2} \|c_{2,k}\|_{Y_2} \\
&\leq \|S_1\|_{X_1 \vee Y_1} \|S_2\|_{X_2 \vee Y_2}
\end{aligned}$$

where the second inequality comes from the definition of the injective cross norm and the third comes from the fact that $\|B\|_{X_1 \wedge X_2} \leq 1$ and $\|C\|_{Y_1 \wedge Y_2} \leq 1$. Taking the supremum over all choices of B and C, we see that

$$\|S_1 \boxtimes S_2\|_{(X_1 \wedge X_2) \vee (Y_1 \wedge Y_2)} \leq \|S_1\|_{X_1 \vee Y_1} \|S_2\|_{X_2 \vee Y_2}.$$

□

Remark 3.15. Note that if all the base norms are the Frobenius norm then Proposition 3.14 tells us that the spectral norm is Kronecker-cross (as we already saw in Lemma 3.3).

To see that

$$\|S_1 \boxtimes S_2\|_{(X_1 \wedge X_2) \vee (Y_1 \wedge Y_2)} = \|S_1 \boxtimes S_2\|_{\sigma},$$

note first that for *simple* tensors u, v, we have $\|u\| = \|u\|_\star$ and $\|v\| = \|v\|_\star$, and hence

$$\|S_1 \boxtimes S_2\|_{(X_1 \wedge X_2) \vee (Y_1 \wedge Y_2)} \geq \|S_1 \boxtimes S_2\|_{\sigma}.$$

Conversely, for u, v with $\|u\|_\star = \|v\|_\star = 1$, we may find decompositions $u = \sum u_i$ and $v = \sum v_j$ where the u_i and v_j are simple tensors and $\sum \|u_i\| = \sum \|v_j\| = 1$, and so

$$\begin{aligned}
|\langle S_1 \boxtimes S_2, u \otimes v \rangle| &= |\langle S_1 \boxtimes S_2, \sum u_i \otimes v_j \rangle| \\
&\leq \sum |\langle S_1 \boxtimes S_2, u_i \otimes v_j \rangle| \\
&\leq \sum \|S_1 \boxtimes S_2\|_{\sigma} \|u_i\| \|v_j\| \\
&\leq \|S_1 \boxtimes S_2\|_{\sigma}.
\end{aligned}$$

Hence

$$\|S_1 \boxtimes S_2\|_{(X_1 \wedge X_2) \vee (Y_1 \wedge Y_2)} = \|S_1 \boxtimes S_2\|_{\sigma}.$$

CHAPTER IV

Slice Spectral and Slice Nuclear Norm

The nuclear norm has seen much use as a convex relaxation of tensor rank. As the slice rank has been of much recent interest, one may ask whether there is a similar convex relaxation of the slice rank. Looking for a norm whose unit ball is the convex hull of unit tensors of slice rank 1, we obtain the *slice nuclear norm*.

Definition 4.1. For $S \in \mathbf{V} = V_1 \otimes \ldots \otimes V_d$, let $\Phi_i(S) : (V_1 \otimes \ldots \otimes \hat{V}_i \otimes \ldots V_d)^\star \to V_i$ be the i-th flattening. The *slice nuclear norm* of S, denoted by $\|S\|_{\tilde{\star}}$ is the infimum of

$$\|\Phi_1(S_1)\|_\star + \|\Phi_2(S_2)\|_\star + \ldots + \|\Phi_d(S_d)\|_\star,$$

taken over all decompositions $S = S_1 + \ldots + S_d$.

Remark 4.2. It's important to note that this is in fact a norm - for the triangle inequality, if $S = S_1 + \ldots + S_d$ and $T = T_1 + \ldots + T_d$ are decompositions of tensors S and T, then a decomposition of $S + T$ is given by $(S_1 + T_1) + \ldots + (S_d + T_d)$, and so

$$\begin{aligned}\|S+T\|_{\tilde{\star}} &\leq \|\Phi_1(S_1+T_1)\|_\star + \ldots + \|\Phi_d(S_d+T_d)\|_\star \\ &\leq \|\Phi_1(S_1)\|_\star + \ldots + \|\Phi_d(S_d)\|_\star + \|\Phi_1(T_1)\|_\star + \ldots + \|\Phi_d(T_d)\|_\star.\end{aligned}$$

Taking infima over the decompositions of S and T, we see that $\|S+T\|_{\tilde{\star}} \leq \|S\|_{\tilde{\star}} + \|T\|_{\tilde{\star}}$

Remark 4.3. When searching for a decomposition for S in the above definition, it is clear that we may restrict ourselves to S_i with $\|S_i\|_\star \leq \|S\|_\star$. Therefore, we can view $\sum_i \|\Phi_i(S_i)\|_\star$ as a continuous function on a compact set, and hence by the Extreme Value Theorem it obtains its infimum. Therefore, we may replace the word 'infimum' in the definition by 'minimum'.

We also have a similar analogue of the spectral norm:

Definition 4.4. For $v \in \mathbf{V} = V_1 \otimes \ldots \otimes V_d$, let $\Phi_i(S) : (V_1 \otimes \ldots \otimes \hat{V}_i \otimes \ldots V_d)^\star \to V_i$ be the i-th flattening. The *slice spectral norm* of S, denoted by $\|S\|_{\tilde{\sigma}}$, is the maximum of $\|\Phi_i(S)\|_\sigma$, taken over all $1 \leq i \leq d$.

Proposition 4.5. The slice nuclear norm and the slice spectral norm are dual, i.e. for all $S, T \in \mathbf{V}$, we have $|\langle S, T\rangle| \leq \|S\|_{\tilde{\sigma}} \|T\|_{\tilde{\star}}$, and for every $S \in \mathbf{V}$, there exists some nonzero $T \in \mathbf{V}$ such that the above inequality becomes an equality.

Proof. First, let $S, T \in \mathbf{V}$ and suppose $T = T_1 + \cdots + T_d$ is a decomposition which satisfies $\|T\|_{\tilde{\star}} = \|\Phi_1(T_1)\|_\star + \cdots + \|\Phi_d(T_d)\|_\star$. Then

$$\begin{aligned}
|\langle S, T\rangle| &= \left| \sum_{i=1}^{d} \langle S, T_i\rangle \right| \\
&\leq \sum_{i=1}^{d} |\langle S, T_i\rangle| \\
&= \sum_{i=1}^{d} |\langle \Phi_i(S), \Phi_i(T_i)\rangle| \\
&\leq \sum_{i=1}^{d} \|\Phi_i(S)\|_\sigma \|\Phi_i(T_i)\|_\star \\
&\leq \max_i \|\Phi_i(S)\|_\sigma \sum_{i=1}^{d} \|\Phi_i(T_i)\|_\star \\
&= \|S\|_{\tilde{\sigma}} \|T\|_{\tilde{\star}}
\end{aligned}$$

where the second inequality follows from the duality of the spectral and nuclear norms.

Given $S \in \mathbf{V}$, let i be such that $\|S\|_{\tilde{\sigma}} = \|\Phi_i(S)\|_\sigma$. From the duality of the nuclear and spectral norms, there exists $T' \in V_1 \otimes \ldots \otimes \hat{V}_i \otimes \ldots V_d$ such that $|\langle \Phi_i(S), T' \rangle| = \|\Phi_i(S)\|_\sigma \|T'\|_\star$. Let $T = \Phi_i^{-1}(T')$. Then, using $T = 0 + \cdots + T + \cdots + 0$ as a decomposition, we have $\|T\|_{\tilde{\star}} \leq \|T'\|_\star$, and so

$$|\langle S, T \rangle| = |\langle \Phi_i(S), T' \rangle| = \|\Phi_i(S)\|_\sigma \|T'\|_\star \geq \|S\|_{\tilde{\sigma}} \|T\|_{\tilde{\star}}.$$

Hence we have found T such that $|\langle S, T \rangle| = \|S\|_{\tilde{\sigma}} \|T\|_{\tilde{\star}}$. □

Remark 4.6. It is natural to ask whether these norms are Kronecker-cross norms. It turns out that the answer is no:

Example 4.7. Consider the tensor $S = e_1 \otimes e_1 \otimes e_1 + e_1 \otimes e_2 \otimes e_2$. Then the singular values of both $\Phi_2(S)$ and $\Phi_3(S)$ are 1 (twice), and the only non-zero singular value of $\Phi_1(S)$ is $\sqrt{2}$. Hence $\|S\|_{\tilde{\sigma}} = \sqrt{2}$ and $\|S\|_{\tilde{\star}} \leq \sqrt{2}$. Similarly, for $T = e_1 \otimes e_1 \otimes e_1 + e_2 \otimes e_2 \otimes e_1$, we have $\|T\|_{\tilde{\sigma}} = \sqrt{2}$ and $\|T\|_{\tilde{\star}} \leq \sqrt{2}$.

However, for $S \boxtimes T$, we see that the non-zero singular values of both $\Phi_1(S \boxtimes T)$ and $\Phi_3(S \boxtimes T)$ are $\sqrt{2}$ (twice), and the non-zero singular values of $\Phi_2(S \boxtimes T)$ are 1 (four times). Hence $\|S \boxtimes T\|_{\tilde{\sigma}} = \sqrt{2}$. Since $\|S \boxtimes T\|^2 = 4$, we see by Proposition 4.5 that $\|S \boxtimes T\|_{\tilde{\star}} \geq \frac{4}{\sqrt{2}} = 2\sqrt{2}$.

It is easy to see that for any S and T, we have

$$\begin{aligned} \|S \boxtimes T\|_{\tilde{\sigma}} &= \max_i \left(\|\Phi_i(S)\|_\sigma \|\Phi_i(T)\|_\sigma \right) \\ &\leq \max_i \|\Phi_i(S)\|_\sigma \max_j \|\Phi_j(T)\|_\sigma \\ &= \|S\|_{\tilde{\sigma}} \|T\|_{\tilde{\sigma}}. \end{aligned}$$

By Lemma 3.4, it follows that $\|S \boxtimes T\|_{\tilde{\star}} \geq \|S\|_{\tilde{\star}} \|T\|_{\tilde{\star}}$. The above example shows that in general we do not have equality for either. We may say more for tensor Kronecker powers under the slice spectral norm, however.

Proposition 4.8. For any tensor S and positive integer n, we have $\|S^{\boxtimes n}\|_{\tilde{\sigma}} = \|S\|_{\tilde{\sigma}}^n$.

Proof. We have

$$\|S^{\boxtimes n}\|_{\tilde{\sigma}} = \max_i \left(\|\Phi_i(S)\|^n\right) = \max_i \left(\|\Phi_i(S)\|\right)^n = \|S\|_{\tilde{\sigma}}^n.$$

□

For the slice nuclear norm, we do not have multiplicativity, even just over tensor Kronecker powers:

Example 4.9. Let $S = \sum_{k=1}^n e_1 \otimes e_k \otimes e_k$ and let $T = \sum_{k=1}^n e_k \otimes e_k \otimes e_1$. Then, arguing as in Example 4.7, we see that $\|S\|_{\tilde{\star}} = \|T\|_{\tilde{\star}} = \|S \boxtimes T\|_{\tilde{\sigma}} = \sqrt{n}$ and so $\|S \boxtimes T\|_{\tilde{\star}} \geq n\sqrt{n}$.

Let $U = S \oplus T$. Then $\|U\|_{\tilde{\star}} = \|S \otimes 0 + 0 \otimes T\|_{\tilde{\star}} \leq \|S\|_{\tilde{\star}} + \|T\|_{\tilde{\star}} = 2\sqrt{n}$, and $\|U \boxtimes U\|_{\tilde{\star}} \geq \|S \boxtimes T\|_{\tilde{\star}} \geq n\sqrt{n}$.

Therefore, for $n \geq 17$, we see that $\|U \boxtimes U\|_{\tilde{\star}} > \|U\|_{\tilde{\star}}^2$ (in particular, there is no equality).

4.1 Stable ranks

There is a well-known surrogate of the rank of a matrix, known as the *stable rank* or *numerical rank* (*Rudelson and Vershynin* (2007)), which we will instead call the *stable spectral rank*. We show how the slice spectral and slice nuclear norms allow us to generalize the concept of stable rank to tensors, in a number of ways.

Definition 4.10. The *stable spectral rank* of a matrix A is

$$\left(\frac{\|A\|}{\|A\|_\sigma}\right)^2.$$

The spectral stable rank finds traction in numerics as it is less sensitive than the usual rank to perturbations in the smallest singular values (in particular to perturbations of any singular values of 0). Although this version of stable rank is the most common, there are some alternate but related notions:

Definition 4.11. The *stable nuclear rank* of a matrix A is

$$\left(\frac{\|A\|_\star}{\|A\|}\right)^2.$$

The *stable nuclear-spectral rank* of a matrix A is

$$\frac{\|A\|_\star}{\|A\|_\sigma}.$$

It is worth considering how these ranks compare for a given matrix:

Proposition 4.12. For any nonzero matrix A of rank r,

$$\left(\frac{\|A\|}{\|A\|_\sigma}\right)^2 \leq \frac{\|A\|_\star}{\|A\|_\sigma} \leq \left(\frac{\|A\|_\star}{\|A\|}\right)^2 \leq r.$$

Proof. Recall that the nuclear and spectral norms are dual, and so $\|A\|^2 \leq \|A\|_\sigma \|A\|_\star$. The first two inequalities result from rearranging this identity.

Suppose that A has rank r and $\{\lambda_1, \cdots, \lambda_r\}$ are the (non-zero) singular values of A, with $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r \geq 0$. Then recall that

$$\|A\|^2 = \sum_{i=1}^r \lambda_i^2$$
$$\|A\|_\star^2 = \left(\sum_{i=1}^r \lambda_i\right)^2.$$

By the generalized mean inequality,

$$\left(\frac{1}{r}\sum_{i=1}^{r}\lambda_i\right)^2 \leq \frac{1}{r}\sum_{i=1}^{r}\lambda_i^2$$

and so

$$\|A\|_\star^2 \leq r\|A\|^2,$$

as required. $\square$

Proposition 4.13. Let A be a matrix with stable nuclear rank s.

1. For $r \leq \frac{s}{2}$ there exists a rank r approximation A' to A with

$$\frac{\|A - A'\|^2}{\|A\|^2} \leq 1 - \frac{r}{s}. \tag{4.1}$$

2. For $r \geq \frac{s}{2}$ there exists a rank r approximation A' to A with

$$\frac{\|A - A'\|^2}{\|A\|^2} \leq \frac{s}{4r}. \tag{4.2}$$

Proof. We may find a singular value decomposition

$$A = \sum_{i=1}^{n}\sigma_i u_i \otimes v_i$$

for A with u_i, v_i unit vectors, $\{u_i\}$ pairwise orthogonal, $\{v_i\}$ pairwise orthogonal, and $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n$ the singular values of A.

Then $\|A\|^2 = \sum_{i=1}^{n}\sigma_i^2$ and $\|A\|_\star^2 = \left(\sum_{i=1}^{n}\sigma_i\right)^2$.

Let

$$A' = \sum_{i=1}^{r}\sigma_i u_i \otimes v_i.$$

Then $\|A - A'\|^2 = \sum_{i=r+1}^{n}\sigma_i^2$.

Firstly, the generalized mean inequality tells us that

$$\frac{1}{r}\left(\sum_{i=1}^{r}\sigma_i\right)^2 \leq \sum_{i=1}^{r}\sigma_i^2$$

and so

$$\frac{1}{r}\left(\sum_{i=1}^{r}\sigma_i\right)^2 \leq \|A\|^2 - \|A - A'\|^2.$$

Therefore,

$$\frac{\|A - A'\|^2}{\|A\|^2} \leq 1 - \frac{1}{r\|A\|^2}\left(\sum_{i=1}^{r}\sigma_i\right)^2. \tag{4.3}$$

However, there exists another upper bound. Since the singular values are in descending order, we certainly have

$$\frac{1}{r}\sum_{i=1}^{r}\sigma_i \geq \sigma_{r+1} \geq \sigma_{r+2} \geq \cdots \geq \sigma_n,$$

and so

$$\sum_{i=r+1}^{n}\sigma_i^2 \leq \frac{1}{r}\sum_{i=1}^{r}\sigma_i\sum_{i=r+1}^{n}\sigma_i = \frac{1}{r}\sum_{i=1}^{r}\sigma_i\left(\|A\|_\star - \sum_{i=1}^{r}\sigma_i\right).$$

Therefore,

$$\frac{\|A - A'\|^2}{\|A\|^2} \leq \frac{1}{r\|A\|^2}\sum_{i=1}^{r}\sigma_i\left(\sqrt{s}\|A\| - \sum_{i=1}^{r}\sigma_i\right). \tag{4.4}$$

Now both 4.3 and 4.4 are quadratics in $\sum_{i=1}^{r}\sigma_i$. It should be clear that the maximum upper bound in 4.4 occurs when $\sum_{i=1}^{r}\sigma_i = \frac{\sqrt{s}\|A\|}{2}$, and in that case, we have

$$\frac{\|A - A'\|^2}{\|A\|^2} \leq \frac{1}{r\|A\|^2}\left(\frac{\sqrt{s}\|A\|}{2}\right)^2 = \frac{s}{4r}.$$

Although this always gives an upper bound for the error in the approximation, it is not always the most efficient one. The bound from 4.3 is a tighter constraint than the one from 4.4 when

$$1 - \frac{1}{r\|A\|^2}\left(\sum_{i=1}^{r}\sigma_i\right)^2 \leq \frac{1}{r\|A\|^2}\sum_{i=1}^{r}\sigma_i\left(\sqrt{s}\|A\| - \sum_{i=1}^{r}\sigma_i\right)$$

which happens when $\sum_{i=1}^{r}\sigma_i \geq \frac{r\|A\|}{\sqrt{s}}$. In that case, from 4.3 we see that

$$\frac{\|A - A'\|^2}{\|A\|^2} \leq 1 - \frac{1}{r\|A\|^2}\left(\frac{r\|A\|}{\sqrt{s}}\right)^2 = 1 - \frac{r}{s}.$$

Of course, this is only relevant if $\sum_{i=1}^{r}\sigma_i$ is sufficiently small, i.e. if $\frac{r\|A\|}{\sqrt{s}} \leq \frac{\sqrt{s}\|A\|}{2}$. This happens when $r \leq \frac{s}{2}$, as claimed. $\square$

We may use our slice nuclear and spectral norms to introduce analogous versions of these stable ranks for higher order tensors.

Definition 4.14. The *stable spectral rank* of a tensor S is

$$\left(\frac{\|S\|}{\|S\|_\sigma}\right)^2.$$

The *stable nuclear rank* of a tensor S is

$$\left(\frac{\|S\|_\star}{\|S\|}\right)^2.$$

The *stable nuclear-spectral rank* of a tensor S is

$$\frac{\|S\|_\star}{\|S\|_\sigma}.$$

Remark 4.15. These stable ranks have a number of nice properties which are not shared by the usual tensor rank. For example, since the Frobenius, spectral and nuclear norms are Kronecker cross norms, these stable ranks are all multiplicative with respect to Kronecker tensor products.

A similar, but weaker, result to that of Proposition 4.12 holds for higher order tensors, again using the duality of the spectral and nuclear norms.

Proposition 4.16. For any nonzero tensor S,

$$\left(\frac{\|S\|}{\|S\|_\sigma}\right)^2 \leq \frac{\|S\|_\star}{\|S\|_\sigma} \leq \left(\frac{\|S\|_\star}{\|S\|}\right)^2.$$

Remark 4.17. A tensor S has stable spectral rank, stable nuclear, and stable nuclear-spectral rank all equal if and only if it satisfies $\|S\|^2 = \|S\|_\sigma \|S\|_\star$. Such tensors are sometimes said to be *unitangent* (*Derksen* (2018)).

Remark 4.18. Note that unlike the case for matrices, it is not true that for a tensor S of rank r, we have $\left(\frac{\|S\|_\star}{\|S\|}\right)^2 \leq r$. In particular, if S is the *matrix multiplication tensor* for $n \times n$ matrices

$$S = \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n} e_{i,j} \otimes e_{j,k} \otimes e_{i,k}$$

where $e_{i,j}$ is the matrix with 1 in the (i,j)-th position and 0s elsewhere, then it is known that $\|S\|^2 = n^3$ and $\|S\|_\star = n^3$ (*Derksen* (2016)), but the rank of S is strictly smaller than n^3 for sufficiently large n (*Landsberg* (2014)).

As in the case with matrices, we should expect that a tensor with small stable nuclear rank might be well approximated by low rank tensors. We will require some technology:

Definition 4.19. A k-sparse combination of vectors $v_1, v_2, \ldots, v_r \in \mathbb{R}^n$ is a vector $\sum_{i=1}^{r} \lambda_i v_i$ where $\lambda_i \neq 0$ for at most k values of i.

Lemma 4.20. Suppose that $w, v_1, v_2, \ldots, v_r \in \mathbb{R}^n$ are nonzero vectors and $w = v_1 + v_2 + \cdots + v_r$. Then there exists a nonzero k-sparse combination u of $v_1, v_2, \ldots, v_r$ such that

$$\sin^2\theta \leq \frac{C^2 - 1}{C^2 - 1 + k}$$

where θ is the angle between w and u and

$$C = \frac{||v_1|| + ||v_2|| + \cdots + ||v_r||}{||w||}.$$

Proof. Let $p_i = \|v_i\|$ for all i. We can scale w, so without loss of generality we may assume that $p_1 + p_2 + \cdots + p_r = 1$. Then $C = \|w\|^{-1}$. Let $u_i = p_i^{-1} v_i$, so that $w = \sum_{i=1}^{r} p_i u_i$ and $\|u_i\| = 1$ for all i. Let $\mathbf{U}_1, \mathbf{U}_2, \ldots, \mathbf{U}_k$ be independent random vectors, where $\mathbb{P}(\mathbf{U}_i = u_j) = p_j$ for all i and j. Then $\|\mathbf{U}_i\| = 1$ and the expected value of $\mathbf{U}_i$ is $\mathbb{E}(\mathbf{U}_i) = w$. We also have $\mathbb{E}(\langle \mathbf{U}_i, \mathbf{U}_j \rangle) = \langle \mathbb{E}(\mathbf{U}_i), \mathbb{E}(\mathbf{U}_j) \rangle = \langle w, w \rangle = \|w\|^2$ for $i \neq j$. Let A be the largest possible value of $\cos(\theta)$ where θ is the angle between w and $\mathbf{U} := \mathbf{U}_1 + \mathbf{U}_2 + \cdots + \mathbf{U}_k$. Note that the outcome of $\mathbf{U}$ is always a k-sparse combination of $v_1, v_2, \ldots, v_r$. We always have

$$\langle w, \mathbf{U} \rangle \leq A \|w\| \|\mathbf{U}\|$$

Taking the expected value gives

$$k\|w\|^2 = \langle w, kw \rangle = \langle w, \mathbb{E}(\mathbf{U}) \rangle \leq A\|w\| \mathbb{E}(\|\mathbf{U}\|).$$

Dividing by $\|w\|$ and squaring gives

$$k^2\|w\|^2 \leq A^2 \mathbb{E}(\|\mathbf{U}\|)^2 \leq A^2 \mathbb{E}(\|\mathbf{U}\|^2) = A^2 \sum_{1 \leq i,j \leq k} \mathbb{E}(\langle \mathbf{U}_i, \mathbf{U}_j \rangle) = A^2(k + (k^2 - k)\|w\|^2)$$

So

$$A^2 \geq \frac{k}{\|w\|^{-2} + (k-1)} = \frac{k}{C^2 - 1 + k}.$$

For some outcome of $\mathbf{U}$, $\cos(\theta) = A$ and

$$\sin^2(\theta) = 1 - A^2 \leq 1 - \frac{k}{C^2 - 1 + k} = \frac{C^2 - 1}{C^2 - 1 + k}.$$

□

Proposition 4.21. Let S be a tensor with stable nuclear rank s. Then there exists a rank r approximation, S_r, to S with

$$\frac{\|S - S_r\|^2}{\|S\|^2} \leq \frac{s-1}{s-1+r}.$$

Proof. Let $S = \sum_{i=1}^{n} v_i$ be a nuclear decomposition of S, i.e. $\|S\|_\star = \sum_{i=1}^{n} \|v_i\|$ and each v_i is a simple tensor. We may apply Lemma 4.20 with $w = S$. This results in a r-sparse combination (and hence a tensor of rank less than or equal to r) S_r with the angle θ between S and S_r satisfying

$$\sin^2 \theta \leq \frac{C^2 - 1}{C^2 - 1 + r}$$

where

$$C^2 = \left(\frac{||v_1|| + ||v_2|| + \cdots + ||v_n||}{||S_r||} \right)^2 = s.$$

Since $\|S - S_r\| = \|S\| \sin \theta$, we thus have

$$\frac{\|S - S_r\|^2}{\|S\|^2} \leq \frac{s-1}{s-1+r}.$$

□

For $r = 1$, this estimate gives a square error bound of $1 - \frac{1}{s}$, which matches the one found for matrices in Proposition 4.13.

4.2 Slice Stable Ranks

Although we have seen one way in which the stable ranks for matrices can be generalized to higher order tensors, our notion of slice nuclear and slice spectral norms provide an alternative. Whereas the stable ranks for matrices may be useful substitutes for the usual matrix rank, our slice stable ranks for higher order tensors are instead useful substitutes for *slice rank*.

Definition 4.22. The *stable slice spectral rank* of a tensor S is

$$\left(\frac{\|S\|}{\|S\|_{\tilde{\sigma}}}\right)^2.$$

The *stable slice nuclear rank* of a tensor S is

$$\left(\frac{\|S\|_{\tilde{\star}}}{\|S\|}\right)^2.$$

The *stable slice nuclear-spectral rank* of a tensor S is

$$\frac{\|S\|_{\tilde{\star}}}{\|S\|_{\tilde{\sigma}}}.$$

Remark 4.23. In Chapter II, we encountered Derksen's (*Derksen* (2020)) G-stable rank, and with $\alpha = (1, \cdots, 1)$, we saw that it can be computed over $\mathbb{C}$ as

$$\mathrm{rk}^G(S) = \sup_{g \in G} \min_i \frac{\|g \cdot S\|^2}{\|\Phi_i(g \cdot S)\|_\sigma^2} = \sup_{g \in G} \frac{\|g \cdot S\|^2}{\max_i \|\Phi_i(g \cdot S)\|_\sigma^2} = \sup_{g \in G} \frac{\|g \cdot S\|^2}{\|g \cdot S\|_{\tilde{\sigma}}^2}.$$

In other words, over $\mathbb{C}$, the G-stable rank of S is simply the supremum of the stable slice spectral rank of all tensors in the G-orbit of S.

Proposition 4.24. For any nonzero tensor S of slice rank s,

$$\left(\frac{\|S\|}{\|S\|_{\tilde{\sigma}}}\right)^2 \leq \frac{\|S\|_{\tilde{\star}}}{\|S\|_{\tilde{\sigma}}} \leq \left(\frac{\|S\|_{\tilde{\star}}}{\|S\|}\right)^2 \leq s.$$

Proof. As in the proof of Proposition 4.12, the first two inequalities follow from the duality of the slice spectral and slice nuclear norms.

To see that the stable slice nuclear norm is less than or equal to the slice rank, suppose S has slice rank s and let $S = S_1 + \cdots + S_r$ be a decomposition with $\mathrm{rk}(\Phi_1(S_1)) + \cdots + \mathrm{rk}(\Phi_r(S_r)) = s$ and such that the S_i are pairwise orthogonal. Then $\|S\|^2 = \sum_j \|S_j\|^2 = \sum_j \|\Phi_j(S_j)\|^2$.

By Proposition 4.12, for each S_i, we have $\mathrm{rk}(\Phi(S_i)) \geq \frac{\|\Phi_i(S_i)\|_\star^2}{\|\Phi_i(S_i)\|^2}$, and so

$$\begin{aligned} \mathrm{srk}(S) = \sum_{i=1}^{r} \mathrm{rk}(\Phi_i(S_i)) &\geq \sum_{i=1}^{r} \frac{\|\Phi_i(S_i)\|_\star^2}{\|\Phi_i(S_i)\|^2} \\ &\geq \frac{\left(\sum_i \|\Phi_i(S_i)\|_\star\right)^2}{\sum_j \|\Phi_j(S_j)\|^2} \\ &= \frac{\left(\sum_i \|\Phi_i(S_i)\|_\star\right)^2}{\|S\|^2} \end{aligned}$$

where the second inequality follows from Sedrakyan's inequality (itself a special case of the Cauchy-Schwartz inequality). Using the definition of the slice nuclear norm, we see that $\mathrm{srk}(S) \geq \frac{\|S\|_{\tilde{\star}}^2}{\|S\|^2}$. □

BIBLIOGRAPHY

BIBLIOGRAPHY

Appellof, C. J., and E. Davidson (1981), Strategies for analyzing data from video fluorometric monitoring of liquid chromatographic effluents, *Anal. Chem.*, *53*, 2053–2056.

Barvinok, A. (2002), *A course in convexity*, *Graduate Studies in Mathematics*, vol. 54, x+366 pp., American Mathematical Society, Providence, RI, doi:10.1090/gsm/054.

Bini, D., M. Capovani, F. Romani, and G. Lotti (1979), $O(n^{2.7799})$ complexity for $n \times n$ approximate matrix multiplication, *Inform. Process. Lett.*, *8*(5), 234–235, doi:10.1016/0020-0190(79)90113-3.

Bürgisser, P., M. Clausen, and M. A. Shokrollahi (1997), *Algebraic complexity theory*, *Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*, vol. 315, xxiv+618 pp., Springer-Verlag, Berlin, doi: 10.1007/978-3-662-03338-8, with the collaboration of Thomas Lickteig.

Candès, E., and B. Recht (2012), Exact matrix completion via convex optimization, *Commun. ACM*, *55*(6), 111–119, doi:10.1145/2184319.2184343.

Christandl, M., A. K. Jensen, and J. Zuiddam (2018), Tensor rank is not multiplicative under the tensor product, *Linear Algebra Appl.*, *543*, 125–139, doi:10.1016/j.laa.2017.12.020.

Christandl, M., F. Gesmundo, and A. K. Jensen (2019), Border rank is not multiplicative under the tensor product, *SIAM J. Appl. Algebra Geom.*, *3*(2), 231–255, doi:10.1137/18M1174829.

De Lathauwer, L., and B. De Moor (1998), From matrix to tensor: Multilinear algebra and signal processing, in *McWhirter and Proudler* (1998), pp. 1–15.

De Lathauwer, L., B. De Moor, and J. Vandewalle (2000), A multilinear singular value decomposition, *SIAM J. Matrix Anal. Appl.*, *21*, 1253–1278.

Defant, A., and K. Floret (1993), *Tensor norms and operator ideals*, *North-Holland Mathematics Studies*, vol. 176, xii+566 pp., North-Holland Publishing Co., Amsterdam.

Derksen, H. (2016), On the nuclear norm and the singular value decomposition of tensors, *Found. Comput. Math.*, *16*(3), 779–811, doi:10.1007/s10208-015-9264-x.

Derksen, H. (2018), A general theory of singular values with applications to signal denoising, *SIAM J. Appl. Algebra Geom.*, *2*(4), 535–596, doi:10.1137/17M1156149.

Derksen, H. (2020), The G-stable rank for tensors, arXiv:2002.08435.

Diestel, J., J. H. Fourie, and J. Swart (2008), *The metric theory of tensor products: Grothendieck's résumé revisited*, x+278 pp., American Mathematical Society, Providence, RI, doi:10.1090/mbk/052.

Ellenberg, J. S., and D. Gijswijt (2017), On large subsets of $\mathbb{F}_q^n$ with no three-term arithmetic progression, *Ann. of Math. (2)*, *185*(1), 339–343, doi:10.4007/annals.2017.185.1.8.

Friedland, S., and L.-H. Lim (2018), Nuclear norm of higher-order tensors, *Math. Comp.*, *87*(311), 1255–1281, doi:10.1090/mcom/3239.

Håstad, J. (1990), Tensor rank is np-complete, *J. Algorithms*, *11*, 644–654.

Heyden, A., G. Sparr, M. Nielsen, and P. Johansen (Eds.) (2002), *EECV 2002: Proceedings of the 7th European Conference on Computer Vision, Lecture Notes in Comput. Sci. 2350*, Springer.

Hitchcock, F. L. (1927), The expression of a tensor or a polyadic as a sum of products, *J. Math Phys.*, *7*, 39–79.

Kolda, T. G., and B. W. Bader (2009), Tensor decompositions and applications, *SIAM Review*, *51*(3), 455–500, doi:10.1137/07070111X.

Landsberg, J. M. (2012), *Tensors: geometry and applications*, *Graduate Studies in Mathematics*, vol. 128, xx+439 pp., American Mathematical Society, Providence, RI, doi:10.1090/gsm/128.

Landsberg, J. M. (2014), New lower bounds for the rank of matrix multiplication, *SIAM J. Comput.*, *43*(1), 144–149, doi:10.1137/120880276.

Lim, L.-H., and P. Comon (2014), Blind multilinear identification, *IEEE Trans. Inform. Theory*, *60*(2), 1260–1280, doi:10.1109/TIT.2013.2291876.

McWhirter, J., and I. Proudler (Eds.) (1998), *Mathematics in Signal Processing IV*, Clarendon Press.

Rudelson, M., and R. Vershynin (2007), Sampling from large matrices: an approach through geometric functional analysis, *J. ACM*, *54*(4), Art. 21, 19, doi:10.1145/1255443.1255449.

Schatten, R. (1950), *A Theory of Cross-Spaces*, Annals of Mathematics Studies, No. 26, vii+153 pp., Princeton University Press, Princeton, N. J.

Tao, T., and W. Sawin (2016), Notes on the "slice rank" of tensors, `https://terrytao.wordpress.com/2016/08/24/`.

Tucker, L. R. (1966), Some mathematical notes on three-mode factor analysis, *Psychometrika*, *31*, 279–311.

Vasilescu, M. A. O., and D. Terzopoulos (2002), Multilinear analysis of image ensembles: Tensorfaces, in *Heyden et al.* (2002), pp. 447–460.