Canonical Forms of $2 \times 2 \times 2$ and $2 \times 2 \times 2 \times 2$ Tensors

A Thesis Submitted to the
College of Graduate Studies and Research
in Partial Fulfillment of the Requirements
for the degree of Master of Science
in the Department of Mathematics and Statistics
University of Saskatchewan
Saskatoon

By

Stavros Georgios Stavrou

Permission to Use

In presenting this thesis in partial fulfilment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

Requests for permission to copy or to make other use of material in this thesis in whole or part should be addressed to:

Head of the Department of Mathematics and Statistics
142 McLean Hall
University of Saskatchewan
Saskatoon, Saskatchewan
Canada
S7N 5E6

Abstract

The rank and canonical forms of a tensor are concepts that naturally generalize that of a matrix. The question of how to determine the rank of a tensor has been widely studied in the literature and has no known solution in general. There are only a few specific cases that are known. In particular, the maximum rank of a $2 \times 2 \times 2$ tensor is 3. This fact was first proved by Kruskal. Later, ten Berge simplified the proof by providing a more straightforward argument. We provide another proof that is more simplified. As a corollary, a new upper bound on the rank of $2 \times \cdots \times 2$ tensors (with $n \geq 3$ factors) is $3 \cdot 2^{n-3}$.

For $2 \times 2 \times 2$ tensors, we consider their canonical forms over $\mathbb{R}, \mathbb{C}$ and some finite fields. We consider the direct product of the general linear groups and verify that over $\mathbb{R}$, these tensors are equivalent to eight canonical forms. When we consider the same problem over $\mathbb{C}$ there are seven canonical forms. These results were discovered independently many times in the literature. Using computer algebra for the case of finite fields, we additionally consider the action of the semidirect product of general linear groups with the symmetric group. For each canonical form, we determine the size of its orbit, and the rank of the tensors in its orbit over $\mathbb{F}_p$ for $p = 2, 3, 5$. These are original results. For larger primes, our computer did not have sufficient memory to finish the computations.

For $2 \times 2 \times 2 \times 2$ tensors, a finite classification of canonical forms over $\mathbb{R}$ and $\mathbb{C}$ is not possible. Instead, we use computer algebra and consider the semidirect product of general linear groups with the symmetric group and determine the canonical forms, the size of its orbit, and the rank of the tensors in its orbit over $\mathbb{F}_p$ for $p = 2, 3$. These are original results. For larger primes, the number of canonical forms becomes too large to be publishable.

Acknowledgements

There are two mathematicians who have greatly influenced my career path, whom I'd like to sincerely thank. Foremost, I would like to thank my supervisor, Professor Murray Bremner for his guidance and support. He is well-known in the department as an excellent teacher who displays enthusiasm, respect, and most of all patience to his colleagues and students. Thank you, Dr. Bremner, for the invaluable direction, motivation, and advice you provided during my time as a graduate student.

Secondly, I would like to thank Dr. Marina Tvalavadze for her positive influence. I had the pleasure of taking an Abstract Algebra course from her during my time as an undergraduate. Her style of teaching and the simple way that she presents material motivated me to study algebra at the graduate level. Thank you, Dr. Tvalavadze, for your helpful insights.

Thank you to the Head of the department, Raj Srinivasin, for ensuring that I always had work during my studies.

I thank my family and friends for their support, especially Tim Bodnarchuk who is always one step ahead of me, providing me with support and serenity.

I dedicate this work to my grandparents: Walter & Grace Kowal. Afierono diatrivi mou sti giagia mou Maria Stavrou, kai ston pappou mou Stavros Mylonas Kyriakou.

Contents

List of Tables

Chapter 1

Preliminaries

1.1 Introduction

A tensor is a multidimensional array of numbers. It was studied by Cayley in the late 1800's and remains an active area of research. The allure begins with trying to extend basic results of matrices from linear algebra to higher-order tensors. Concepts like tensor rank, canonical forms, and (hyper)determinants become complicated when they are generalized. Tensors have a wide variety of applications, ranging from chemometrics, multiway data analysis, signal processing, numerical analysis, and complexity theory. Although we do not elaborate on their applications, see [9, 22, 15].

Chapter 1 introduces the different topics that will be discussed. We define a kth-order tensor as an element of the outer product of k vector spaces. A first-order tensor is a vector and a second-order tensor is a matrix. Third-order tensors of size $p \times q \times r$ can be represented as r side-by-side $p \times q$ frontal matrix slices. Higher order tensors ($k \geq 3$) can be flattened so that the entries are re-written in matrix or vector form. We also discuss multiplication of tensors. In particular, given a third-order tensor, we can multiply by matrices on each of its three sides. This multilinear multiplication allows us to transform one tensor to another. Our next object of study is tensor rank, a concept generalizing that of a matrix. However, the properties of tensor rank are very different. Recall that for a matrix, its row rank, column rank, and outer product rank are all equal. For tensors, this is not true in general, and determining the outer product rank is a difficult problem with no known solution for arbitrary tensors. The multilinear rank of a tensor generalizes the concept of row and column rank of a matrix. Our focus is on outer product rank so when there is no risk of ambiguity we will refer to the outer product rank simply as rank. Another interesting feature is the dependence tensor rank has on the choice of base field.

Chapter 2 gives an analysis of $2 \times 2 \times 2$ tensors. We begin by discussing the maximum rank of such a tensor. Kruskal [24] proved first that the maximum rank of a $2 \times 2 \times 2$ tensor is 3. Later, J. ten Berge [36] provided a more simple proof of this fact. We also provide our own proof, which we believe is even more basic. As a corollary, an upper bound of the rank of a kth order $2 \times \cdots \times 2$ tensor is given by the formula $3 \cdot 2^{k-3}$.

The next important concept that allows us to classify $2 \times 2 \times 2$ tensors is the hyperdeterminant. Cayley's hyperdeterminant (also called Kruskal's polynomial) is a degree four polynomial in the eight coefficients of the tensor, which has 12 terms. Over $\mathbb{R}$, this hyperdeterminant gives information about the rank of the tensor: if the sign of the hyperdeterminant is negative then the rank is 3; if the sign is positive then the rank is 2; if the hyperdeterminant is zero then we do not know anything about its rank.

The next item of interest is canonical forms. We would like to know if there is a finite list of tensors such that every $2 \times 2 \times 2$ tensor is equivalent to exactly one of these. This classification has been done independently many times over $\mathbb{R}$ and $\mathbb{C}$. By applying the direct product of the general linear groups acting by simultaneous changes of basis along the three direction, we can reduce any tensor to exactly one of the canonical forms. Over $\mathbb{R}$ there are 8 forms, and over $\mathbb{C}$ there are 7. We use computer algebra to classify $2 \times 2 \times 2$ tensors over prime fields $\mathbb{F}_p$, for $p = 2, 3, 5$. Just as for the case of $\mathbb{R}$ and $\mathbb{C}$ we consider the action of the direct product of the general linear groups, which we call the small symmetry group. We also consider the action of the semidirect product of general linear groups with the symmetric group, which we call the large symmetry group. The actions of the symmetry groups decompose the set of tensors over finite fields into a disjoint union of orbits, where the tensors in each orbit are equivalent under the group action. The canonical form is the minimal element of each orbit. Our computer did not have enough memory for $p > 5$.

Given a general $2 \times 2 \times 2$ tensor, we can see that there is a relationship between the components of each vector in the outer product decomposition and the entries of the tensor. This allows us to see how the entries of the tensor effect its rank. For rank-1 tensors with only non-zero entries, we will see that the vectors in mode i (for $i = 1, 2, 3$) are scalar multiples of each other. We also consider the cases when there are zero entries. For a general rank-2 decomposition, we perform a series of substitutions in the equations relating the components of the vectors with the entries of the tensor and see that Cayley's hyperdeterminant shows up.

Chapter 3 gives a description of polynomial invariants and canonical forms for $2 \times 2 \times 2 \times 2$ tensors. We start by summarizing results by Thibon and Luque in which they construct a set of invariants which allows them to construct a closed form for the hyperdeterminant of a $2 \times 2 \times 2 \times 2$ tensor.

We determine the canonical forms of these tensors over the finite fields $\mathbb{F}_2$ and $\mathbb{F}_3$. The Maple program used for $2 \times 2 \times 2$ tensors is not efficient when it is generalized to $2 \times 2 \times 2 \times 2$ tensors. In order to compute the ranks and canonical forms, more sophisticated programming techniques are used. Using the large symmetry group we first compute the orbits, and then determine the rank of each orbit. For $p = 2$, the maximum rank is 6 and there are 65536 tensors partitioned between 30 orbits. For $p = 3$, the maximum rank is 5 and there are 43046721 tensors partitioned between 49 orbits.

1.2 Matrix Products

We begin by discussing different types of matrix products, which we collect from [22]. Then we explain how tensors are represented, tensor products, and introduce the notion of tensor rank.

A *matrix*, which we will denote by an uppercase letter (e.g. A), is a collection of elements ordered by rows and columns. The matrix

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$$

has m rows and n columns, so we say it is an "m by n" matrix. For each entry of the matrix, the first subscript refers to the row and the second subscript refers to the column. Sometimes we use the abbreviation $A = [a_{ij}]_{i,j=1}^{i=m,j=n}$ for a matrix with elements a_{ij}. Let $\mathbb{F}^{m\times n}$ denote the set of $m \times n$ matrices with entries in the field $\mathbb{F}$. A vector is a matrix with one row or one column. We will denote vectors by lowercase letters (e.g. u), and represent the space of length n vectors by $\mathbb{F}^n$.

Definition 1.1. The *tensor product* of two column vectors $u \in \mathbb{F}^n$ and $v \in \mathbb{F}^m$, denoted $u \otimes v$, is a column vector of length nm given by

$$u \otimes v = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \otimes \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_m \end{bmatrix} = \begin{bmatrix} u_1 v \\ u_2 v \\ \vdots \\ u_n v \end{bmatrix}$$

Definition 1.2. The *Kronecker product* of two matrices $A \in \mathbb{F}^{I\times J}$ and $B \in \mathbb{F}^{K\times L}$, denoted $A \otimes B$, is a matrix of size $IK \times JL$ defined by

$$A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & \dots & a_{1J}B \\ a_{21}B & a_{22}B & \dots & a_{2J}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{I1}B & a_{I2}B & \dots & a_{IJ}B \end{bmatrix}$$
$$= \begin{bmatrix} a_1 \otimes b_1 & a_1 \otimes b_2 & \dots & a_J \otimes b_{L-1} & a_J \otimes b_L \end{bmatrix}$$

where $a_1, \dots, a_J$ and $b_1, \dots, b_L$ are the columns of matrix A and B, respectively.

Informally, we refer to this matrix product as the *tensor product* or *outer product.*

Example 1.3. Let $A \in \mathbb{R}^{2\times 2}$ and $B \in \mathbb{R}^{2\times 3}$. The result is a matrix of size 4×6.

$$A \otimes B = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \otimes \begin{bmatrix} 5 & 6 \\ 7 & 8 \\ 9 & 10 \end{bmatrix} = \begin{bmatrix} 1 \cdot B & 2 \cdot B \\ 3 \cdot B & 4 \cdot B \end{bmatrix} = \begin{bmatrix} 5 & 6 & 10 & 12 \\ 7 & 8 & 14 & 16 \\ 9 & 10 & 18 & 20 \\ 15 & 18 & 20 & 24 \\ 21 & 24 & 28 & 32 \\ 27 & 30 & 36 & 40 \end{bmatrix}$$

Definition 1.4. The *Khatri-Rao product* of two matrices $A \in \mathbb{F}^{I\times K}$ and $B \in \mathbb{F}^{J\times K}$, denoted $A \odot B$, is a matrix of size $IJ \times K$ defined by

$$A \odot B = \begin{bmatrix} a_1 \otimes b_1 & a_2 \otimes b_2 & \dots & a_K \otimes b_K \end{bmatrix}$$

Example 1.5. Using the same A and B from the previous example (since the column dimension of both matrices are equal), their Khatri-Rao product yeilds a 4×3 matrix:

$$A \odot B = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \odot \begin{bmatrix} 5 & 6 \\ 7 & 8 \\ 9 & 10 \end{bmatrix} = \begin{bmatrix} 5 & 12 \\ 7 & 16 \\ 9 & 20 \\ 15 & 24 \\ 21 & 32 \\ 27 & 40 \end{bmatrix}$$

Definition 1.6. The *Hadamard product* of $A, B \in \mathbb{F}^{I \times J}$, denoted $A \star B$, is an element-wise product that yields an $I \times J$ matrix defined by

$$A \star B = \begin{bmatrix} a_{11}b_{11} & a_{12}b_{12} & \dots & a_{1J}b_{1J} \\ a_{21}b_{21} & a_{22}b_{22} & \dots & a_{2J}b_{2J} \\ \vdots & \vdots & \ddots & \vdots \\ a_{I1}b_{I1} & a_{I2}b_{I2} & \dots & a_{IJ}b_{IJ} \end{bmatrix}$$

Example 1.7. Using matrix B from the previous examples,

$$B \star B = \begin{bmatrix} 5 & 6 \\ 7 & 8 \\ 9 & 10 \end{bmatrix} \star \begin{bmatrix} 5 & 6 \\ 7 & 8 \\ 9 & 10 \end{bmatrix} = \begin{bmatrix} 25 & 36 \\ 49 & 64 \\ 81 & 100 \end{bmatrix}$$

1.3 Representing Tensors

Definition 1.8. An order-k *tensor* $\boldsymbol{X}$ is an element of the *tensor product* of k vector spaces $V_1 \otimes V_2 \otimes \cdots \otimes V_k$, where the *order* refers to the number of dimensions or modes.

This is an abstract definition that does not depend on a choice of basis in each vector space. This is one of the motivations for studying the canonical forms of tensors, since they represent properties of tensors which do not depend on the choice of basis. Once we fix a basis in each vector space, we can associate to each order-k tensor a k-dimensional array. Then an order-1 tensor is a vector, and an order-2 tensor is a matrix. Recall that we represent vectors by plain lowercase letters (e.g., u), and matrices by uppercase letters (e.g., A). We represent tensors whose order is 3 or greater by boldface capital letters (e.g., $\boldsymbol{X}$). The $(i_1, \dots, i_k)$-th entry of the order-k tensor $\boldsymbol{X}$ is denoted by $x_{i_1,\dots,i_k}$. A matrix (order-2 tensor) is two dimensional so we can consider the set of row vectors and the set of column vectors. For a general order-k tensor there are k dimensions (or modes) and thus k mutually orthogonally displayed sets of vectors.

Definition 1.9. [22] Let $\boldsymbol{X}$ be an order-k tensor. A *fibre* is a vector in one of the k modes of the tensor, formed by fixing $k-1$ of the indices.

These fibres are the higher-order analogue of matrix rows and columns. Let $\boldsymbol{X} \in \mathbb{R}^{d_1 \times d_2 \times d_3}$ be an order-3 tensor. We refer to the mode-1,2,3 fibres as the column, row, and tube fibres, respectively. The set of column, row, and tube fibres are given respectively by

$$\{x_{\bullet i_2 i_3} \in \mathbb{R}^{d_1} \mid 1 \leq i_2 \leq d_2, 1 \leq i_3 \leq d_3\},$$
$$\{x_{i_1 \bullet i_3} \in \mathbb{R}^{d_2} \mid 1 \leq i_1 \leq d_1, 1 \leq i_3 \leq d_3\},$$
$$\{x_{i_1 i_2 \bullet} \in \mathbb{R}^{d_3} \mid 1 \leq i_1 \leq d_1, 1 \leq i_2 \leq d_2\}.$$

Definition 1.10. [22] Let $\boldsymbol{X}$ be an order-k tensor, with $k > 2$. A *slice* or *slab* is a matrix in one of the k modes of the tensor, formed by fixing $k-2$ of the indices.

For order-3 tensors, the three sets of matrices

$$\{\boldsymbol{X}_{\bullet\bullet i_3} \in \mathbb{F}^{d_1 \times d_2} \mid 1 \leq i_3 \leq d_3\},$$
$$\{\boldsymbol{X}_{\bullet i_2 \bullet} \in \mathbb{F}^{d_2 \times d_3} \mid 1 \leq i_2 \leq d_2\},$$
$$\{\boldsymbol{X}_{i_1 \bullet\bullet} \in \mathbb{F}^{d_1 \times d_3} \mid 1 \leq i_1 \leq d_1\}.$$

give the frontal, vertical, and horizontal slices of the tensor, respectively. This notation allows us to represent an order-3 tensor in terms of its frontal slices. Let $\boldsymbol{X} \in \mathbb{F}^{d_1 \times d_2 \times d_3}$. Then $\boldsymbol{X}$ has d_1 horizontal slices, d_2 vertical slices, and d_3 frontal slices:

$$\boldsymbol{X} = \left[\begin{array}{c|c|c} \boldsymbol{X}_1 & \dots & \boldsymbol{X}_{d_3} \end{array}\right]$$
$$= \left[\begin{array}{ccc|ccc|ccc} x_{111} & \dots & x_{1d_21} & \dots & & \dots & x_{11d_3} & \dots & x_{1d_2d_3} \\ \vdots & & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ x_{d_111} & \dots & x_{d_1d_21} & \dots & & \dots & x_{d_11d_3} & \dots & x_{d_1d_2d_3} \end{array}\right]$$

1.4 Transforming a Tensor into a Matrix or Vector

Definition 1.11. [22] *Matricization* is the *flattening* of a tensor by reordering the elements into a matrix.

An order-N tensor $\boldsymbol{X} \in \mathbb{F}^{d_1 \times \cdots \times d_N}$ can be matricized in N different ways (that is, along the N different modes). The mode-n matricization, denoted $\boldsymbol{X}_{(n)}$, maps the tensor element $(i_1, i_2, \dots, i_N)$ to the matrix element (i_n, j), where

$$j = 1 + \sum_{k=1, k\neq n}^{N} (i_k - 1) J_k \quad \text{with} \quad J_k = \prod_{m=1, m\neq n}^{k-1} i_m.$$

The size of $\boldsymbol{X}_{(n)}$ is $i_n \times i_1 \dots i_{n-1} i_{n+1} \dots i_N$ [22]. This notation is awkward but the concept is easy to understand. Consider the following example.

Example 1.12. Let $\boldsymbol{X} \in \mathbb{R}^{4\times3\times2}$ be given in terms of its frontal slices:

$$\boldsymbol{X} = \left[\begin{array}{ccc|ccc} 0 & 1 & 2 & 3 & 4 & 5 \\ 6 & 7 & 8 & 9 & 10 & 11 \\ 12 & 13 & 14 & 15 & 16 & 17 \\ 18 & 19 & 20 & 21 & 22 & 23 \end{array}\right]$$

Since the order of this tensor is 3, there are 3 ways to matricize this tensor. The mode-1 matricization writes the columns of the tensor as the columns of a 4×6

matrix:

$$X_{(1)} = \begin{bmatrix} 0 & 1 & 2 & 3 & 4 & 5 \\ 6 & 7 & 8 & 9 & 10 & 11 \\ 12 & 13 & 14 & 15 & 16 & 17 \\ 18 & 19 & 20 & 21 & 22 & 23 \end{bmatrix}$$

The mode-2 matricization writes the rows of the tensor as the columns of a 3×8 matrix:

$$X_{(2)} = \begin{bmatrix} 0 & 6 & 12 & 18 & 3 & 9 & 15 & 21 \\ 1 & 7 & 13 & 19 & 4 & 10 & 16 & 22 \\ 2 & 8 & 14 & 20 & 5 & 11 & 17 & 23 \end{bmatrix}$$

The mode-3 matricization writes the tubes of the tensor as the columns of a 2×12 matrix:

$$X_{(3)} = \begin{bmatrix} 0 & 6 & 12 & 18 & 1 & 7 & 13 & 19 & 2 & 8 & 14 & 20 \\ 3 & 9 & 15 & 21 & 4 & 10 & 16 & 22 & 5 & 11 & 17 & 23 \end{bmatrix}$$

Definition 1.13. [22] The *vectorization* of a tensor writes the columns of the tensor as a vector.

Example 1.14. Let $\boldsymbol{X} \in \mathbb{R}^{2\times2\times2}$ be given by

$$\boldsymbol{X} = \left[\begin{array}{cc|cc} 1 & 3 & 5 & 7 \\ 2 & 4 & 6 & 8 \end{array}\right]$$

Then the vectorization of $\boldsymbol{X}$, denoted $\text{vec}(\boldsymbol{X})$, is

$$\text{vec}(\boldsymbol{X}) = \begin{bmatrix} 1 \\ 2 \\ \vdots \\ 8 \end{bmatrix}$$

1.5 Tensor Multiplication

In this section we discuss tensor-matrix multiplication, also known as multilinear matrix multiplication, as well as tensor-tensor multiplication via the outer product. Recall that when multiplying a matrix by another matrix, there are two independent multiplication operations: left-multiplication and right-multiplication. When we multiply an order-k tensor by a matrix, there are k different multiplication operations: one per mode [9]. If $\boldsymbol{X} = [x_{i_1\ldots i_k}] \in \mathbb{F}^{d_1\times\cdots\times d_k}$ and $A_1 = [a_{ij}^{(1)}] \in \mathbb{F}^{a_1\times d_1}, \ldots, A_k = [a_{ij}^{(k)}] \in \mathbb{F}^{a_k\times d_k}$, then $\boldsymbol{Y} = (A_1,\ldots,A_k)\cdot\boldsymbol{X}$ is the new tensor $\boldsymbol{Y} = [y_{i_1\ldots i_k}] \in \mathbb{F}^{a_1\times\cdots\times a_k}$ defined by

$$y_{i_1\ldots i_k} = \sum_{i_1,\ldots,i_k=1}^{d_1,\ldots,d_k} a_{i_1j_1}\ldots a_{i_kj_k}x_{j_1\ldots j_k},$$

where the matrix A_l acts on the mode-l fibres. So in particular, if $\boldsymbol{X} = [x_{ijk}] \in \mathbb{F}^{d_1\times d_2\times d_3}$ and $L = [\lambda_{pi}] \in \mathbb{F}^{c_1\times d_1}, M = [\mu_{qj}] \in \mathbb{R}^{c_2\times d_2}, N = [\nu_{rk}] \in \mathbb{F}^{c_3\times d_3}$, then the tensor $\boldsymbol{X}$ can be transformed into a new tensor $\boldsymbol{Y} = [y_{pqr}] \in \mathbb{F}^{c_1\times c_2\times c_3}$ via the multiplication $\boldsymbol{Y} = (L,M,N)\cdot\boldsymbol{X}$ defined by

$$y_{pqr} = \sum_{i,j,k=1}^{d_1,d_2,d_3} \lambda_{pi}\mu_{qj}\nu_{rk}x_{ijk}.$$

We make this particular case explicit because we will use this multiplication on order-3 tensors often throughout this and later discussions. Let's see an example of how matrices act on the modes of an order-3 tensor.

Example 1.15. Let $\boldsymbol{X} \in \mathbb{R}^{3\times3\times2}$ be given by

$$\boldsymbol{X} = \left[\begin{array}{ccc|ccc} 1 & 2 & 1 & 0 & 2 & 2 \\ 0 & 1 & 1 & 1 & 4 & 5 \\ 2 & 3 & 0 & 3 & 3 & 0 \end{array}\right]$$

and let $L \in \mathbb{R}^{2\times 3}$, $M \in \mathbb{R}^{4\times 3}$, and $N \in \mathbb{R}^{3\times 2}$ be the following matrices

$$L = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 0 & 2 \end{bmatrix}, \qquad M = \begin{bmatrix} 0 & 0 & 1 \\ 2 & 1 & 3 \\ 5 & 5 & 1 \\ 1 & 0 & 8 \end{bmatrix}, \qquad N = \begin{bmatrix} 3 & 5 \\ 5 & 7 \\ 7 & 0 \end{bmatrix}.$$

Then the multilinear multiplication of $\boldsymbol{X}$ by matrices L, M, N yields a $2\times 4\times 3$ tensor. We'll compute this in steps, starting with multiplication of the tensor $\boldsymbol{X}$ by the matrix L. For this we want the mode-1 fibres to be columns in the matrix representation of the tensor. We have

$$L\boldsymbol{X} = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 0 & 2 \end{bmatrix} \left[\begin{array}{ccc|ccc} 1 & 2 & 1 & 0 & 2 & 2 \\ 0 & 1 & 1 & 1 & 4 & 5 \\ 2 & 3 & 0 & 3 & 3 & 0 \end{array}\right] = \left[\begin{array}{ccc|ccc} 3 & 6 & 2 & 4 & 9 & 7 \\ 4 & 10 & 2 & 6 & 10 & 4 \end{array}\right] =: \boldsymbol{P}.$$

We now take the $2 \times 3 \times 2$ tensor $\boldsymbol{P}$ and write it so that the mode-2 fibres are the columns:

$$\boldsymbol{P} = \left[\begin{array}{cc|cc} 3 & 4 & 4 & 6 \\ 6 & 10 & 9 & 10 \\ 2 & 2 & 7 & 4 \end{array}\right].$$

We can now multiply (on the left) by the matrix M:

$$MP = \begin{bmatrix} 0 & 0 & 1 \\ 2 & 1 & 3 \\ 5 & 5 & 1 \\ 1 & 0 & 8 \end{bmatrix} \left[\begin{array}{cc|cc} 3 & 4 & 4 & 6 \\ 6 & 10 & 9 & 10 \\ 2 & 2 & 7 & 4 \end{array}\right] = \left[\begin{array}{cc|cc} 2 & 2 & 7 & 4 \\ 18 & 24 & 38 & 34 \\ 47 & 72 & 72 & 84 \\ 19 & 20 & 60 & 38 \end{array}\right] =: \boldsymbol{Q}.$$

Putting this back into standard form of the frontal slices, we obtain

$$\boldsymbol{Q} = \left[\begin{array}{cccc|cccc} 2 & 18 & 47 & 19 & 7 & 38 & 72 & 60 \\ 2 & 24 & 72 & 20 & 4 & 34 & 84 & 38 \end{array}\right].$$

We now rewrite the $2 \times 4 \times 2$ tensor $\boldsymbol{Q}$ in terms of its horizontal slices so that its mode-3 fibres are the columns:

$$\boldsymbol{Q} = \left[\begin{array}{cccc|cccc} 2 & 18 & 47 & 19 & 2 & 24 & 72 & 20 \\ 7 & 38 & 72 & 60 & 4 & 34 & 84 & 38 \end{array}\right].$$

We now multiply (on the left) by the matrix N to obtain

$$\begin{aligned} N\boldsymbol{Q} &= \begin{bmatrix} 3 & 5 \\ 5 & 7 \\ 7 & 0 \end{bmatrix} \left[\begin{array}{cccc|cccc} 2 & 18 & 47 & 19 & 2 & 24 & 72 & 20 \\ 7 & 38 & 72 & 60 & 4 & 34 & 84 & 38 \end{array}\right] \\ &= \left[\begin{array}{cccc|cccc} 41 & 244 & 501 & 357 & 26 & 242 & 636 & 250 \\ 59 & 356 & 739 & 515 & 38 & 358 & 948 & 366 \\ 14 & 126 & 329 & 133 & 14 & 168 & 504 & 140 \end{array}\right] =: \boldsymbol{R}. \end{aligned}$$

We must now re-write $\boldsymbol{R}$ in its standard form (i.e. we must write the columns as the mode-3 fibres). This gives the result

$$\boldsymbol{Y} = \left[\begin{array}{cccc|cccc|cccc} 41 & 244 & 501 & 357 & 59 & 356 & 739 & 515 & 14 & 126 & 329 & 133 \\ 26 & 242 & 636 & 250 & 38 & 358 & 948 & 366 & 14 & 168 & 504 & 140 \end{array}\right].$$

Definition 1.16. [9] Let $\boldsymbol{X} \in \mathbb{F}^{d_1 \times \cdots \times d_k}$ be an order-k tensor and $\boldsymbol{Y} \in \mathbb{F}^{c_1 \times \cdots \times c_h}$ be an order-h tensor; then the outer product of $\boldsymbol{X}$ and $\boldsymbol{Y}$ is the tensor $\boldsymbol{Z} := \boldsymbol{X} \otimes \boldsymbol{Y} \in \mathbb{F}^{d_1 \times \cdots \times d_k \times c_1 \times \cdots \times c_h}$ of order $k + h$ defined by

$$z_{i_1 \ldots i_k j_1 \ldots j_h} = x_{i_1 \ldots i_k} y_{j_1 \ldots j_h}. \tag{1.1}$$

We can also define the direct sum of the order-k tensors $\boldsymbol{X} \in \mathbb{F}^{d_1 \times \cdots \times d_k}$ and $\boldsymbol{Y} \in \mathbb{F}^{c_1 \times \cdots \times c_k}$ (notice that addition is defined only for tensors of the same order). The result is an order-k tensor $\boldsymbol{Z} := \boldsymbol{X} \oplus \boldsymbol{Y} \in \mathbb{F}^{(c_1+d_1) \times \cdots \times (c_k+d_k)}$ defined by

$$z_{i_1,\ldots,i_k} = \begin{cases} x_{i_1,\ldots,i_k} & \text{if } 1 \leq i_\alpha \leq d_\alpha,\ \alpha = 1, \ldots, k \\ y_{i_1-d_1,\ldots,i_k-d_k} & \text{if } d_\alpha + 1 \leq i_\alpha \leq c_\alpha + d_\alpha,\ \alpha = 1, \ldots, k \\ 0 & \text{otherwise.} \end{cases}$$

For matrices, the direct sum of $A \in \mathbb{F}^{m_1 \times n_1}$ and $B \in \mathbb{F}^{m_2 \times n_2}$ is the block-diagonal matrix

$$A \oplus B = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \in \mathbb{F}^{(m_1+m_2)\times(n_1+n_2)}.$$

The direct sum of the two tensors $\boldsymbol{X}$ and $\boldsymbol{Y}$ defined above can be written symbolically as

$$\boldsymbol{Z} = \boldsymbol{X} \oplus \boldsymbol{Y} = \left[\begin{array}{cc|cc} \boldsymbol{X}_{c_1,c_2,c_3} & \boldsymbol{O} & \boldsymbol{O} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} & \boldsymbol{O} & \boldsymbol{Y}_{d_1,d_2,d_3} \end{array}\right].$$

Example 1.17. Let $\boldsymbol{X} \in \mathbb{R}^{2\times3\times2}$ and $\boldsymbol{Y} \in \mathbb{R}^{2\times2\times2}$ be order-3 tensors represented by

$$\boldsymbol{X} = \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 2 & 2 & 0 \\ 1 & 0 & 2 & 0 & 0 & 1 \end{array}\right]$$

and

$$\boldsymbol{Y} = \left[\begin{array}{cc|cc} 1 & 2 & 3 & 0 \\ 2 & 0 & 1 & 4 \end{array}\right].$$

Then $\boldsymbol{Z} = \boldsymbol{X} \otimes \boldsymbol{Y} \in \mathbb{R}^{2\times3\times2\times2\times2\times2}$. $\boldsymbol{Z}$ is an order-6 tensor with 96 entries. The first entry is $z_{111111} = x_{111}y_{111} = (1)(1) = 1$, entry $z_{211122} = x_{211}y_{122} = (1)(0) = 0$, and the last entry is $z_{232222} = x_{232}y_{222} = (1)(4) = 4$. Using the same tensors $\boldsymbol{X}$ and $\boldsymbol{Y}$ defined above, since they are both order-3, we can perform the direct sum to obtain the order-3 tensor $\boldsymbol{X} \oplus \boldsymbol{Y} \in \mathbb{R}^{4\times5\times4}$ given by

$$\begin{aligned} \boldsymbol{X} \oplus \boldsymbol{Y} &= \left[\begin{array}{cc|cc} \boldsymbol{X}_{2,3,2} & \boldsymbol{O} & \boldsymbol{O} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} & \boldsymbol{O} & \boldsymbol{Y}_{2,2,2} \end{array}\right] \\ &= \left[\begin{array}{ccccc|ccccc|ccccc|ccccc} 1 & 2 & 3 & 0 & 0 & 2 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 2 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 2 & 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 1 & 4 \end{array}\right] \end{aligned}$$

1.6 Tensor Rank

In this section we'll discuss various notions of tensor rank and the related issues. We begin by considering multilinear rank. This concept generalizes the column and row ranks of a matrix to higher order tensors. We will see that, unlike for matrices where the row rank, column rank, and outer product rank are all equal, this is not the case for higher order tensors [9]. In general, to describe the multilinear rank of an order-k tensor in $\mathbb{F}^{d_1 \times \cdots \times d_k}$ we fix $k-1$ of the indices thereby creating k sets of fibres:

$$\begin{aligned} \boldsymbol{X}_{\bullet i_2 \dots i_k} &:= [x_{i_1 \dots i_k}]_{i_1=1}^{d_1} \in \mathbb{F}^{d_1} \\ \boldsymbol{X}_{i_1 \bullet \dots i_k} &:= [x_{i_1 \dots i_k}]_{i_2=1}^{d_2} \in \mathbb{F}^{d_2} \\ &\vdots \\ \boldsymbol{X}_{i_1 \cdots \bullet} &:= [x_{i_1 \dots i_k}]_{i_k=1}^{d_k} \in \mathbb{F}^{d_k} \end{aligned}$$

Then in analogy with row and column rank, define

$$\begin{aligned} r_1(\boldsymbol{X}) &:= \dim(\operatorname{span}_{\mathbb{F}}\{X_{\bullet \dots i_k} \mid 1 \leq i_j \leq d_j, j = 2, \dots, k\}) \\ r_2(\boldsymbol{X}) &:= \dim(\operatorname{span}_{\mathbb{F}}\{X_{i_1 \bullet \dots i_k} \mid 1 \leq i_j \leq d_j, j = 1, \dots, k, j \neq 2\}) \\ &\vdots \\ r_k(\boldsymbol{X}) &:= \dim(\operatorname{span}_{\mathbb{F}}\{X_{i_1 \cdots \bullet} \mid 1 \leq i_j \leq d_j, j = 1, \dots, k-1\}) \end{aligned}$$

Definition 1.18. [9] The *multilinear rank* of $\boldsymbol{X} \in \mathbb{F}^{d_1 \times \cdots \times d_k}$, denoted $\operatorname{rank}_{\boxplus}(\boldsymbol{X})$, is the k-tuple $(r_1(\boldsymbol{X}), \dots, r_k(\boldsymbol{X}))$.

Example 1.19. Let's see an example of how the multilinear rank of a tensor is computed. Let $\boldsymbol{X} \in \mathbb{R}^{2 \times 2 \times 3}$ be given by

$$\left[\begin{array}{cc|cc|cc} 1 & 2 & 2 & 2 & 1 & 0 \\ 3 & 4 & 0 & 1 & 4 & 10 \end{array}\right]$$

Since $\boldsymbol{X}$ is a third order tensor, we should expect the multilinear rank to be a 3-tuple given by the row, column, and tube rank. By fixing two of the indices we can create three sets of fibres. The set of column fibres is determined by fixing $i_2 \in \{1,2\}, i_3 \in \{1,2,3\}$:

$$\left\{ \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ 4 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 4 \end{bmatrix}, \begin{bmatrix} 0 \\ 10 \end{bmatrix} \right\}$$

The set of row fibres is determined by fixing $i_1 \in \{1,2\}, i_3 \in \{1,2,3\}$:

$$\left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 3 \\ 4 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 4 \\ 10 \end{bmatrix} \right\}$$

The set of tube fibres is determined by fixing $i_1, i_2 \in \{1,2\}$:

$$\left\{ \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 0 \\ 4 \end{bmatrix}, \begin{bmatrix} 4 \\ 1 \\ 10 \end{bmatrix} \right\}$$

Then we compute the dimension of the spanning sets,

$$r_1(\boldsymbol{X}) = \dim \left(\operatorname{span}_{\mathbb{R}} \left\{ \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ 4 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 4 \end{bmatrix}, \begin{bmatrix} 0 \\ 10 \end{bmatrix} \right\} \right) = 2$$

$$r_2(\boldsymbol{X}) = \dim \left(\operatorname{span}_{\mathbb{R}} \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 3 \\ 4 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 4 \\ 10 \end{bmatrix} \right\} \right) = 2$$

$$r_3(\boldsymbol{X}) = \dim \left(\operatorname{span}_{\mathbb{R}} \left\{ \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 0 \\ 4 \end{bmatrix}, \begin{bmatrix} 4 \\ 1 \\ 10 \end{bmatrix} \right\} \right) = 3$$

We conclude that $\operatorname{rank}_{\boxplus}(\boldsymbol{X}) = (2,2,3)$.

Before we summarize some properties of multilinear rank, we introduce the notion of *outer product rank*.

Definition 1.20. [20] A tensor $\boldsymbol{X} \in \mathbb{F}^{d_1 \times \cdots \times d_k}$ is *decomposable* if it can be written as

$$\boldsymbol{X} = u^{(1)} \otimes u^{(2)} \otimes \cdots \otimes u^{(k)},$$

with non-zero $u^{(i)} \in \mathbb{F}^{d_i}$ for $i = 1, \ldots, k$. The $(i_1, i_2, \ldots, i_k)$th entry of $\boldsymbol{X}$ is

$$x_{i_1 i_2 \cdots i_k} = u_{i_1}^{(1)} u_{i_2}^{(2)} \cdots u_{i_k}^{(k)}.$$

A decomposable tensor is also called a *simple tensor* or a *rank-1 tensor*.

Definition 1.21. [20] A tensor has *outer product rank* r if it can be written as a sum of r - and no fewer - decomposable tensors,

$$\boldsymbol{X} = \sum_{i=1}^{r} = u_i^{(1)} \otimes \cdots \otimes u_i^{(k)} = u_1^{(1)} \otimes \cdots \otimes u_1^{(k)} + \cdots + u_r^{(1)} \otimes \cdots \otimes u_r^{(k)}$$

where $u_i^{(1)} \in \mathbb{F}^{d_1}, \ldots, u_i^{(k)} \in \mathbb{F}^{d_k}, i = 1, \ldots, r$. We write $\text{rank}_{\otimes}(\boldsymbol{X})$ to denote the outer product rank of $\boldsymbol{X}$.

The only rank-0 tensor is the zero tensor.

Definition 1.22. The *maximum rank* is defined to be the $\max\{\text{rank}_{\otimes}(\boldsymbol{X}) \mid \boldsymbol{X} \in \mathbb{F}^{d_1 \times \cdots \times d_k}\}$.

The maximum rank of an $I \times J$ matrix is the minimum of $\{I, J\}$. However, for arbitrary $\boldsymbol{X} \in \mathbb{F}^{d_1 \times \cdots \times d_k}$, $k \geq 3$ the maximum rank is unknown and always exceeds the minimum of $\{d_1, \ldots, d_k\}$. In most cases there are weak upper bounds [22]. Of particular interest to us is the fact that the rank of any $2 \times 2 \times 2$ tensor is at most three [36]. Kruskal proved this fact, and in the next chapter we will provide our own proof. We summarize the discussion from page 1094 of [9] as a proposition.

Proposition 1.23. *[9] The following inequality relates the multilinear rank with the outer product rank and the dimensions of any $d_1 \times \cdots \times d_k$ tensor:*

$$r_p(\mathbf{X}) \leq \min\{\text{rank}_{\otimes}(\mathbf{X}), d_p\}$$

where $p = 1, \ldots, k$.

Proof. Suppose $\text{rank}_{\otimes}(\boldsymbol{X}) = r$. Then we can write

$$\boldsymbol{X} = \sum_{i=1}^{r} u_i^{(1)} \otimes \cdots \otimes u_i^{(k)} = u_1^{(1)} \otimes \cdots \otimes u_1^{(k)} + \cdots + u_r^{(1)} \otimes \cdots \otimes u_r^{(k)}$$

For $p = 1$, each vector $\boldsymbol{X}_{\bullet i_2 \dots i_k} = [x_{i_1 \dots i_k}]_{i_1=1}^{d_1} \in \text{span}(u_1^{(1)}, \dots, u_r^{(1)})$. Then $r_1(\boldsymbol{X})$ is at most as large as the number of terms in the decomposition: $r_1(\boldsymbol{X}) \leq r$. Also, the dimension of the span of the vectors $\boldsymbol{X}_{\bullet i_2 \dots i_k}$ cannot be larger than d_1 since $\boldsymbol{X}_{\bullet i_2 \dots i_k} \in \mathbb{R}^{d_1}$. That is, $r_1(\boldsymbol{X}) \leq d_1$. A similar argument for $p = 2, \dots, k$ proves the statement.

□

Consequently, we have the inequality:

$$\max\{r_p(\boldsymbol{X}) \mid p = 1, \dots, k\} \leq \text{rank}_{\otimes}(\boldsymbol{X})$$

Let us focus our attention to the outer product rank. When not otherwise specified, *rank* will henceforth refer to outer product rank. In 1927 Hitchcock [20] proposed the definition of the rank of a tensor $\boldsymbol{X}$ as we defined it above: the smallest number of rank-1 tensors that generate $\boldsymbol{X}$ as their sum. The primary difference between matrix rank and tensor rank is that there is no known algorithm to determine the rank of a given tensor [22]. In fact, the problem is NP-hard [22] [18]. Furthermore, tensor rank is field dependent. That is, the rank of a real-valued tensor is in general different over $\mathbb{R}$ and $\mathbb{C}$. This was proved by Kruskal [24], and we will see an example of this later.

Recall that the set of invertible $n \times n$ matrices forms a group under matrix multiplication, denoted by $GL_n(\mathbb{F})$. Fix a basis on $V_1 \otimes \cdots \otimes V_k$ so that we may regard an order-k tensor as a multidimensional array, and consider the following definition.

Definition 1.24. The group $GL_{d_1,\dots,d_k}(\mathbb{F}) := GL_{d_1}(\mathbb{F}) \times \cdots \times GL_{d_k}(\mathbb{F})$ acts on an order-k tensor by changes of basis along the k modes.

Lemma 1.25. *Lemma 2.3, page 1092 of [9]*
If $\mathbf{X} \in \mathbb{R}^{d_1 \times \cdots \times d_k}$ *and* $(A_1, \dots, A_k) \in \mathbb{R}^{c_1 \times d_1} \times \cdots \times \mathbb{R}^{c_k \times d_k}$, *then*

$$\text{rank}_{\otimes}((A_1, \dots, A_k) \cdot \mathbf{X}) \leq \text{rank}_{\otimes}(\mathbf{X}),$$

with equality when $(A_1, \ldots, A_k) \in \mathrm{GL}_{d_1,\ldots,d_k}(\mathbb{R})$.

Proof. Suppose some tensor X has rank r. Consider the decomposition of $\boldsymbol{X}$ as a sum of rank-1 tensors: $\boldsymbol{X} = \boldsymbol{X}_1 + \cdots + \boldsymbol{X}_r$. It is easy to see that multilinear matrix multiplication of decomposable tensors obeys

$$(A_1, \ldots, A_k) \cdot \boldsymbol{X}_i = (A_1, \ldots, A_k) \cdot (x_i^{(1)} \otimes \cdots \otimes x_i^{(k)}) = A_1 x_i^{(1)} \otimes \cdots \otimes A_k x_i^{(k)}.$$

That is, for any decomposable (rank-1) tensor, the action of invertible matrices results in a decomposable tensor (none of the factors in the tensor product are 0). By linearity,

$$\begin{aligned}
&(A_1, \ldots, A_k) \cdot (\boldsymbol{X}_1 + \cdots + \boldsymbol{X}_r) \\
&= (A_1, \ldots, A_k) \cdot \boldsymbol{X}_1 + \cdots + (A_1, \ldots, A_k) \cdot \boldsymbol{X}_r \\
&= (A_1, \ldots, A_k) \cdot (x_1^{(1)} \otimes \cdots \otimes x_1^{(k)}) + \cdots + (A_1, \ldots, A_k) \cdot (x_r^{(1)} \otimes \cdots \otimes x_r^{(k)}) \\
&= A_1 x_1^{(1)} \otimes \cdots \otimes A_k x_1^{(k)} + \cdots + A_1 x_r^{(1)} \otimes \cdots \otimes A_k x_r^{(k)}.
\end{aligned}$$

Then we have that

$$\mathrm{rank}_\otimes((A_1, \ldots, A_k) \cdot \boldsymbol{X}) \le \mathrm{rank}_\otimes(\boldsymbol{X}),$$

If the A_i are invertible then

$$\boldsymbol{X} = (A_1^{-1}, \ldots, A_k^{-1}) \cdot [(A_1, \ldots, A_k) \cdot \boldsymbol{X}],$$

and so $\mathrm{rank}_\otimes(\boldsymbol{X}) \le \mathrm{rank}_\otimes((A_1, \ldots, A_k) \cdot \boldsymbol{X})$ and hence $\mathrm{rank}_\otimes((A_1, \ldots, A_k) \cdot \boldsymbol{X}) = \mathrm{rank}_\otimes(\boldsymbol{X})$. □

More generally, we have the multilinear rank equivalents of the previous lemma. First we clarify that for two lists, $(r_1, \ldots, r_k)$ and $(s_1, \ldots, s_k)$, we say $(r_1, \ldots, r_k) \le (s_1, \ldots, s_k)$ if $r_1 \le s_1, \ldots, r_k \le s_k$.

Lemma 1.26. *Page 1094 of [9]*
If $\mathbf{X} \in \mathbb{F}^{d_1 \times \cdots \times d_k}$ *and* $(A_1, \ldots, A_k) \in \mathbb{F}^{c_1 \times d_1} \times \cdots \times \mathbb{F}^{c_k \times d_k}$, *then*

$$\mathrm{rank}_\boxplus((A_1, \ldots, A_k) \cdot \mathbf{X}) \le \mathrm{rank}_\boxplus(\mathbf{X}),$$

with equality when $(A_1, \ldots, A_k) \in \mathrm{GL}_{d_1,\ldots,d_k}(\mathbb{F})$.

Definition 1.27. [9] Let the base field be $\mathbb{R}$ or $\mathbb{C}$. The rank of the tensor format $d_1 \times \cdots \times d_k$ is called *typical* if it holds true on a set of positive volume.

This definition supposes that some topology (see Section 2.1.2 Border Rank) has been defined on $\mathbb{F}^{d_1 \times \cdots \times d_k}$ [8], and volume is integration in $\mathbb{R}^n$ with respect to the standard Euclidean metric [23].

Tensors in $\mathbb{R}^{2\times2\times2}$ have typical ranks of two and three. Monte Carlo experiments have demonstrated that the $2 \times 2 \times 2$ tensors with a rank of two fill approximately 79% of the space, while those with a rank of three fill about 21% [22]. Recent work by Bergqvist [1] has determined exact values of typical ranks for $n \times n \times 2$ tensors (with $n = 2, 3$), whose elements are independent random variables, normally distributed with mean 0 and variance 1. For $n = 2$, the probability of having rank 2 is $\frac{\pi}{4}$. Then, the probability of a $2 \times 2 \times 2$ tensor having rank 3 is $1 - \frac{\pi}{4}$. For $n = 3$, the probability of obtaining a tensor with typical rank 3 is $\frac{1}{2}$.

Tables 3.2, 3.3, 3.4 and 3.5 on pages 465 and 466 of [22] give a comparison of maximum ranks and typical ranks of general order-3 tensors.

Definition 1.28. Given an order-k tensor of format $d_1 \times \cdots \times d_k$ over $\mathbb{R}$ or $\mathbb{C}$, if there is a single typical rank then this rank is called the *generic* rank of that format.

This definition is field dependent, since there is a single generic rank if the underlying field is algebraically closed (as in $\mathbb{C}$), but there may be several typical ranks if the field is $\mathbb{R}$. So a generic rank is typical, but a typical rank need not be generic. See Comon et al. [35, 8] for tables summarizing the generic and typical ranks of various orders of tensors. We note that for $k = 2$, the generic rank of a matrix is always equal to its maximum rank.

Given dimensions $d_1, \ldots, d_k$, the determination of the existence of the generic rank is an open problem in general [9]. This problem does not occur for multilinear rank, however. The following proposition shows that a generic multilinear rank always exists (since computing the multilinear rank involves computing the rank of matrix slices, which always have a generic rank) and depends only on $d_1, \ldots, d_k$, not the base field.

Proposition 1.29. *Page 1119 of [9] The multilinear rank of a tensor is independent of the choice of base field. If $\bar{\mathbb{F}}$ is an extension field of $\mathbb{F}$, the value of rank$_{\boxplus}(\mathbf{X})$ is the same whether the tensor $\mathbf{X}$ is regarded as an element of $\mathbb{F}^{d_1 \times \cdots \times}$ or of $\bar{\mathbb{F}}^{d_1 \times \cdots \times}$*

Proof. Let $\operatorname{rank}_{\boxplus}(\boldsymbol{X}) = (r_1(\boldsymbol{X}), \ldots, r_k(\boldsymbol{X}))$. Let

$$f_i : \mathbb{F}^{d_1 \times \cdots \times d_k} \to \mathbb{F}^{d_i \times \prod_{j \neq i} d_j}$$

be the map that flattens a tensor into a matrix in the ith mode. Then $r_i(\boldsymbol{X}) = \operatorname{rank}(f_i(\boldsymbol{X}))$, where the right side denotes the matrix rank. Since matrix rank is independent of the base field, so is the multilinear rank and the result follows. □

Chapter 2

On $2 \times 2 \times 2$ Tensors

Recall that we represent order-3 tensors of size $I \times J \times K$ pictorially as K side-by-side $I \times J$ matrix slices. The general tensor $\boldsymbol{X} \in \mathbb{F}^{2\times2\times2}$ is represented in terms of its frontal slices,

$$\boldsymbol{X} = \left[\begin{array}{c|c} \boldsymbol{X}_1 & \boldsymbol{X}_2 \end{array} \right] = \left[\begin{array}{cc|cc} x_{111} & x_{121} & x_{112} & x_{122} \\ x_{211} & x_{221} & x_{212} & x_{222} \end{array} \right],$$

by fixing the last index. By fixing the second and first indices, we obtain the vertical slices:

$$\begin{bmatrix} x_{111} & x_{112} \\ x_{211} & x_{212} \end{bmatrix}, \qquad \begin{bmatrix} x_{121} & x_{122} \\ x_{221} & x_{222} \end{bmatrix}$$

and the horizontal slices

$$\begin{bmatrix} x_{111} & x_{112} \\ x_{121} & x_{122} \end{bmatrix}, \qquad \begin{bmatrix} x_{211} & x_{212} \\ x_{221} & x_{222} \end{bmatrix}$$

respectively.

2.1 Rank

2.1.1 Outer Product Rank

Kruskal proved in a rather complicated way that the outer product rank of a $2 \times 2 \times 2$ tensor is at most 3. J. ten Berge provided what he called a more straightforward constructive proof in [36]. Professor Murray Bremner summarizes ten Berge's paper in [3], which we use to discuss the proof of the maximum rank of $2 \times 2 \times 2$ tensors. Consider the outer product $a \otimes b \otimes c$ of three nonzero column vectors in $\mathbb{F}^2$. We obtain the $2 \times 2 \times 2$ rank-1 tensor whose (i, j, k) entry is $a_i b_j c_k$:

$$a \otimes b \otimes c = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \otimes \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \otimes \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \left[\begin{array}{cc|cc} a_1 b_1 c_1 & a_1 b_2 c_1 & a_1 b_1 c_2 & a_1 b_2 c_2 \\ a_2 b_1 c_1 & a_2 b_2 c_1 & a_2 b_1 c_2 & a_2 b_2 c_2 \end{array} \right]$$

Lemma 2.1. *Let* $\mathbf{X}$ *be a* $2 \times 2 \times 2$ *tensor. The rank of* $\mathbf{X}$ *is 0 if and only if every entry of* $\mathbf{X}$ *is zero. The rank of* $\mathbf{X}$ *is 1 if and only if* $\mathbf{X}$ *is the outer product of three nonzero vectors.*

Proof. Immediate. □

Lemma 2.2. *The rank of a nonzero* $2 \times 2 \times 2$ *tensor* $\mathbf{X}$ *is the smallest positive integer* r *such that the frontal slices can be written as*

$$X_1 = AC_1B^t, \qquad X_2 = AC_2B^t,$$

where A *and* B *are* $2 \times r$ *matrices and* C_1 *and* C_2 *are* $r \times r$ *diagonal matrices.*

Proof. Consider the decomposition of $\boldsymbol{X}$ as a sum of outer products of nonzero vectors:

$$\boldsymbol{X} = \sum_{i=1}^{r} a^{(i)} \otimes b^{(i)} \otimes c^{(i)}$$

where $a^{(i)}, b^{(i)}, c^{(i)}, i = 1, \ldots, r$ are column vectors of length 2. Consider three $2 \times r$ matrices A, B, C written in terms of column vectors:

$$\begin{aligned} A &= \left[\begin{array}{cccc} a^{(1)} & a^{(2)} & \ldots & a^{(r)} \end{array} \right] = [A_{ij}] \\ B &= \left[\begin{array}{cccc} b^{(1)} & b^{(2)} & \ldots & b^{(r)} \end{array} \right] = [B_{ij}] \\ C &= \left[\begin{array}{cccc} c^{(1)} & c^{(2)} & \ldots & c^{(r)} \end{array} \right] = [C_{ij}] \end{aligned}$$

Then the first frontal slice of the decomposition of $\boldsymbol{X}$ above is

$$\begin{aligned}
&\sum_{i=1}^{r} \begin{bmatrix} a_1^{(i)} b_1^{(i)} c_1^{(i)} & a_1^{(i)} b_2^{(i)} c_1^{(i)} \\ a_2^{(i)} b_1^{(i)} c_1^{(i)} & a_2^{(i)} b_2^{(i)} c_1^{(i)} \end{bmatrix} = \sum_{i=1}^{r} c_1^{(i)} \begin{bmatrix} a_1^{(i)} b_1^{(i)} & a_1^{(i)} b_2^{(i)} \\ a_2^{(i)} b_1^{(i)} & a_2^{(i)} b_2^{(i)} \end{bmatrix} \\
&= \sum_{i=1}^{r} c_1^{(i)} \begin{bmatrix} A_{1i} B_{1i} & A_{1i} B_{2i} \\ A_{2i} B_{1i} & A_{2i} B_{2i} \end{bmatrix} = \sum_{i=1}^{r} c_1^{(i)} \begin{bmatrix} A_{1i} B_{i1}^t & A_{1i} B_{i2}^t \\ A_{2i} B_{i1}^t & A_{2i} B_{i2}^t \end{bmatrix} \\
&= \sum_{i=1}^{r} \begin{bmatrix} A_{1i} c_1^{(i)} B_{i1}^t & A_{1i} c_1^{(i)} B_{i2}^t \\ A_{2i} c_1^{(i)} B_{i1}^t & A_{2i} c_1^{(i)} B_{i2}^t \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{r} A_{1i} c_1^{(i)} B_{i1}^t & \sum_{i=1}^{r} A_{1i} c_1^{(i)} B_{i2}^t \\ \sum_{i=1}^{r} A_{2i} c_1^{(i)} B_{i1}^t & \sum_{i=1}^{r} A_{2i} c_1^{(i)} B_{i2}^t \end{bmatrix} \\
&= A C_1 B^t
\end{aligned}$$

where C_1 is the $r \times r$ diagonal matrix whose diagonal entries are the entries

$$c_1^{(1)}, c_1^{(2)}, \ldots, c_1^{(r)}$$

in the first row of C. A similar calculation shows that the second frontal slice equals AC_2B^t where C_2 is the $r \times r$ diagonal matrix whose diagonal entries are the entries $c_2^{(1)}, c_2^{(2)}, \ldots, c_2^{(r)}$ in the second row of C:

$$C_1 = \begin{bmatrix} c_1^{(1)} & 0 & \ldots & 0 \\ 0 & c_1^{(2)} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & c_1^{(r)} \end{bmatrix}, \qquad C_2 = \begin{bmatrix} c_2^{(1)} & 0 & \ldots & 0 \\ 0 & c_2^{(2)} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & c_2^{(r)}. \end{bmatrix}$$

Conversely, if the two frontal slices can be written as AC_1B^t and AC_2B^t where A, B and C are the matrices defined above, then $\boldsymbol{X} = \sum_{i=1}^{r} a^{(i)} \otimes b^{(i)} \otimes c^{(i)}$. □

Definition 2.3. A $2 \times 2 \times 2$ tensor is called *superdiagonal* if it has one of the following forms for nonzero elements $\alpha, \beta \in \mathbb{F}$:

$$\left[\begin{array}{cc|cc} \alpha & 0 & 0 & 0 \\ 0 & 0 & 0 & \beta \end{array}\right], \quad \left[\begin{array}{cc|cc} 0 & \alpha & 0 & 0 \\ 0 & 0 & \beta & 0 \end{array}\right], \quad \left[\begin{array}{cc|cc} 0 & 0 & 0 & \beta \\ \alpha & 0 & 0 & 0 \end{array}\right], \quad \left[\begin{array}{cc|cc} 0 & 0 & \beta & 0 \\ 0 & \alpha & 0 & 0 \end{array}\right]$$

Lemma 2.4. *A superdiagonal* $2 \times 2 \times 2$ *tensor has rank 2.*

Proof. A superdiagonal tensor has rank at most 2, since

$$\left[\begin{array}{cc|cc} \alpha & 0 & 0 & 0 \\ 0 & 0 & 0 & \beta \end{array}\right] = \begin{bmatrix} \alpha \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ \beta \end{bmatrix} \otimes \begin{bmatrix} 0 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

The other superdiagonal tensors are decomposed similarly. We want to find the general form of a rank-1 tensor described in the previous lemma. To this end, set

$$A = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}, \qquad B = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}, \qquad C_1 = \begin{bmatrix} c_1 \end{bmatrix}, \qquad C_2 = \begin{bmatrix} c_2 \end{bmatrix}.$$

Then we have

$$X_1 = AC_1B^t = \begin{bmatrix} a_1c_1b_1 & a_1c_1b_2 \\ a_2c_1b_1 & a_2c_1b_2 \end{bmatrix} = c_1 \begin{bmatrix} a_1b_1 & a_1b_2 \\ a_2b_1 & a_2b_2 \end{bmatrix} = c_1(A \otimes B)$$
$$X_2 = AC_2B^t = \begin{bmatrix} a_1c_2b_1 & a_1c_2b_2 \\ a_2c_2b_1 & a_2c_2b_2 \end{bmatrix} = c_2 \begin{bmatrix} a_1b_1 & a_1b_2 \\ a_2b_1 & a_2b_2 \end{bmatrix} = c_2(A \otimes B),$$

where $A, B \in \mathbb{F}^2$. That is, X_1 and X_2 are scalar multiples of the same rank-1 matrix. Since any superdiagonal tensor does not satisfy this, it must have rank strictly greater than 1. Thus, we conclude a superdiagonal tensor has rank ≥ 2. □

Lemma 2.5. *The nonzero* $2 \times 2 \times 2$ *tensor* $\mathbf{X}$ *has rank one if and only if all six of its slices are singular and it is not a superdiagonal tensor.*

Proof. ($\Rightarrow$) Suppose that either $\boldsymbol{X}$ has a non-singular slice or it is a superdiagonal tensor. It remains to prove that if $\boldsymbol{X}$ has a non-singular slice then its rank is greater than one. By interchanging the appropriate slices, we may assume that the first frontal slice is non-singular. For $r \leq 1$, we have that $X_1 = AC_2B^t = c_1(a \otimes b)$, $a, b \in \mathbb{F}^2 - \{0\}$ is a matrix with rank at most one. But this means the tensor is singular (has at least one singular slice in any of the directions).

($\Leftarrow$) Assume that all six slices of $\boldsymbol{X}$ are singular and that $\boldsymbol{X}$ is not superdiagonal. Consider the following cases.

Case 1: Some slice is zero. We may assume that X_1 is the zero slice. Then since X_2 is singular as well, it must have a rank equal to one, and thus $\boldsymbol{X}$ has rank 1. We can write the singular, non-zero slice X_2 as the outer product of two non-zero vectors:

$$X_2 = a \otimes b \qquad \Longrightarrow \qquad \boldsymbol{X} = a \otimes b \otimes \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Case 2: No slice is zero. Then since X_1 is nonzero and singular, we have

$$X_1 = a \otimes b, \qquad a \neq 0, \qquad b = \begin{bmatrix} 1 \\ \lambda \end{bmatrix}$$

That is, a is the first column of X_1 and λa is the second column.

Subcase 2(a): $\lambda = 0$. Since the vertical slices are singular and nonzero we have either

$$(i) \qquad X_2 = a \otimes c, \qquad c = \begin{bmatrix} \mu \\ \nu \end{bmatrix}, \qquad \mu, \nu \neq 0$$

(that is, μa is the first column of X_2 and νa is the second column), or

$$(ii) \qquad X_2 = d \otimes c, \qquad d \neq 0, \qquad c = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

(that is, the zero vector is the first column and d is the second column).

In subcase (i), since the horizontal slices are nonzero and singular, the vectors b and c are linearly dependent, say $c = \rho b, \rho \in \mathbb{F}$, and hence $\boldsymbol{X}$ has rank 1, since

$$\boldsymbol{X} = a \otimes b \otimes \begin{bmatrix} 1 \\ \rho \end{bmatrix}.$$

In subcase (ii), since the horizontal slices are nonzero and singular, $\boldsymbol{X}$ must be a superdiagonal tensor, which we needn't consider.

Subcase 2(b): $\lambda \neq 0$. Just as in subcase 2(a)(i), since the vertical slices are singular and nonzero we have

$$X_2 = a \otimes c, \qquad c = \begin{bmatrix} \mu \\ \nu \end{bmatrix}, \qquad \mu, \nu \neq 0,$$

(that is, μa is the first column of X_2 and νa is the second column). As before, b and c are linearly independent vectors, and hence

$$X = a \otimes b \otimes \begin{bmatrix} 1 \\ \rho \end{bmatrix},$$

which proves $\boldsymbol{X}$ has rank 1. $\square$

Theorem 2.6. *The rank of a $2 \times 2 \times 2$ tensor X is at most 3.*

Proof. It remains to prove that a tensor $\boldsymbol{X}$, whose first frontal slice X_1 is non-singular, has rank ≤ 3. We write

$$X_1 = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix}, \qquad Y_2 = X_2 X_1^{-1} = \begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{bmatrix}.$$

Consider the following matrices:

$$A = \begin{bmatrix} 1 & 0 & y_{12} \\ 0 & 1 & y_{21} \end{bmatrix}, \qquad B = \begin{bmatrix} x_{11} & x_{21} & x_{11} + x_{21} \\ x_{12} & x_{22} & x_{12} + x_{22} \end{bmatrix} = X_1^t \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix},$$

$$C_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad C_2 = \begin{bmatrix} y_{11} - y_{12} & 0 & 0 \\ 0 & y_{22} - y_{21} & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

We then verify by direct calculation that

$$AC_1B^t = \begin{bmatrix} 1 & 0 & y_{12} \\ 0 & 1 & y_{21} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} X_1 = X_1,$$

$$AC_2B^t = \begin{bmatrix} 1 & 0 & y_{12} \\ 0 & 1 & y_{21} \end{bmatrix} \begin{bmatrix} y_{11} - y_{12} & 0 & 0 \\ 0 & y_{22} - y_{21} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} X_1 = Y_2 X_1 = X_2.$$

Then Lemma 2.2 completes the proof.

$\square$

We now provide our own proof below, which we believe is more simple since it only relies on basic Linear Algebra.

Theorem 2.7. *Let $\mathbf{X} \in V \otimes V \otimes V$ where $V = \mathbb{F}^2$ and the standard basis vectors are denoted e_1, e_2. Then the outer product rank of $\mathbf{X}$ is less than or equal to 3.*

Proof. Let $\boldsymbol{X} \in V \otimes V \otimes V$. Then we can write

$$\begin{aligned}\boldsymbol{X} &= \sum_{i,j,k=1}^{2} x_{ijk} e_i \otimes e_j \otimes e_k \\ &= (x_{111} e_1 \otimes e_1 \otimes e_1 + x_{112} e_1 \otimes e_1 \otimes e_2) + (x_{121} e_1 \otimes e_2 \otimes e_1 + x_{122} e_1 \otimes e_2 \otimes e_2) \\ &\quad + (x_{211} e_2 \otimes e_1 \otimes e_1 + x_{212} e_2 \otimes e_1 \otimes e_2) + (x_{221} e_2 \otimes e_2 \otimes e_1 + x_{222} e_2 \otimes e_2 \otimes e_2) \\ &= e_1 \otimes e_1 \otimes v_1 + e_1 \otimes e_2 \otimes v_2 + e_2 \otimes e_1 \otimes v_3 + e_2 \otimes e_2 \otimes v_4\end{aligned}$$

where the v_i for $1 \leq i \leq 4$ are given by

$$v_1 = x_{111} e_1 + x_{112} e_2, \qquad v_2 = x_{121} e_1 + x_{122} e_2,$$
$$v_3 = x_{211} e_1 + x_{212} e_2, \qquad v_4 = x_{221} e_1 + x_{222} e_2.$$

Clearly, the rank of $\boldsymbol{X}$ is less than or equal 4. Let us show that for any tensor $\boldsymbol{X}$ its rank is less than or equal to 3. We assume that $v_i \neq 0$ for $i = 1, 2, 3, 4$, otherwise the rank is less than or equal to 3. Consider the following two cases:

Case 1. If v_2, v_3 are linearly independent, then $v_4 = \lambda v_2 + \mu v_3$. Hence, we have

$$\begin{aligned}\boldsymbol{X} &= e_1 \otimes e_1 \otimes v_1 + e_1 \otimes e_2 \otimes v_2 + e_2 \otimes e_1 \otimes v_3 + e_2 \otimes e_2 \otimes (\lambda v_2 + \mu v_3) \\ &= e_1 \otimes e_1 \otimes v_1 + e_1 \otimes e_2 \otimes v_2 + e_2 \otimes e_1 \otimes v_3 + \lambda e_2 \otimes e_2 \otimes v_2 + \mu e_2 \otimes e_2 \otimes v_3 \\ &= e_1 \otimes e_1 \otimes v_1 + (e_1 + \lambda e_2) \otimes e_2 \otimes v_2 + e_2 \otimes (e_1 + \mu e_2) \otimes v_3.\end{aligned}$$

Therefore, the rank of $\boldsymbol{X}$ is less than or equal to 3, as required.

Case 2. If v_2, v_3 are dependent, then $v_3 = tv_2$, where t is a scalar ($v_2 \neq 0$, otherwise the sum is at most 3). Then we have

$$\boldsymbol{X} = e_1 \otimes e_1 \otimes v_1 + e_1 \otimes e_2 \otimes v_2 + te_2 \otimes e_1 \otimes v_2 + e_2 \otimes e_2 \otimes v_4.$$

Consider the following two subcases:

(i) let v_1, v_2 be dependent. Then $v_2 = qv_1$ where q is a scalar (if either vectors are zero, then again the sum is at most 3). Thus,

$$\begin{aligned}\boldsymbol{X} &= e_1 \otimes e_1 \otimes v_1 + qe_1 \otimes e_2 \otimes v_1 + qte_2 \otimes e_1 \otimes v_1 + e_2 \otimes e_2 \otimes v_4 \\ &= e_1 \otimes (e_1 + qe_2) \otimes v_1 + qte_2 \otimes e_1 \otimes v_1 + e_2 \otimes e_2 \otimes v_4.\end{aligned}$$

In this case, the rank of $\boldsymbol{X}$ is less than or equal to 3.

(ii) let v_1, v_2 be independent. Then $v_4 = \lambda v_1 + \mu v_2$. Therefore,

$$\boldsymbol{X} = e_1 \otimes e_1 \otimes v_1 + e_1 \otimes e_2 \otimes v_2 + te_2 \otimes e_1 \otimes v_2 + \lambda e_2 \otimes e_2 \otimes v_1 + \mu e_2 \otimes e_2 \otimes v_2.$$

Next we add and subtract the following terms: $qe_1 \otimes e_1 \otimes v_2$ and $se_2 \otimes e_2 \otimes v_2$ where q, s are scalars such that $q = \frac{t}{s+\mu}$. Hence, we have

$$\begin{aligned}\boldsymbol{X} = {} & e_1 \otimes e_1 \otimes v_1 - qe_1 \otimes e_1 \otimes v_2 + qe_1 \otimes e_1 \otimes v_2 + e_1 \otimes e_2 \otimes v_2 + te_2 \otimes e_1 \otimes v_2 \\ & + \lambda e_2 \otimes e_2 \otimes v_1 + \mu e_2 \otimes e_2 \otimes v_2 + se_2 \otimes e_2 \otimes v_2 - se_2 \otimes e_2 \otimes v_2 \\ & = e_1 \otimes e_1 \otimes (v_1 - qv_2) + e_2 \otimes e_2 \otimes (\lambda v_1 - sv_2) \\ & + (qe_1 \otimes e_1 \otimes v_2 + se_2 \otimes e_2 \otimes v_2 + e_1 \otimes e_2 \otimes v_2 + \mu e_2 \otimes e_2 \otimes v_2 + te_2 \otimes e_1 \otimes v_2).\end{aligned}$$

Let's simplify the expression in brackets:

$$\begin{aligned}& qe_1 \otimes e_1 \otimes v_2 + se_2 \otimes e_2 \otimes v_2 + e_1 \otimes e_2 \otimes v_2 + \mu e_2 \otimes e_2 \otimes v_2 + te_2 \otimes e_1 \otimes v_2 \\ & = e_1 \otimes (qe_1 + e_2) \otimes v_2 + e_2 \otimes ((s+\mu)e_2 + te_1) \otimes v_2 \\ & = e_1 \otimes (qe_1 + e_2) \otimes v_2 + (s+\mu)e_2 \otimes (\frac{t}{s+\mu}e_1 + e_2) \otimes v_2 \\ & = e_1 \otimes (qe_1 + e_2) \otimes v_2 + (s+\mu)e_2 \otimes (qe_1 + e_2) \otimes v_2 \\ & = (e_1 + (s+\mu)e_2) \otimes (qe_1 + e_2) \otimes v_2.\end{aligned}$$

Then we obtain the following simplified result:

$$
\begin{aligned}
\boldsymbol{X} &= e_1 \otimes e_1 \otimes v_1 - qe_1 \otimes e_1 \otimes v_2 + qe_1 \otimes e_1 \otimes v_2 + e_1 \otimes e_2 \otimes v_2 + te_2 \otimes e_1 \otimes v_2 \\
&\quad + \lambda e_2 \otimes e_2 \otimes v_1 + \mu e_2 \otimes e_2 \otimes v_2 + se_2 \otimes e_2 \otimes v_2 - se_2 \otimes e_2 \otimes v_2 \\
&= e_1 \otimes e_1 \otimes (v_1 - qv_2) + e_2 \otimes e_2 \otimes (\lambda v_1 - sv_2) + (e_1 + (s+\mu)e_2) \otimes (qe_1 + e_2) \otimes v_2
\end{aligned}
$$

This implies the rank is once again smaller than or equal to 3, completing the proof.

□

The following corollary gives an upper bound on the rank of order-k tensors of format $2 \times \cdots \times 2$.

Corollary 2.8. *Let* $\mathbf{X} \in V^{\otimes n}$, $n \geq 3$, *where* $V = \mathbb{F}^2$ *and* $\{e_1, e_2\}$ *are the standard basis vectors. Then* $\mathrm{rank}_{\otimes}(\mathbf{X}) \leq 3 \cdot 2^{n-3}$, $n \geq 3$.

Proof. We will prove this by induction on n, the order of the tensor. If $n = 4$, write $\boldsymbol{X} = \left[\begin{array}{c|c} \boldsymbol{X}_1 & \boldsymbol{X}_2 \end{array}\right]$, where $\boldsymbol{X}_1, \boldsymbol{X}_2 \in V \otimes V \otimes V$. Then both of them have rank smaller than or equal to 3, and hence we can write

$$
\begin{aligned}
\boldsymbol{X}_1 &= a_1 \otimes b_1 \otimes c_1 + a_2 \otimes b_2 \otimes c_2 + a_3 \otimes b_3 \otimes c_3 \\
\boldsymbol{X}_2 &= a_1' \otimes b_1' \otimes c_1' + a_2' \otimes b_2' \otimes c_2' + a_3' \otimes b_3' \otimes c_3'
\end{aligned}
$$

We can lift $\boldsymbol{X}_1, \boldsymbol{X}_2$ to new tensors $\left[\begin{array}{c|c} \boldsymbol{X}_1 & \boldsymbol{O} \end{array}\right]$ and $\left[\begin{array}{c|c} \boldsymbol{O} & \boldsymbol{X}_2 \end{array}\right]$, where $\boldsymbol{O}$ is the zero $2 \times 2 \times 2$ tensor, and

$$\left[\begin{array}{c|c} \boldsymbol{X}_1 & \boldsymbol{O} \end{array}\right] = a_1 \otimes b_1 \otimes c_1 \otimes e_1 + a_2 \otimes b_2 \otimes c_2 \otimes e_1 + a_3 \otimes b_3 \otimes c_3 \otimes e_1$$

$$\left[\begin{array}{c|c} \boldsymbol{O} & \boldsymbol{X}_2 \end{array}\right] = a_1' \otimes b_1' \otimes c_1' \otimes e_2 + a_2' \otimes b_2' \otimes c_2' \otimes e_2 + a_3' \otimes b_3' \otimes c_3' \otimes e_2.$$

Since $\boldsymbol{X} = \left[\begin{array}{c|c} \boldsymbol{X}_1 & \boldsymbol{O} \end{array}\right] + \left[\begin{array}{c|c} \boldsymbol{O} & \boldsymbol{X}_2 \end{array}\right]$, we have a decomposition of $\boldsymbol{X}$ into a sum of $3 \cdot 2^{4-3} = 6$ terms.

Suppose when $n = k$, $\operatorname{rank}_{\otimes}(\boldsymbol{X}) \leq 3 \cdot 2^{k-3} =: r$. Then we can write $\boldsymbol{X}_1, \boldsymbol{X}_2$ as a sum of r terms with each term containing an outer product of k vectors:

$$\boldsymbol{X}_1 = a_1 \otimes \cdots \otimes z_1 + \cdots + a_r \otimes \cdots \otimes z_r$$
$$\boldsymbol{X}_2 = a_1' \otimes \cdots \otimes z_1' + \cdots + a_r' \otimes \cdots \otimes z_r'$$

Then when $n = k+1$, we can write $\boldsymbol{X} = \left[\begin{array}{c|c} \boldsymbol{X}_1 & \boldsymbol{X}_2 \end{array}\right]$, where $\boldsymbol{X}_1, \boldsymbol{X}_2 \in V^{\otimes k}$ are given as above. Once again, we lift $\boldsymbol{X}_1, \boldsymbol{X}_2$ to new tensors $\left[\begin{array}{c|c} \boldsymbol{X}_1 & \boldsymbol{O} \end{array}\right]$ and $\left[\begin{array}{c|c} \boldsymbol{O} & \boldsymbol{X}_2 \end{array}\right]$, where $\boldsymbol{O}$ is the order-k $2 \times \cdots \times 2$ tensor of zeros, and

$$\left[\begin{array}{c|c} \boldsymbol{X}_1 & \boldsymbol{O} \end{array}\right] = a_1 \otimes \cdots \otimes z_1 \otimes e_1 + \cdots + a_r \otimes \cdots \otimes z_r \otimes e_1$$

$$\left[\begin{array}{c|c} \boldsymbol{O} & \boldsymbol{X}_2 \end{array}\right] = a_1' \otimes \cdots \otimes z_1' \otimes e_2 + \cdots + a_r' \otimes \cdots \otimes z_r' \otimes e_2.$$

Since $\boldsymbol{X} = \left[\begin{array}{c|c} \boldsymbol{X}_1 & \boldsymbol{O} \end{array}\right] + \left[\begin{array}{c|c} \boldsymbol{O} & \boldsymbol{X}_2 \end{array}\right]$, we have a decomposition of $\boldsymbol{X}$ into

$$3 \cdot 2^{k-3} + 3 \cdot 2^{k-3} = 3 \cdot 2^{(k+1)-3}$$

terms, implying $\operatorname{rank}_{\otimes}(\boldsymbol{X}) \leq 3 \cdot 2^{(k+1)-3}$. $\square$

When we look at the canonical forms of $2 \times 2 \times 2 \times 2$ tensors, we will see that the maximum rank over $\mathbb{F}_2$ is 6, and the maximum rank over $\mathbb{F}_3$ is 5. This suggests the maximum rank over $\mathbb{R}$ and $\mathbb{C}$ is probably 5. In Table 3 of [35], Comon et al. show that the smallest typical rank is 4.

Definition 2.9. An order-k tensor $[x_{i_1 \dots i_k}] \in \mathbb{F}^{n \times \dots \times n}$ is called *symmetric* if

$$x_{i_{\sigma(1)} \dots i_{\sigma(k)}} = x_{i_1 \dots i_k}, \qquad i_1, \dots, i_k \in \{1, \dots, n\},$$

for all permutations $\sigma \in S_k$.

For example, the third-order tensor $[x_{ijk}] \in \mathbb{F}^{n \times n \times n}$ is symmetric if

$$x_{ijk} = x_{ikj} = x_{jik} = x_{jki} = x_{kij} = x_{kji}$$

for all $i, j, k \in \{1, \dots, n\}$. Comon et al. [8] give a table that includes the generic and typical ranks of symmetric tensors over fields with characteristic zero. The smallest typical rank is 3 (Table 6 of [35]).

2.1.2 Border Rank

The concept of border rank is used to define weak solutions to the approximation of a tensor's rank-r decomposition. We begin by discussing the topology of tensor rank. The discussion is from [9].

Definition 2.10. [23] Let V be a vector space. A *norm* on V is a function $|| \cdot || : V \to [0, \infty)$ satisfying:

$$\begin{aligned} (i) &\quad ||v|| = 0 \text{ if and only if } v = 0, \\ (ii) &\quad ||\alpha v|| = |\alpha| \cdot ||v|| \;\; \forall\, \alpha \in \mathbb{F}, v \in V, \\ (iii) &\quad ||v + w|| \leq ||v|| + ||w|| \;\; \forall\, v, w \in V. \end{aligned}$$

Definition 2.11. Let $\boldsymbol{X} = [x_{i_1 \dots i_k}] \in \mathbb{F}^{d_1 \times \dots \times d_k}$. The *Frobenius norm* of $\boldsymbol{X}$, denoted $||\boldsymbol{X}||_F$, and its associated inner product are defined by

$$||\boldsymbol{X}||_F^2 := \sum_{i_1,\dots,i_k=1}^{d_1,\dots,d_k} |x_{i_1 \dots i_k}|^2, \qquad \langle \boldsymbol{X}, \boldsymbol{Y} \rangle_F := \sum_{i_1,\dots,i_k=1}^{d_1,\dots,d_k} x_{i_1 \dots i_k} y_{i_1 \dots i_k}.$$

Definition 2.12. The l^2-norm (sometimes called Euclidean Norm) of a vector $x = \begin{bmatrix} x_1 & \dots & x_k \end{bmatrix}^t \in \mathbb{F}^k$, denoted $||x||_2$, is defined by

$$||x||_2 = \sqrt{\sum_{i=1}^{k} |x_i|^2} \qquad \text{for } \mathbb{F} = \mathbb{C},$$

where $|x_i|$ denotes the complex modulus, and is defined by

$$||x||_2 = \sqrt{\sum_{i=1}^{k} x_i^2} \qquad \text{for } \mathbb{F} = \mathbb{R},$$

For a decomposable tensor, the Frobenius norm satisfies

$$||u_1 \otimes u_2 \otimes \dots \otimes u_k||_F = ||u_1||_2 \; ||u_2||_2 \; \cdots \; ||u_k||_2$$

where $|| \cdot ||_2$ denotes the l^2-norm of a vector. For arbitrary tensors $\boldsymbol{X}$ and $\boldsymbol{Y}$,

$$||\boldsymbol{X} \otimes \boldsymbol{Y}||_F = ||\boldsymbol{X}||_F ||\boldsymbol{Y}||_F.$$

Consider the Orthogonal Group $O_n(\mathbb{F}) = \{M \in GL_n(\mathbb{F}) \mid M^t M = M M^t = I\}$, a subgroup of the General Linear group whose transformations preserve the Euclidean inner product. For $(A_1, \dots, A_k) \in O_{d_1,\dots,d_k}(\mathbb{F})$, we have orthogonal invariance:

$$||(A_1, \dots, A_k) \cdot \boldsymbol{X}||_F = ||\boldsymbol{X}||_F$$

for any tensor $\boldsymbol{X}$.

Definition 2.13. Let P be a set and τ a collection of subsets of P. Then τ is called a *topology* on P if:

(i) P and the empty set are elements of τ,

(ii) any union of elements of τ is an element of τ,

(iii) a finite intersection of elements of τ is an element of τ.

There are other norms on tensor spaces but since $\mathbb{F}^{d_1\times\cdots\times d_k}$ is finite dimensional, every norm induces the same topology:

Theorem 2.14. *[23] All norms on finite-dimensional vector spaces over $\mathbb{R}$ or $\mathbb{C}$ are equivalent.*

We define the following topological subspaces of $\mathbb{F}^{d_1\times\cdots\times d_k}$:

$$\mathcal{S}_r(d_1,\dots,d_k) = \{\boldsymbol{X} \in \mathbb{F}^{d_1\times\cdots\times d_k} \mid \operatorname{rank}_\otimes(\boldsymbol{X}) \leq r\},$$
$$\bar{\mathcal{S}}_r(d_1,\dots,d_k) = \text{closure of } \mathcal{S}_r(d_1,\dots,d_k) \subset \mathbb{F}^{d_1\times\cdots\times d_k}$$

We need $\bar{\mathcal{S}}_r$ because $\mathcal{S}_r$ is usually not closed.

Definition 2.15. An order-k tensor $\boldsymbol{X} \in \mathbb{F}^{d_1\times\cdots\times d_k}$ has *border rank* r, denoted $\underline{\operatorname{rank}}_\otimes(\boldsymbol{X})$, if $\boldsymbol{X} \in \bar{\mathcal{S}}_r(d_1,\dots,d_k)$ and $\boldsymbol{X} \notin \bar{\mathcal{S}}_{r-1}(d_1,\dots,d_k)$.

Remark 2.16. We have that $\underline{\operatorname{rank}}_\otimes(\boldsymbol{X}) \leq \operatorname{rank}_\otimes(\boldsymbol{X})$ for all tensors $\boldsymbol{X}$ (see [9] for a proof of this fact).

The purpose of border rank goes back to our discussion of finding an exact rank-r decomposition of a tensor. In particular, the problem of finding a best rank-r approximation of an order-k tensor, for $k \geq 3$, $r = 2,\dots,\min\{d_1,\dots,d_k\}$ has no solution in general. See Theorem 4.10 in [9] for a proof of this fact. Moreover, the set of tensors that fail to have a best low-rank approximation has positive volume [9]. We do not summarize how limit points are characterized because the discussion is technical, so we refer you to [9]. We will close this section with an example of a limit of rank-2 tensors that converges to a rank-3 tensor (see Theorem 1.1 in [9]). For $d_1, d_2, d_3 \geq 2$, let

$$\boldsymbol{X}_n = n\left(x_1 + \frac{1}{n}y_1\right) \otimes \left(x_2 + \frac{1}{n}y_2\right) \otimes \left(x_3 + \frac{1}{n}y_3\right) - nx_1 \otimes x_2 \otimes x_3$$

be a sequence of tensors in $\mathbb{R}^{d_1 \times d_2 \times d_3}$ with $\operatorname{rank}_{\otimes}(\boldsymbol{X}_n) \leq 2$. The sequence converges to a rank-3 tensor:

$$\lim_{n\to\infty} \boldsymbol{X}_n = x_1 \otimes x_2 \otimes y_3 + x_1 \otimes y_2 \otimes x_3 + y_1 \otimes x_2 \otimes x_3 =: \boldsymbol{X}$$

In particular, $\boldsymbol{X}$ is an example of a tensor that has no best rank-2 approximation.

2.2 Cayley's Hyperdeterminant

Analogous to the determinant of a matrix we have the *hyperdeterminant* of a higher-order tensor. For $2 \times 2 \times 2$ tensors, this hyperdeterminant is known as Cayley's hyperdeterminant (or Kruskal's polynomial), discovered in 1845 [5]. This quartic polynomial, which we will denote by Δ, governs the structure of a $2\times2\times2$ tensor: the sign of this polynomial gives us insight into the rank of a $2\times2\times2$ tensor. We will derive Δ by applying transformations from $\mathrm{GL}_{2,2,2}(\mathbb{R})$. Recall from the previous chapter that these transformations are rank-invariant, and act by simultaneous changes of basis along the three directions.

By applying slice operations to $\boldsymbol{X} = \left[\begin{array}{c|c} \boldsymbol{X}_1 & \boldsymbol{X}_2 \end{array}\right]$ we can create new 2×2 slices of the form $\boldsymbol{X}^* = \beta_1 \boldsymbol{X}_1 + \beta_2 \boldsymbol{X}_2$. Then

$$\det(\boldsymbol{X^*}) = \det(\beta_1 \boldsymbol{X}_1 + \beta_2 \boldsymbol{X}_2)$$

$$= \det\left(\begin{bmatrix} \beta_1 x_{111} & \beta_1 x_{121} \\ \beta_1 x_{211} & \beta_1 x_{221} \end{bmatrix} + \begin{bmatrix} \beta_2 x_{112} & \beta_2 x_{122} \\ \beta_2 x_{212} & \beta_2 x_{222} \end{bmatrix}\right)$$

$$= \det\left(\begin{bmatrix} \beta_1 x_{111} + \beta_2 x_{112} & \beta_1 x_{121} + \beta_2 x_{122} \\ \beta_1 x_{211} + \beta_2 x_{212} & \beta_1 x_{221} + \beta_2 x_{222} \end{bmatrix}\right)$$

$$\begin{aligned} &= (\beta_1 x_{111} + \beta_2 x_{112})(\beta_1 x_{221} + \beta_2 x_{222}) - (\beta_1 x_{211} + \beta_2 x_{212})(\beta_1 x_{121} + \beta_2 x_{122}) \\ &= \beta_1^2 x_{111} x_{221} + \beta_1 \beta_2 x_{111} x_{222} + \beta_1 \beta_2 x_{112} x_{221} + \beta_2^2 x_{112} x_{222} \\ &\quad - (\beta_1^2 x_{211} x_{121} + \beta_1 \beta_2 x_{211} x_{122} + \beta_1 \beta_2 x_{212} x_{121} + \beta_2^2 x_{212} x_{122}) \end{aligned}$$

Collecting like terms of β_1, β_2 yields the following polynomial:

$$\det(\boldsymbol{X^*}) = \beta_1^2 \det(\boldsymbol{X}_1) + \beta_1 \beta_2 \frac{\det(\boldsymbol{X}_1 + \boldsymbol{X}_2) - \det(\boldsymbol{X}_1 - \boldsymbol{X}_2)}{2} + \beta_2^2 \det(\boldsymbol{X}_2)$$

Then the discriminant of this polynomial gives the hyperdeterminant, Δ.

$$\Delta(\boldsymbol{X}) = \Delta([\boldsymbol{X}_1 \mid \boldsymbol{X}_2]) = \left[\frac{\det(\boldsymbol{X}_1 + \boldsymbol{X}_2) - \det(\boldsymbol{X}_1 - \boldsymbol{X}_2)}{2}\right]^2 - 4\det(\boldsymbol{X}_1)\det(\boldsymbol{X}_2)$$

In terms of the eight entries of $\boldsymbol{X}$, we have

$$\begin{aligned} \Delta(\boldsymbol{X}) = &(x_{111}^2 x_{222}^2 + x_{112}^2 x_{221}^2 + x_{121}^2 x_{212}^2 + x_{122}^2 x_{211}^2) \\ &- 2(x_{111} x_{112} x_{221} x_{222} + x_{111} x_{121} x_{212} x_{222} + x_{111} x_{122} x_{211} x_{222} + x_{112} x_{121} x_{212} x_{221} \\ &+ x_{112} x_{122} x_{221} x_{211} + x_{121} x_{122} x_{212} x_{211}) + 4(x_{111} x_{122} x_{212} x_{221} + x_{112} x_{121} x_{211} x_{222}), \end{aligned}$$

a homogeneous polynomial of degree four.

Observe the following useful results.

Proposition 2.17. *Proposition 5.4, page 1107 of [9]*
If $\mathbf{X} \in \mathbb{R}^{d_1 \times \cdots \times d_k}$ *has rank 1, then* $\mathbf{X} = (A_1, \ldots, A_k) \cdot \mathbf{Y}$*, where* $(A_1, \ldots, A_k) \in \mathrm{GL}_{d_1,\ldots,d_k}(\mathbb{R})$ *and* $\mathbf{Y} = e_1 \otimes \cdots \otimes e_k$*.*

Proof. Write $\boldsymbol{X} = x_1 \otimes \cdots \otimes x_k$ and choose the A_i such that $A_i e_i = x_i$ for all i. □

Proposition 2.18. *Proposition 5.5, page 1107 of [9]*
Assume $d_i \geq 2$ *for all* i*. If* $\mathbf{X} \in \mathbb{R}^{d_1 \times \cdots \times d_k}$ *has rank 2, then* $\mathbf{X} = (A_1, \ldots, A_k) \cdot \mathbf{Y}$*, where* $(A_1, \ldots, A_k) \in \mathrm{GL}_{d_1,\ldots,d_k}(\mathbb{R})$ *and* $\mathbf{Y} \in \mathbb{R}^{2 \times \cdots \times 2}$ *is of the form* $\mathbf{Y} = e_1 \otimes \cdots \otimes e_1 + f_1 \otimes \cdots \otimes f_k$*. Here,* e_1 *is the standard basis vector, and each* f_i *is equal either to* e_1 *or* e_2*, and at least two of the* f_i *are equal to* e_2*.*

Proof. Since $\boldsymbol{X}$ has rank 2, we can write

$$\boldsymbol{X} = x_1 \otimes \cdots \otimes x_k + y_1 \otimes \cdots \otimes y_k$$

where each vector in the decomposition is non-zero. The vectors x_i and y_i must be linearly independent for at least two different indices i. Otherwise, suppose $y_i = \lambda_i x_i$ for $k-1$ of the indices, say, $i = 1, \ldots, k-1$. Then

$$\begin{aligned}
\boldsymbol{X} &= x_1 \otimes \cdots \otimes x_{k-1} \otimes x_k + \lambda_1 x_1 \otimes \cdots \otimes \lambda_{k-1} x_{k-1} \otimes y_k \\
&= x_1 \otimes \cdots \otimes x_{k-1} \otimes x_k + (\lambda_1 \cdots \lambda_{k-1}) x_1 \otimes \cdots \otimes x_{k-1} \otimes y_k \\
&= x_1 \otimes \cdots \otimes x_{k-1} \otimes x_k + x_1 \otimes \cdots \otimes x_{k-1} \otimes (\lambda_1 \cdots \lambda_{k-1}) y_k \\
&= x_1 \otimes \cdots \otimes x_{k-1} \otimes (x_k + (\lambda_1 \cdots \lambda_{k-1}) y_k)
\end{aligned}$$

This contradicts the assumption that the rank is 2.

For each i choose $A_i : \mathbb{R}^2 \to \mathbb{R}^{d_i}$ such that $A_i e_1 = x_i$ and such that $A_i e_2 = y_i$ if y_i is linearly independent of x_i (otherwise $A_i e_2$ may be arbitrary). It is easy to check that

$$(A_1, \ldots, A_k)^{-1} \cdot \boldsymbol{X} = e_1 \otimes \cdots \otimes e_1 + \lambda f_1 \otimes \cdots \otimes f_k,$$

where the f_i are as specified in the proposition, and λ is the product of the λ_i over the indices where $y_i = \lambda_i x_i$. We need to get rid of λ to get the correct form. We do this by replacing $A_i e_2 = y_i$ with $A_i e_2 = \lambda y_i$ at one of the indices i for which x_i, y_i are linearly independent. This completes the proof. $\square$

Proposition 2.19. *Proposition 5.6, page 1108 of [9]*
Let $\mathbf{X} \in \mathbb{R}^{2\times2\times2}$*, and let* $\mathbf{X}'$ *be obtained from* $\mathbf{X}$ *by permuting the three factors in the outer product, and let* $(A_1, A_2, A_3) \in \mathrm{GL}_{2,2,2}(\mathbb{R})$*. Then* $\Delta(\mathbf{X}') = \Delta(\mathbf{X})$ *and*

$$\Delta((A_1, A_2, A_3) \cdot \mathbf{X}) = \det(A_1)^2 \det(A_2)^2 \det(A_3)^2 \Delta(\mathbf{X}).$$

Proof. To prove that Δ is invariant under all permutations of the factors of $\mathbb{R}^{2\times2\times2}$, it suffices to check invariance in the cases of two distinct transpositions, since any two distinct transpositions generate the symmetric group on 3 elements. First let's show that Δ is invariant under the transposition of the first and second factors. Recall

$$\Delta\left(\left[\boldsymbol{X}_1 \mid \boldsymbol{X}_2\right]\right) = \left[\frac{\det(\boldsymbol{X}_1 + \boldsymbol{X}_2) - \det(\boldsymbol{X}_1 - \boldsymbol{X}_2)}{2}\right]^2 - 4\det(\boldsymbol{X}_1)\det(\boldsymbol{X}_2)$$

Then replacing $\boldsymbol{X}_1, \boldsymbol{X}_2$ with their transposes $\boldsymbol{X}_1^t, \boldsymbol{X}_2^t$ shows that Δ is invariant under the transposition of the first and second factors. Next, to show Δ is invariant under transposition of the second and third factors, write

$$\boldsymbol{X} = \left[\begin{array}{cc|cc} x_{11} & x_{12} & x_{21} & x_{22} \end{array}\right]$$

where the x_{ij} are column vectors. Then one can check that the equation

$$\begin{aligned}\Delta(\boldsymbol{X}) = {} & \det\begin{bmatrix} x_{11} & x_{22}\end{bmatrix}^2 + \det\begin{bmatrix} x_{21} & x_{12}\end{bmatrix}^2 \\ & - 2\det\begin{bmatrix} x_{11} & x_{12}\end{bmatrix}\det\begin{bmatrix} x_{21} & x_{22}\end{bmatrix} - 2\det\begin{bmatrix} x_{11} & x_{21}\end{bmatrix}\det\begin{bmatrix} x_{12} & x_{22}\end{bmatrix}\end{aligned}$$

has the desired symmetry.

Now we will consider the second claim. It is enough to verify using the case $(A_1, A_2, A_3) = (A_1, I, I)$. Then

$$(A_1, A_2, A_3) \cdot \boldsymbol{X} = \left[\begin{array}{c|c} A_1\boldsymbol{X}_1 & A_1\boldsymbol{X}_2 \end{array}\right]$$

and thus

$$\Delta\left(\left[A_1\boldsymbol{X}_1 \mid A_1\boldsymbol{X}_2\right]\right) = \left[\frac{\det(A_1\boldsymbol{X}_1 + A_1\boldsymbol{X}_2) - \det(A_1\boldsymbol{X}_1 - A_1\boldsymbol{X}_2)}{2}\right]^2$$
$$- 4\det(A_1\boldsymbol{X}_1)\det(A_1\boldsymbol{X}_2)$$
$$= \det(A_1)^2\left[\frac{\det(\boldsymbol{X}_1 + \boldsymbol{X}_2) - \det(\boldsymbol{X}_1 - \boldsymbol{X}_2)}{2}\right]^2$$
$$- 4\det(A_1)^2\det(\boldsymbol{X}_1)\det(\boldsymbol{X}_2)$$
$$= \det(A_1)^2\det(\boldsymbol{X}).$$

We see that a factor of $\det(A_1)^2$ appears, as required. □

Corollary 2.20. *Corollary 5.7, page 1108 of [9]*

The sign of Δ is invariant under the action of $GL_{2,2,2}(\mathbb{R})$.

The following propositions show the relationship between the hyperdeterminant and the rank of a $2 \times 2 \times 2$ tensor.

Proposition 2.21. *Proposition 5.9, page 1109 of [9]*

The base field is $\mathbb{R}$. If $\Delta(\mathbf{X}) > 0$ then $\text{rank}_{\otimes}(\mathbf{X}) \leq 2$.

Proof. If $\Delta(\boldsymbol{X}) > 0$ then the homogenous quadratic equation

$$\det(\boldsymbol{X^*}) = \beta_1^2\det(\boldsymbol{X}_1) + \beta_1\beta_2\frac{\det(\boldsymbol{X}_1 + \boldsymbol{X}_2) - \det(\boldsymbol{X}_1 - \boldsymbol{X}_2)}{2} + \beta_2^2\det(\boldsymbol{X}_2)$$

has two linearly independent root pairs (β_{11}, β_{12}) and (β_{21}, β_{22}). We can use slab operations to transform

$$\left[\begin{array}{c|c} \boldsymbol{X}_1 & \boldsymbol{X}_2 \end{array}\right] \to \left[\begin{array}{c|c} \boldsymbol{Y}_1 & \boldsymbol{Y}_2 \end{array}\right]$$

where $\boldsymbol{Y}_i = \beta_{i1}\boldsymbol{X}_1 + \beta_{i2}\boldsymbol{X}_2$. By construction $\det(\boldsymbol{Y}_i) = 0$, and so we can write $\boldsymbol{Y}_i = \mathbf{f_i} \otimes \mathbf{g_i}$ for some $\mathbf{f_i}, \mathbf{g_i} \in \mathbb{R}^{\mathbf{2}}$, possibly zero. It follows that

$$\left[\begin{array}{c|c} \boldsymbol{Y}_1 & \boldsymbol{Y}_2 \end{array}\right] = f_1 \otimes g_1 \otimes e_1 + f_2 \otimes g_2 \otimes e_2,$$

and thus $\text{rank}_{\otimes}(\boldsymbol{X}) = \text{rank}_{\otimes}\left(\left[\begin{array}{c|c} \boldsymbol{Y}_1 & \boldsymbol{Y}_2 \end{array}\right]\right) \leq 2.$ □

Proposition 2.22. *Proposition 5.10, page 1109 of [9]*
The base field is $\mathbb{R}$. *If* $rank_{\otimes}(\mathbf{X}) \leq 2$ *then* $\Delta(\mathbf{X}) \geq 0$.

Proof. If $\operatorname{rank}_{\otimes}(\boldsymbol{X}) \leq 1$ then either $\boldsymbol{X} = 0$ or $\boldsymbol{X} = (A_1, A_2, A_3) \cdot (e_1 \otimes e_1 \otimes e_1)$ and thus $\Delta(\boldsymbol{X}) = 0$. It remains to show that $\Delta(\boldsymbol{X})$ is non-negative when the rank equals 2. Proposition 2.18 implies that $\boldsymbol{X}$ can be transformed by an element of $\mathrm{GL}_{2,2,2}(\mathbb{R})$ into one of

$$T_1 = \left[\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right] \quad \text{or} \quad T_2 = \left[\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{array}\right]$$

Since $\Delta(T_1) = 1$ and $\Delta(T_2) = 0$, we have $\Delta(\boldsymbol{X}) \geq 0$. □

2.3 General Rank-r Tensor

Given a $2 \times 2 \times 2$ tensor $\boldsymbol{X} \in \mathbb{R}$ we can see that there is a relationship between the components of each vector in the outer product decomposition and the entries of the tensor. By considering a general tensor, Rovi [31] determined how the entries of the tensor effected its rank. In the next few sections we follow her approach.

2.4 General Rank-1 Tensor

We begin by looking at what makes a $2 \times 2 \times 2$ tensor rank-1. Let $\boldsymbol{X}$ be a rank-1 tensor with decomposition

$$\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \otimes \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \otimes \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \left[\begin{array}{cc|cc} x_{111} & x_{121} & x_{112} & x_{122} \\ x_{211} & x_{221} & x_{212} & x_{222} \end{array}\right]$$

We have the following relations between the entries of the tensor and the components of the vectors in the decomposition:

$$a_1b_1c_1 = x_{111}, \qquad a_1b_2c_1 = x_{121}, \qquad a_1b_1c_2 = x_{112}, \qquad a_1b_2c_2 = x_{122},$$
$$a_2b_1c_1 = x_{211}, \qquad a_2b_2c_1 = x_{221}, \qquad a_2b_1c_2 = x_{212}, \qquad a_2b_2c_2 = x_{222}.$$

There is at least one non-zero entry in this tensor, so we can interchange rows, columns and slices so that $x_{111} \neq 0$. That is, $a_1, b_1, c_1 \neq 0$. Then we can apply the transformations $\left(\frac{1}{a_1}\boldsymbol{X}_{1\bullet\bullet}\right)$, $\left(\frac{1}{b_1}\boldsymbol{X}_{\bullet1\bullet}\right)$ and $\left(\frac{1}{c_1}\boldsymbol{X}_{\bullet\bullet1}\right)$ so that $a_1, b_1, c_1 = 1$. Now we will consider cases depending on whether the entries are zero or non-zero. First consider when $a_2, b_2, c_2 \neq 0$. We can rearrange the equations in terms of a_2, b_2, and c_2:

$$a_2 = \frac{x_{211}}{b_1c_1} = \frac{x_{221}}{b_2c_1} = \frac{x_{212}}{b_1c_2} = \frac{x_{222}}{b_2c_2}, \qquad b_2 = \frac{x_{121}}{a_1c_1} = \frac{x_{221}}{a_2c_1} = \frac{x_{122}}{a_1c_2} = \frac{x_{222}}{a_2c_2},$$

$$c_2 = \frac{x_{112}}{a_1b_1} = \frac{x_{212}}{a_2b_1} = \frac{x_{122}}{a_1b_2} = \frac{x_{222}}{a_2c_2}.$$

Then a_2 is given by the ratio of the mode-1 entries, b_2 is given by the ratio of the mode-2 entries, and c_2 is given by the ratio of the mode-3 entries.

Example 2.23. Consider the tensor $\boldsymbol{X} \in \mathbb{R}^{2\times2\times2}$, given by

$$\boldsymbol{X} = \left[\begin{array}{cc|cc} 2 & 4 & -6 & -12 \\ 3 & 6 & -9 & -18 \end{array}\right]$$

In this example, $x_{111} = 2 \neq 0$ and so we do not have to rearrange any of the slices. We can scale the tensor to make this entry equal to 1:

$$\left[\begin{array}{cc|cc} 1 & 2 & -3 & -6 \\ \frac{3}{2} & 3 & -\frac{9}{2} & -9 \end{array}\right]$$

We see that the ratio between row one and row two is $\frac{3}{2}$; We also see that column two is twice column one, and slice two is negative three times slice one. Thus, the tensor is rank-1 and its decomposition is given by

$$\begin{bmatrix} 1 \\ \frac{3}{2} \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 2 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ -3 \end{bmatrix} = \left[\begin{array}{cc|cc} 1 & 2 & -3 & -6 \\ \frac{3}{2} & 3 & -\frac{9}{2} & -9 \end{array}\right]$$

Let's consider the case when one of a_2, b_2 or c_2 is zero. Suppose $b_2 = 0$:

$$\begin{bmatrix} 1 \\ a_2 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ c_2 \end{bmatrix} = \left[\begin{array}{cc|cc} x_{111} & 0 & x_{112} & 0 \\ x_{211} & 0 & x_{212} & 0 \end{array}\right]$$

We can see that the second vertical slice $\boldsymbol{X}_{\bullet 2 \bullet}$ has all entries equal to zero. Hence we can write the following equations

$$a_2 = x_{211}, \qquad a_2 c_2 = x_{212}, \qquad c_2 = x_{112},$$

with solutions given by

$$a_2 = \frac{x_{211}}{x_{111}} = \frac{x_{212}}{x_{112}}, \qquad c_2 = x_{112}.$$

Similarly, we can find the general form when the other entries of the vectors are zero. We conclude that by checking the ratios between the entries of each mode we can determine if a given tensor is rank-1.

2.5 General Rank-2 Tensor

Given a rank-2 tensor $\boldsymbol{X} \in \mathbb{R}^{2\times2\times2}$, we can write it as a sum of two decomposable tensors,

$$\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \otimes \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \otimes \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} + \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} \otimes \begin{bmatrix} e_1 \\ e_2 \end{bmatrix} \otimes \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} = \left[\begin{array}{cc|cc} x_{111} & x_{121} & x_{112} & x_{122} \\ x_{211} & x_{221} & x_{212} & x_{222} \end{array}\right]$$

Just as for the case of rank-1 tensors, we know there is at least one non-zero entry in each vector. The first mode corresponds to the vectors

$$\begin{bmatrix} a_1 \\ a_2 \end{bmatrix}, \qquad \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}.$$

We can apply an invertible linear transformation that sends these vectors to

$$\begin{bmatrix} 1 \\ a_2 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 \\ d_2 \end{bmatrix}.$$

After a similar computation in the second mode, we obtain the decomposition

$$\begin{bmatrix} 1 \\ a_2 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ b_2 \end{bmatrix} \otimes \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} + \begin{bmatrix} 1 \\ d_2 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ e_2 \end{bmatrix} \otimes \begin{bmatrix} f_1 \\ f_2 \end{bmatrix}$$

Then we have the equations,

$$c_1 + f_1 = x_{111}, \qquad c_2 + f_2 = x_{112}, \tag{2.1}$$

$$a_2c_1 + d_2f_1 = x_{211}, \qquad a_2c_2 + d_2f_2 = x_{212}, \tag{2.2}$$

$$b_2c_1 + e_2f_1 = x_{121}, \qquad b_2c_2 + e_2f_2 = x_{122}, \tag{2.3}$$

$$a_2b_2c_1 + d_2e_2f_1 = x_{221}, \qquad a_2b_2c_2 + d_2e_2f_2 = x_{222}. \tag{2.4}$$

We have to consider solving the equations under two different conditions: when the mode-i (for $i = 1, 2, 3$) vectors are linearly independent, and linearly dependent. That is, when the entries of the component vectors are different, and when they are equal. First we'll consider when they are not equal. In particular, let's look at the case when $a_2 \neq d_2$:

Rearranging (2.1) as $f_1 = x_{111} - c_1$ and $f_2 = x_{112} - c_2$, then substituting these into the other relations, we obtain a new system:

$$f_1 = x_{111} - c_1 \qquad\qquad f_2 = x_{112} - c_2 \tag{2.5}$$

$$a_2c_1 + d_2(x_{111} - c_1) = x_{211} \qquad a_2c_2 + d_2(x_{112} - c_2) = x_{212} \tag{2.6}$$

$$b_2c_1 + e_2(x_{111} - c_1) = x_{121} \qquad b_2c_2 + e_2(x_{112} - c_2) = x_{122} \tag{2.7}$$

$$a_2b_2c_1 + d_2e_2(x_{111} - c_1) = x_{221} \qquad a_2b_2c_2 + d_2e_2(x_{112} - c_2) = x_{222} \tag{2.8}$$

Rearranging (2.6) to get

$$c_1 = \frac{x_{211} - d_2x_{111}}{a_2 - d_2}, \qquad c_2 = \frac{x_{212} - d_2x_{112}}{a_2 - d_2},$$

then substituting these equations into the other relations gives:

$$f_1 = x_{111} - \left(\frac{x_{211} - d_2x_{111}}{a_2 - d_2}\right) \tag{2.9}$$

$$c_1 = \frac{x_{211} - d_2x_{111}}{a_2 - d_2} \tag{2.10}$$

$$b_2\left(\frac{x_{211} - d_2x_{111}}{a_2 - d_2}\right) + e_2\left(x_{111} - \left(\frac{x_{211} - d_2x_{111}}{a_2 - d_2}\right)\right) = x_{121} \tag{2.11}$$

$$a_2b_2\left(\frac{x_{211} - d_2x_{111}}{a_2 - d_2}\right) + d_2e_2\left(x_{111} - \left(\frac{x_{211} - d_2x_{111}}{a_2 - d_2}\right)\right) = x_{221} \tag{2.12}$$

$$f_2 = x_{112} - \left(\frac{x_{212} - d_2x_{112}}{a_2 - d_2}\right) \tag{2.13}$$

$$c_2 = \frac{x_{212} - d_2x_{112}}{a_2 - d_2} \tag{2.14}$$

$$b_2\left(\frac{x_{212} - d_2x_{112}}{a_2 - d_2}\right) + e_2\left(x_{112} - \left(\frac{x_{212} - d_2x_{112}}{a_2 - d_2}\right)\right) = x_{122} \tag{2.15}$$

$$a_2b_2\left(\frac{x_{212} - d_2x_{112}}{a_2 - d_2}\right) + d_2e_2\left(x_{112} - \left(\frac{x_{212} - d_2x_{112}}{a_2 - d_2}\right)\right) = x_{222} \tag{2.16}$$

To determine d_2 we will first write b_2 explicitly. To do this we'll multiply (2.11) by d_2 and subtract it from (2.12), then we'll multiply (2.15) by d_2 and subtract it from (2.16). This leaves us with:

$$b_2 = \frac{x_{221} - d_2x_{121}}{x_{211} - d_2x_{111}}, \qquad b_2 = \frac{x_{222} - d_2x_{122}}{x_{212} - d_2x_{112}}$$

Now we can put these two equations together to get a quadratic equation in the variable d_2:

$$(x_{121}x_{112} - x_{111}x_{122})d_2^2 + (x_{122}x_{211} + x_{111}x_{222} - x_{112}x_{221} - x_{121}x_{212})d_2 \\ +x_{221}x_{212} - x_{211}x_{222} = 0$$

This gives us two solutions:

$$d_2 = \frac{x_{112}x_{221} + x_{121}x_{212} - x_{122}x_{211} - x_{111}x_{222} \pm \sqrt{\Delta}}{2(x_{121}x_{112} - x_{111}x_{122})}$$

where the discriminant Δ is given by:

$$\begin{aligned}\Delta &= (x_{122}x_{211} + x_{111}x_{222} - x_{112}x_{221} - x_{121}x_{212})^2 \\ &\quad - 4(x_{121}x_{112} - x_{111}x_{122})(x_{221}x_{212} - x_{211}x_{222}) \\ &= (x_{111}^2x_{222}^2 + x_{112}^2x_{221}^2 + x_{121}^2x_{212}^2 + x_{122}^2x_{211}^2) \\ &\quad - 2(x_{111}x_{112}x_{221}x_{222} + x_{111}x_{121}x_{212}x_{222} + x_{111}x_{122}x_{211}x_{222} + x_{112}x_{121}x_{212}x_{221} \\ &\quad + x_{112}x_{122}x_{221}x_{211} + x_{121}x_{122}x_{212}x_{211}) + 4(x_{111}x_{122}x_{212}x_{221} + x_{112}x_{121}x_{211}x_{222}).\end{aligned}$$

Notice that Δ is exactly Cayley's hyperdeterminant of a $2 \times 2 \times 2$ tensor.

Next, we'll use the same idea to compute a_2. That is, we'll first compute two equations for e_2 then set these equations equal. To do this we'll multiply (2.11) by a_2 and subtract it from (2.12), then we'll multiply (2.15) by a_2 and subtract it from (2.16). This leaves us with:

$$e_2 = \frac{x_{221} - a_2x_{121}}{x_{211} - a_2x_{111}}, \qquad e_2 = \frac{x_{222} - a_2x_{122}}{x_{212} - d_2x_{112}}$$

Now we can put these two equations together to get a quadratic equation in the variable a_2. In fact the coefficients are the same as those in the quadratic equation above.

$$(x_{121}x_{112} - x_{111}x_{122})a_2^2 + (x_{122}x_{211} + x_{111}x_{222} - x_{112}x_{221} - x_{121}x_{212})a_2 \\ +x_{221}x_{212} - x_{211}x_{222} = 0$$

Then we will obtain the same two solutions as above:

$$a_2 = \frac{x_{112}x_{221} + x_{121}x_{212} - x_{122}x_{211} - x_{111}x_{222} \pm \sqrt{\Delta}}{2(x_{121}x_{112} - x_{111}x_{122})}.$$

Since a_2 and d_2 correspond to the same solutions, we can set a_2 to one solution, and d_2 to be the other. The order doesn't matter since this would correspond to a change in the position between both rank-1 components. This quadratic equation provides useful information about the tensor's rank [31]:

1. If the equation has distinct real roots, then the tensor will have rank-2.

2. If the roots are complex then we have computed the rank-2 decomposition over $\mathbb{C}$ of a real rank-3 tensor.

3. The equation cannot have equal roots since this would yield a contradiction because we have assumed that $a_2 \neq d_2$. The equation has equal roots when $\Delta = 0$. Below are the explicit equations for the entries of the vectors in the rank-2 decomposition:

$$a_2 = \frac{x_{112}x_{221} + x_{121}x_{212} - x_{122}x_{211} - x_{111}x_{222} + \sqrt{\Delta}}{2(x_{121}x_{112} - x_{111}x_{122})}$$

$$b_2 = \frac{x_{221} - d_2x_{121}}{x_{211} - d_2x_{111}} = \frac{x_{222} - d_2x_{122}}{x_{212} - d_2x_{112}}$$

$$c_1 = \frac{x_{211} - d_2x_{111}}{a_2 - d_2}$$

$$c_2 = \frac{x_{212} - d_2x_{112}}{a_2 - d_2}$$

$$d_2 = \frac{x_{112}x_{221} + x_{121}x_{212} - x_{122}x_{211} - x_{111}x_{222} - \sqrt{\Delta}}{2(x_{121}x_{112} - x_{111}x_{122})}$$

$$e_2 = \frac{x_{221} - a_2x_{121}}{x_{211} - a_2x_{111}} = \frac{x_{222} - a_2x_{122}}{x_{212} - a_2x_{112}}$$

$$f_1 = x_{111} - \left(\frac{x_{211} - d_2x_{111}}{a_2 - d_2}\right)$$

$$f_2 = x_{112} - \left(\frac{x_{212} - d_2x_{112}}{a_2 - d_2}\right)$$

Next we will consider what happens when some of the values of the vectors are equal. There are three cases to consider: $a_2 = d_2$, $b_2 = e_2$, and $\frac{c_1}{c_2} = \frac{f_1}{f_2}$. Let's expose only the first case, since the others can be solved similarly. If we set $a_2 = d_2$ the original relations reduce to:

$$\begin{aligned} c_1 + f_1 &= x_{111}, & c_2 + f_2 &= x_{112}, \\ d_2c_1 + d_2f_1 &= x_{211}, & d_2c_2 + d_2f_2 &= x_{212}, \\ b_2c_1 + e_2f_1 &= x_{121}, & b_2c_2 + e_2f_2 &= x_{122}, \\ d_2b_2c_1 + d_2e_2f_1 &= x_{221}, & d_2b_2c_2 + d_2e_2f_2 &= x_{222}. \end{aligned}$$

After factoring out d_2 we see that some equations are multiples of others:

$$\begin{aligned} c_1 + f_1 &= x_{111}, & c_2 + f_2 &= x_{112}, \\ d_2(c_1 + f_1) &= x_{211}, & d_2(c_2 + f_2) &= x_{212}, \\ b_2c_1 + e_2f_1 &= x_{121}, & b_2c_2 + e_2f_2 &= x_{122}, \\ d_2(b_2c_1 + e_2f_1) &= x_{221}, & d_2(b_2c_2 + e_2f_2) &= x_{222}. \end{aligned}$$

Eliminating the redundant relations, we're left with:

$$c_1 + f_1 = x_{111}, \quad b_2c_1 + e_2f_1 = x_{121}, \quad c_2 + f_2 = x_{112}, \quad b_2c_2 + e_2f_2 = x_{122}.$$

This corresponds to the following rank-2 decomposition:

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ b_2 \end{bmatrix} \otimes \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ e_2 \end{bmatrix} \otimes \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} = \begin{bmatrix} 1 \\ b_2 \end{bmatrix} \otimes \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} + \begin{bmatrix} 1 \\ e_2 \end{bmatrix} \otimes \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} = \begin{bmatrix} x_{111} & x_{112} \\ x_{121} & x_{122} \end{bmatrix}$$

Since there are four equations and six unknowns the decomposition is not unique when $a_2 = d_2$. Similarly, if we set $b_2 = e_2$ or $\frac{c_1}{c_2} = \frac{f_1}{f_2}$ and follow the same procedure as above we'll find that the decomposition is not unique. In terms of the tensor's slices, if $a_2 = d_2$ and the ratio between the mode-1 entries equals a_2, then the decomposition is not unique. Similarly, if $b_2 = e_2$ and the ratio between the mode-2 entries equals b_2, then once again the decomposition is not unique.

Example 2.24. Using the equations above, verify the decomposition of the following rank-2 tensor:

$$\begin{bmatrix}1\\4\end{bmatrix}\otimes\begin{bmatrix}1\\2\end{bmatrix}\otimes\begin{bmatrix}-1\\1\end{bmatrix}+\begin{bmatrix}1\\3\end{bmatrix}\otimes\begin{bmatrix}1\\-3\end{bmatrix}\otimes\begin{bmatrix}2\\-3\end{bmatrix}=\left[\begin{array}{cc|cc}1 & -8 & -2 & 11\\ 2 & -26 & -5 & 35\end{array}\right]$$

The strategy is to compute a_2 and d_2 first since they depend only on the entries of the tensor, then the other components can be determined easily. Using the equations we derived above, we substitute the entries of the tensor and find that $a_2 = 4$, $d_2 = 3$, and:

$$b_2 = \frac{x_{222} - d_2 x_{122}}{x_{212} - d_2 x_{112}} = \frac{35 - 3(11)}{-5 - 3(-2)} = 2$$

$$e_2 = \frac{x_{222} - a_2 x_{122}}{x_{212} - a_2 x_{112}} = \frac{35 - 4(11)}{-5 + 8} = -3$$

$$c_1 = \frac{x_{211} - d_2 x_{111}}{a_2 - d_2} = \frac{2 - 3(1)}{4 - 3} = -1$$

$$c_2 = \frac{x_{212} - d_2 x_{112}}{a_2 - d_2} = \frac{-5 - 3(-2)}{4 - 3} = 1$$

$$f_1 = x_{111} - \left(\frac{x_{211} - d_2 x_{111}}{a_2 - d_2}\right) = 1 - (-1) = 2$$

$$f_2 = x_{112} - \left(\frac{x_{212} - d_2 x_{112}}{a_2 - d_2}\right) - 2 - 1 = -3$$

2.6 General Rank-3 Tensor

Consider the general rank-3 decomposition of a $2 \times 2 \times 2$ tensor $\boldsymbol{X} \in \mathbb{R}^{2\times2\times2}$,

$$\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \otimes \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \otimes \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} + \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} \otimes \begin{bmatrix} e_1 \\ e_2 \end{bmatrix} \otimes \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} + \begin{bmatrix} g_1 \\ g_2 \end{bmatrix} \otimes \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} \otimes \begin{bmatrix} i_1 \\ i_2 \end{bmatrix} = \left[\begin{array}{cc|cc} x_{111} & x_{121} & x_{112} & x_{122} \\ x_{211} & x_{221} & x_{212} & x_{222} \end{array}\right].$$

We can apply an invertible linear transformation to the mode-1 vectors so that

$$\left\{ \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}, \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}, \begin{bmatrix} g_1 \\ g_2 \end{bmatrix} \right\} \mapsto \left\{ \begin{bmatrix} 1 \\ a_2 \end{bmatrix}, \begin{bmatrix} 1 \\ d_2 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}.$$

A similar transformation in the second mode allows us to write the rank-3 decomposition as

$$\begin{bmatrix} 1 \\ a_2 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ b_2 \end{bmatrix} \otimes \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} + \begin{bmatrix} 1 \\ d_2 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ e_2 \end{bmatrix} \otimes \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} 0 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} g \\ h \end{bmatrix} = \left[\begin{array}{cc|cc} x_{111} & x_{121} & x_{112} & x_{122} \\ x_{211} & x_{221} & x_{212} & x_{222} \end{array}\right].$$

We have the following relations between the entries of the tensor and the components of the vectors in the decomposition:

$$\begin{aligned} c_1 + f_1 &= x_{111}, & c_2 + f_2 &= x_{112}, \\ a_2c_1 + d_2f_1 &= x_{211}, & a_2c_2 + d_2f_2 &= x_{212}, \\ b_2c_1 + e_2f_1 &= x_{121}, & b_2c_2 + e_2f_2 &= x_{122}, \\ a_2b_2c_1 + d_2e_2f_1 + g &= x_{221}, & a_2b_2c_2 + d_2e_2f_2 &= x_{222}. \end{aligned}$$

Just as in the previous two cases, a series of substitutions gives the new system:

$$f_1 = x_{111} - c_1 \tag{2.17}$$

$$c_1 = \frac{x_{211} - d_2x_{111}}{a_2 - d_2} \tag{2.18}$$

$$b_2\left(\frac{x_{211} - d_2x_{111}}{a_2 - d_2}\right) + e_2\left(x_{111} - \left(\frac{x_{211} - d_2x_{111}}{a_2 - d_2}\right)\right) = x_{121} \tag{2.19}$$

$$a_2b_2\left(\frac{x_{211} - d_2x_{111}}{a_2 - d_2}\right) + d_2e_2\left(x_{111} - \left(\frac{x_{211} - d_2x_{111}}{a_2 - d_2}\right)\right) + i_1 = x_{221} \tag{2.20}$$

$$f_2 = x_{112} - c_2 \tag{2.21}$$

$$c_2 = \frac{x_{212} - d_2 x_{112}}{a_2 - d_2} \tag{2.22}$$

$$b_2 \left(\frac{x_{212} - d_2 x_{112}}{a_2 - d_2} \right) + e_2 \left(x_{112} - \left(\frac{x_{212} - d_2 x_{112}}{a_2 - d_2} \right) \right) = x_{122} \tag{2.23}$$

$$a_2 b_2 \left(\frac{x_{212} - d_2 x_{112}}{a_2 - d_2} \right) + d_2 e_2 \left(x_{112} - \left(\frac{x_{212} - d_2 x_{112}}{a_2 - d_2} \right) \right) + i_2 = x_{222} \tag{2.24}$$

Rearranging (2.19) and (2.23) so that

$$b_2(x_{211} - d_2 x_{111}) + e_2(a_2 x_{111} - x_{211}) = (a_2 - d_2) x_{121},$$

$$b_2(x_{212} - d_2 x_{112}) + e_2(a_2 x_{112} - x_{212}) = (a_2 - d_2) x_{122},$$

and setting $a_2 = 1$ and $d_2 = -1$ (we have two parameters since the number of variables still exceeds the number of equations by two) gives us

$$b_2(x_{211} + x_{111}) + e_2(x_{111} - x_{211}) = 2x_{121},$$

$$b_2(x_{212} + x_{112}) + e_2(x_{112} - x_{212}) = 2x_{122}.$$

Substituting one relation into the other gives

$$e_2 = \frac{x_{121}(x_{112} + x_{212}) - x_{122}(x_{111} + x_{211})}{x_{111} x_{212} - x_{211} x_{112}}.$$

We also obtain

$$b_2 = \frac{x_{111}^2 x_{122} - x_{211}^2 x_{122} + x_{111} x_{121} x_{212} - x_{211} x_{121} x_{112} - x_{111} x_{121} x_{112} + x_{211} x_{121} x_{212}}{(x_{111} + x_{211})(x_{111} x_{212} - x_{211} x_{112})}.$$

Then the remaining relations reduce to

$$c_1 = \frac{x_{211} + x_{111}}{2}, \quad c_2 = \frac{x_{212} + x_{112}}{2}, \quad f_1 = \frac{x_{111} - x_{211}}{2}, \quad f_2 = \frac{x_{112} - x_{212}}{2},$$

$$g = x_{221} - b_2 \left(\frac{x_{211} - d_2 x_{111}}{a_2 - d_2} \right) + e_2 \left(x_{111} - \left(\frac{x_{211} - d_2 x_{111}}{a_2 - d_2} \right) \right),$$

$$h = x_{222} - b_2 \left(\frac{x_{212} - d_2 x_{112}}{a_2 - d_2} \right) + e_2 \left(x_{112} - \left(\frac{x_{212} - d_2 x_{112}}{a_2 - d_2} \right) \right).$$

The equations above determine a non-unique solution to a rank-3 decomposition of a $2 \times 2 \times 2$ real tensor. That is, these results are particular to the case when $a_2 = 1$ and $d_2 = -1$, and so different equations will result from different choices for these entries.

Let's consider the tensor below to see how the rank of a real-valued tensor may be different over $\mathbb{R}$ and $\mathbb{C}$.

Example 2.25. Let $\boldsymbol{X} \in \mathbb{R}^{2\times2\times2}$ be given by

$$\boldsymbol{X} = \left[\begin{array}{cc|cc} 1 & 0 & 0 & 1 \\ 0 & 1 & -1 & 0 \end{array}\right].$$

The hyperdeterminant of this tensor is negative, $\Delta(\boldsymbol{X}) = -4$, and so we conclude that this tensor has rank-3 over $\mathbb{R}$. Substituting the entries of the tensor into

$$\begin{aligned}(x_{121}x_{112} - x_{111}x_{122})a_2^2 + (x_{122}x_{211} + x_{111}x_{222} - x_{112}x_{221} - x_{121}x_{212})a_2 \\ +x_{221}x_{212} - x_{211}x_{222} = 0,\end{aligned}$$

we see that $a_2^2 + 1 = 0$ with complex solutions $a_2 = \pm i$. Using the other equations we derived above we find that the decomposition over $\mathbb{C}$ is

$$\boldsymbol{X} = \begin{bmatrix} 1 \\ -i \end{bmatrix} \otimes \begin{bmatrix} 1 \\ i \end{bmatrix} \otimes \begin{bmatrix} \frac{1}{2} \\ \frac{i}{2} \end{bmatrix} + \begin{bmatrix} 1 \\ i \end{bmatrix} \otimes \begin{bmatrix} 1 \\ -i \end{bmatrix} \otimes \begin{bmatrix} \frac{1}{2} \\ -\frac{i}{2} \end{bmatrix}.$$

Over $\mathbb{R}$ this tensor has rank-3, and the decomposition is given by

$$\boldsymbol{X} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ -1 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} 0 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ -1 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

2.7 Canonical Forms

This section explains how to classify the canonical forms of $2 \times 2 \times 2$ tensors by studying their *equivalence* under multilinear matrix multiplication.

Theorem 2.26. *Let A be a nonzero $m \times n$ matrix. There exist invertible $m \times m$ and $n \times n$ matrices P and Q, such that the product PAQ is*

$$PAQ = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix},$$

where I_r is the $r \times r$ identity matrix, $r \leq m, n$.

The matrix PAQ is called the *Smith normal form* of the matrix A. This is a natural and familiar idea from basic linear algebra: for any 2×2 matrix, we can reduce it to exactly one of the following three distinct Smith normal forms

$$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

using left and right matrix multiplication. The canonical forms of $2 \times 2 \times 2$ arrays over $\mathbb{R}$ and $\mathbb{C}$ have been independently rediscovered on many occasions. The canonical forms over $\mathbb{C}$ were first determined by Le Paige [25] in 1881. In 1922 the problem was presented in greater detail by Schwartz [33]. Oldenburger provided simpler calculations [29] in 1932; later, the same author considered the real case [30]. The complex case has been re-analyzed more recently by Ehrenborg [11]. They classified the orbits of the group action $\mathrm{GL}_{2,2,2}(\mathbb{C})$ on $\mathbb{C}^{2\times2\times2}$ and found seven distinct orbits. As a demonstration of how such results depend on the field, V. de Silva and L.-H. Lim do a similar analysis of the group action $\mathrm{GL}_{2,2,2}(\mathbb{R})$ on $\mathbb{R}^{2\times2\times2}$ and discover eight distinct orbits. We'll consider the problem over $\mathbb{R}, \mathbb{C}$ and some finite fields $\mathbb{F}_p$, where p is prime.

2.8 Canonical Forms over $\mathbb{R}$

Since we're interested in how matrices act on order-3 tensors, recall the following:

If $\boldsymbol{X} = [x_{ijk}] \in \mathbb{R}^{d_1 \times d_2 \times d_3}$ and $L = [\lambda_{pi}] \in \mathbb{R}^{c_1 \times d_1}, M = [\mu_{qj}] \in \mathbb{R}^{c_2 \times d_2}, N = [\nu_{rk}] \in \mathbb{R}^{c_3 \times d_3}$, then the tensor $\boldsymbol{X}$ can be transformed into a new tensor $\boldsymbol{Y} = [y_{pqr}] \in \mathbb{R}^{c_1 \times c_2 \times c_3}$

via the multiplication $\boldsymbol{Y} = (L, M, N) \cdot \boldsymbol{X}$ defined by

$$y_{pqr} = \sum_{i,j,k=1}^{d_1,d_2,d_3} \lambda_{pi}\mu_{qj}\nu_{rk}x_{ijk}$$

Definition 2.27. Two tensors $\boldsymbol{X}$, $\boldsymbol{Y} \in \mathbb{R}^{2\times2\times2}$ are said to be $\mathrm{GL}_{2,2,2}(\mathbb{R})$-*equivalent* (or simply *equivalent*) if there exists $A_1, A_2, A_3 \in \mathrm{GL}_{2,2,2}(\mathbb{R})$ such that $\boldsymbol{Y} = (A_1, A_2, A_3) \cdot \boldsymbol{X}$.

In the proof of the following theorem, de Silva and Lim use the term generic (denoted G) and degenerate (denoted D) to mean an orbit of positive measure and an orbit of measure zero, respectively. The term *degenerate* means different things to different authors. In most cases, a degenerate tensor is one which does not have a best approximation by tensors of lower rank.

Theorem 2.28. *Theorem 7.1, page 1114 of [9]*
Every tensor in $\mathbb{R}^{2\times2\times2}$ is equivalent via a transformation in $GL_{2,2,2}(\mathbb{R})$ to precisely one of the eight distinct canonical forms indicated in Table 2.1, with its invariants taking the values shown.

Proof. Let $\boldsymbol{X} = \left[\begin{array}{c|c} \boldsymbol{X}_1 & \boldsymbol{X}_2 \end{array}\right]$, and consider the following cases:

Case 1. If the rank of $\boldsymbol{X}_1$ is 0, then we have

$$\left[\begin{array}{cc|cc} 0 & 0 & \times & \times \\ 0 & 0 & \times & \times \end{array}\right].$$

The use of "$\times$" indicates that the entry is not important. We can reduce $\boldsymbol{X}_2$ using matrix operations. Depending on the rank of $\boldsymbol{X}_2$, $\boldsymbol{X}$ has the following three possible forms:

$$\left[\begin{array}{cc|cc} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right], \qquad \left[\begin{array}{cc|cc} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right], \qquad \left[\begin{array}{cc|cc} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right].$$

tensor	sign(Δ)	$\text{rank}_\boxplus$	$\text{rank}_\otimes$	$\underline{\text{rank}}_\otimes$
$D_0 = \left[\begin{array}{cc\|cc} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$	0	$(0,0,0)$	0	0
$D_1 = \left[\begin{array}{cc\|cc} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$	0	$(1,1,1)$	1	1
$D_2 = \left[\begin{array}{cc\|cc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{array}\right]$	0	$(2,2,1)$	2	2
$D_2' = \left[\begin{array}{cc\|cc} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right]$	0	$(1,2,2)$	2	2
$D_2'' = \left[\begin{array}{cc\|cc} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right]$	0	$(2,1,2)$	2	2
$G_2 = \left[\begin{array}{cc\|cc} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right]$	$+$	$(2,2,2)$	2	2
$D_3 = \left[\begin{array}{cc\|cc} 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{array}\right]$	0	$(2,2,2)$	3	2
$G_3 = \left[\begin{array}{cc\|cc} 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \end{array}\right]$	$-$	$(2,2,2)$	3	3

Table 2.1: Canonical forms of $2 \times 2 \times 2$ tensors over $\mathbb{R}$

These correspond to D_0, D_1, D_2 (D stands for degenerate), respectively (after reordering the slices).

Case 2. Suppose the rank of $\boldsymbol{X}_1$ is 1. We can assume $x_{111} \neq 0$ (otherwise swapping rows and columns will bring any nonzero entry of $\boldsymbol{X}_1$ to the $(1,1,1)$-th position). Then applying operations to reduce $\boldsymbol{X}_1$ to row canonical form gives

$$\left[\begin{array}{cc|cc} 1 & 0 & a & b \\ 0 & 0 & c & d \end{array}\right].$$

We have two subcases to consider.

(i) If $d \neq 0$ then $(-\frac{b}{d} \cdot \boldsymbol{X}_{2\bullet\bullet} + \boldsymbol{X}_{1\bullet\bullet}), (-\frac{c}{d} \cdot \boldsymbol{X}_{\bullet 2\bullet} + \boldsymbol{X}_{\bullet 1\bullet})$, results in d eliminating b and c, respectively. Next, $(-a' \cdot \boldsymbol{X}_{\bullet\bullet 1} + \boldsymbol{X}_{\bullet\bullet 2})$ will eliminate a'. Then $(\frac{1}{d} \cdot \boldsymbol{X}_{\bullet\bullet 2})$ gives the next form, G_2 (G stands for generic.

$$\left[\begin{array}{cc|cc} 1 & 0 & a & b \\ 0 & 0 & c & d \end{array}\right] \rightarrow \left[\begin{array}{cc|cc} 1 & 0 & a' & 0 \\ 0 & 0 & 0 & d \end{array}\right] \rightarrow \left[\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & d \end{array}\right] \rightarrow \left[\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right].$$

(ii) If $d = 0$ then $(-a \cdot \boldsymbol{X}_{\bullet\bullet 1} + \boldsymbol{X}_{\bullet\bullet 2})$ gives

$$\left[\begin{array}{cc|cc} 1 & 0 & a & b \\ 0 & 0 & c & 0 \end{array}\right] \rightarrow \left[\begin{array}{cc|cc} 1 & 0 & 0 & b \\ 0 & 0 & c & 0 \end{array}\right].$$

According to whether b or c are zero we get

$$\left[\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right], \quad \left[\begin{array}{cc|cc} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right], \quad \left[\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right], \quad \left[\begin{array}{cc|cc} 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{array}\right].$$

These correspond to D_1, D_2', D_2'' and D_3, respectively.

Case 3. If the rank of $\boldsymbol{X}_1$ is 2, then applying operations to reduce $\boldsymbol{X}_1$ to row canonical form gives

$$\boldsymbol{X} = \left[\begin{array}{cc|cc} 1 & 0 & \times & \times \\ 0 & 1 & \times & \times \end{array}\right].$$

In this case, the entries of $\boldsymbol{X}_2$ are irrelevant to the argument. By applying a transformation of the form (A, A^{-1}, I), we can keep $\boldsymbol{X}_1$ fixed while conjugating $\boldsymbol{X}_2$ into Jordan canonical form. There are four subcases to consider.

(i) If $\boldsymbol{X}_2$ has repeated real eigenvalues and is diagonalizable, then $(-\alpha\boldsymbol{X}_{\bullet\bullet 1}+\boldsymbol{X}_{\bullet\bullet 2})$ gives D_2.

$$\boldsymbol{X} = \left[\begin{array}{cc|cc} 1 & 0 & \alpha & 0 \\ 0 & 1 & 0 & \alpha \end{array}\right] \to \left[\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{array}\right].$$

(ii) If $\boldsymbol{X}_2$ has repeated real eigenvalues and is *not* diagonalizable, $(-\alpha\boldsymbol{X}_{\bullet\bullet 1}+\boldsymbol{X}_{\bullet\bullet 2})$ gives D_3 again (after interchanging slices and columns).

$$\boldsymbol{X} = \left[\begin{array}{cc|cc} 1 & 0 & \alpha & 1 \\ 0 & 1 & 0 & \alpha \end{array}\right] \to \left[\begin{array}{cc|cc} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{array}\right] \to \left[\begin{array}{cc|cc} 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{array}\right].$$

(iii) If $\boldsymbol{X}_2$ has distinct real eigenvalues, then performing $(\boldsymbol{X}_{\bullet\bullet 2} - \alpha \cdot \boldsymbol{X}_{\bullet\bullet 1})$, $(\frac{1}{\beta-\alpha} \cdot \boldsymbol{X}_{\bullet\bullet 2})$, and $(\boldsymbol{X}_{\bullet\bullet 1} - \boldsymbol{X}_{\bullet\bullet 2})$ gives G_2 again.

$$\left[\begin{array}{cc|cc} 1 & 0 & \alpha & 0 \\ 0 & 1 & 0 & \beta \end{array}\right] \to \left[\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & \beta-\alpha \end{array}\right] \to \left[\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \end{array}\right] \to \left[\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right].$$

(iv) If $\boldsymbol{X}_2$ has distinct complex conjugate eigenvalues, then $(-a\boldsymbol{X}_{\bullet\bullet 1} + \boldsymbol{X}_{\bullet\bullet 2})$, followed by $(\frac{1}{b} \cdot \boldsymbol{X}_{\bullet\bullet 2})$ gives G_3.

$$\left[\begin{array}{cc|cc} 1 & 0 & a & -b \\ 0 & 1 & b & a \end{array}\right] \to \left[\begin{array}{cc|cc} 1 & 0 & 0 & -b \\ 0 & 1 & b & 0 \end{array}\right] \to \left[\begin{array}{cc|cc} 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \end{array}\right].$$

This completes the proof.

□

Remark 2.29. We will verify that the canonical form

$$G_3 = \left[\begin{array}{cc|cc} 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \end{array}\right]$$

has rank 3. The outer product decomposition is:

$$\begin{aligned}G_3 &= e_1 \otimes e_1 \otimes e_1 + e_2 \otimes e_2 \otimes e_1 + e_2 \otimes e_1 \otimes e_2 - e_1 \otimes e_2 \otimes e_2 \\ &= e_1 \otimes e_1 \otimes e_1 + e_2 \otimes e_2 \otimes e_1 + e_2 \otimes e_1 \otimes e_2 - e_1 \otimes e_2 \otimes e_2 \\ &\quad + (e_2 \otimes e_2 \otimes e_2 - e_2 \otimes e_2 \otimes e_2 + e_2 \otimes e_1 \otimes e_1 - e_2 \otimes e_1 \otimes e_1) \\ &= (-e_1 - e_2) \otimes e_2 \otimes e_2 + (e_1 - e_2) \otimes e_1 \otimes e_1 + e_2 \otimes (e_1 + e_2) \otimes (e_1 + e_2).\end{aligned}$$

Thus, the rank is at most 3. To prove that it is exactly 3, we employ the hyperdeterminant. Since $\Delta(G_3) < 0$, the rank must be exactly 3.

2.9 Canonical Forms over $\mathbb{C}$

Now we'll consider the group action $\mathrm{GL}_{2,2,2}(\mathbb{C})$ on $\mathbb{C}^{2\times2\times2}$. We'll summarize the ideas from [2].

Theorem 2.30. *A $2\times2\times2$ tensor $\mathbf{X}$ can be transformed to a form in which $\mathbf{X}_{\bullet\bullet1}$ has a non-zero diagonal, $\mathbf{X}_{\bullet\bullet2}$ has at least one zero entry behind the non-zero diagonal of $\mathbf{X}_{\bullet\bullet1}$, and there is at least one zero in the $(i,j,2)$-th position of $\mathbf{X}_{\bullet\bullet2}$ corresponding to a non-zero entry in the $(i,j,1)$-th position of $\mathbf{X}_{\bullet\bullet1}$. The position of the non-zero elements are unique up to permutation of the 2×2 slices.*

Proof. Since the tensor has three spatial dimensions, at least three zero entries can be created. Write $\boldsymbol{X}$ in terms of its frontal slices

$$\boldsymbol{X} = \left[\begin{array}{cc|cc} a & b & c & d \\ e & f & g & h \end{array}\right]$$

We can assume $a \neq 0$, otherwise we can interchange rows, columns, or slices. First we want to eliminate e (if $e = 0$, skip this). Applying the transformation $\left(-\frac{e}{a}\boldsymbol{X}_{1\bullet\bullet} + \boldsymbol{X}_{2\bullet\bullet}\right)$ yields

$$\left[\begin{array}{cc|cc} a & b & c & d \\ e & f & g & h \end{array}\right] \to \left[\begin{array}{cc|cc} a & b & c & d \\ 0 & f' & g' & h' \end{array}\right].$$

Now we want to eliminate b (if $b = 0$, skip this). The transformation $\left(-\frac{b}{a}\boldsymbol{X}_{\bullet 1 \bullet} + \boldsymbol{X}_{\bullet 2 \bullet}\right)$ yields

$$\left[\begin{array}{cc|cc} a & b & c & d \\ 0 & f' & g' & h' \end{array}\right] \rightarrow \left[\begin{array}{cc|cc} a & 0 & c & d' \\ 0 & f' & g' & h'' \end{array}\right].$$

Finally, we can eliminate exactly one of the non-zero entries in the second slice that lies behind the diagonal of the first slice. If $c = 0$ and $h'' = 0$ then we're finished, so without loss of generality, assume $c \neq 0$. The transformation $\left(-\frac{c}{a}\boldsymbol{X}_{\bullet\bullet 1} + \boldsymbol{X}_{\bullet\bullet 2}\right)$ gives

$$\left[\begin{array}{cc|cc} a & 0 & c & d' \\ 0 & f' & g' & h'' \end{array}\right] \rightarrow \left[\begin{array}{cc|cc} a & 0 & 0 & d' \\ 0 & f' & g' & h'' \end{array}\right].$$

We have transformed a tensor to a form that has at least three zero entries. □

Definition 2.31. Transforming a $2 \times 2 \times 2$ tensor to the form

$$\left[\begin{array}{cc|cc} \times & 0 & 0 & \times \\ 0 & \times & \times & \times \end{array}\right]$$

(up to a permutation of its slices) is called the *standard reduction* of a $2 \times 2 \times 2$ tensor.

Now we consider the different possibilities of the rank, which we know is at most 3 for $2 \times 2 \times 2$ tensors.

Lemma 2.32. *Lemma 5.5, page 18 of [2]. A $2 \times 2 \times 2$ rank-1 tensor can be transformed to a tensor with exactly one non-zero entry.*

Proof. Any $2 \times 2 \times 2$, rank-1 tensor $\boldsymbol{X}$ can be written as an outer product of three non-zero vectors. Since at least one entry in each vector must be non-zero, we can assume $\boldsymbol{X}$ has the decomposition

$$\begin{bmatrix} 1 \\ a_2 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ b_2 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ c_2 \end{bmatrix}.$$

That is, we can interchange rows, columns and slices so that the non-zero entry (at least one must exist) is in the $(1,1,1)$th position, then we can scale the vectors so that the first component of each vector is 1. If $a_2 \neq 0$ (otherwise skip this step), apply the transformation $(-a_2 \boldsymbol{X}_{1\bullet\bullet} + \boldsymbol{X}_{2\bullet\bullet})$:

$$\begin{bmatrix}1\\a_2\end{bmatrix} \otimes \begin{bmatrix}1\\b_2\end{bmatrix} \otimes \begin{bmatrix}1\\c_2\end{bmatrix} \to \begin{bmatrix}1\\0\end{bmatrix} \otimes \begin{bmatrix}1\\b_2\end{bmatrix} \otimes \begin{bmatrix}1\\c_2\end{bmatrix}.$$

If $b_2 \neq 0$ (otherwise skip this step), apply the transformation $(-b_2 \boldsymbol{X}_{\bullet 1\bullet} + \boldsymbol{X}_{\bullet 2\bullet})$:

$$\begin{bmatrix}1\\0\end{bmatrix} \otimes \begin{bmatrix}1\\b_2\end{bmatrix} \otimes \begin{bmatrix}1\\c_2\end{bmatrix} \to \begin{bmatrix}1\\0\end{bmatrix} \otimes \begin{bmatrix}1\\0\end{bmatrix} \otimes \begin{bmatrix}1\\c_2\end{bmatrix}.$$

If $c_2 \neq 0$ (otherwise skip this step), apply the transformation $(-c_2 \boldsymbol{X}_{\bullet\bullet 1} + \boldsymbol{X}_{\bullet\bullet 2})$:

$$\begin{bmatrix}1\\0\end{bmatrix} \otimes \begin{bmatrix}1\\0\end{bmatrix} \otimes \begin{bmatrix}1\\c_2\end{bmatrix} \to \begin{bmatrix}1\\0\end{bmatrix} \otimes \begin{bmatrix}1\\0\end{bmatrix} \otimes \begin{bmatrix}1\\0\end{bmatrix}.$$

$\boldsymbol{X}$ has been transformed to a tensor with exactly one non-zero entry. □

Lemma 2.33. *Lemma 5.6, page 19 of [2]. A $2 \times 2 \times 2$, rank-2 tensor can be transformed to a tensor with exactly two non-zero entries.*

Proof. Let $\boldsymbol{X}$ be a $2 \times 2 \times 2$, rank-2 tensor. Then $\boldsymbol{X}$ is the sum of two decomposable tensors, $\boldsymbol{Y}$ and $\boldsymbol{Z}$. By the previous lemma, we can apply transformations to reduce $\boldsymbol{Y}$ to a tensor with exactly one non-zero entry. Since these transformations are linear, $\boldsymbol{Z}$ will be transformed also. Thus, $\boldsymbol{X}$ can be transformed to

$$\begin{bmatrix}1\\0\end{bmatrix} \otimes \begin{bmatrix}1\\0\end{bmatrix} \otimes \begin{bmatrix}1\\0\end{bmatrix} + \begin{bmatrix}d_1\\d_2\end{bmatrix} \otimes \begin{bmatrix}e_1\\e_2\end{bmatrix} \otimes \begin{bmatrix}f_1\\f_2\end{bmatrix}$$

We want to apply transformations to reduce $\boldsymbol{X}$ such that $\boldsymbol{Y}$ remains fixed and $\boldsymbol{Z}$ is reduced to a form with exactly one non-zero entry. (i) If one of d_1 or d_2 is zero, then skip to (ii). Otherwise, the transformation $\left(-\frac{d_1}{d_2} \boldsymbol{X}_{2\bullet\bullet} + \boldsymbol{X}_{1\bullet\bullet}\right)$ gives:

$$\begin{bmatrix}1\\0\end{bmatrix} \otimes \begin{bmatrix}1\\0\end{bmatrix} \otimes \begin{bmatrix}1\\0\end{bmatrix} + \begin{bmatrix}d_1\\d_2\end{bmatrix} \otimes \begin{bmatrix}e_1\\e_2\end{bmatrix} \otimes \begin{bmatrix}f_1\\f_2\end{bmatrix} \to \begin{bmatrix}1\\0\end{bmatrix} \otimes \begin{bmatrix}1\\0\end{bmatrix} \otimes \begin{bmatrix}1\\0\end{bmatrix} + \begin{bmatrix}0\\d_2\end{bmatrix} \otimes \begin{bmatrix}e_1\\e_2\end{bmatrix} \otimes \begin{bmatrix}f_1\\f_2\end{bmatrix}.$$

(ii) If one of e_1 or e_2 is zero, then skip to (iii). Otherwise, the transformation $\left(-\frac{e_1}{e_2}\boldsymbol{X}_{\bullet 2\bullet} + \boldsymbol{X}_{\bullet 1\bullet}\right)$ gives:

$$\begin{bmatrix}1\\0\end{bmatrix}\otimes\begin{bmatrix}1\\0\end{bmatrix}\otimes\begin{bmatrix}1\\0\end{bmatrix}+\begin{bmatrix}0\\d_2\end{bmatrix}\otimes\begin{bmatrix}e_1\\e_2\end{bmatrix}\otimes\begin{bmatrix}f_1\\f_2\end{bmatrix}\to\begin{bmatrix}1\\0\end{bmatrix}\otimes\begin{bmatrix}1\\0\end{bmatrix}\otimes\begin{bmatrix}1\\0\end{bmatrix}+\begin{bmatrix}0\\d_2\end{bmatrix}\otimes\begin{bmatrix}0\\e_2\end{bmatrix}\otimes\begin{bmatrix}f_1\\f_2\end{bmatrix}.$$

(iii) If one of f_1 or f_2 is zero, then we're finished. Otherwise, the transformation $\left(-\frac{f_1}{f_2}\boldsymbol{X}_{\bullet\bullet 2} + \boldsymbol{X}_{\bullet\bullet 1}\right)$ gives:

$$\begin{bmatrix}1\\0\end{bmatrix}\otimes\begin{bmatrix}1\\0\end{bmatrix}\otimes\begin{bmatrix}1\\0\end{bmatrix}+\begin{bmatrix}0\\d_2\end{bmatrix}\otimes\begin{bmatrix}0\\e_2\end{bmatrix}\otimes\begin{bmatrix}f_1\\f_2\end{bmatrix}\to\begin{bmatrix}1\\0\end{bmatrix}\otimes\begin{bmatrix}1\\0\end{bmatrix}\otimes\begin{bmatrix}1\\0\end{bmatrix}+\begin{bmatrix}0\\d_2\end{bmatrix}\otimes\begin{bmatrix}0\\e_2\end{bmatrix}\otimes\begin{bmatrix}0\\f_2\end{bmatrix}$$

We have transformed $\boldsymbol{X}$ to a tensor with exactly two non-zero entries. □

Definition 2.34. The tensor

$$\boldsymbol{X}=\begin{bmatrix}1\\0\end{bmatrix}\otimes\begin{bmatrix}1\\0\end{bmatrix}\otimes\begin{bmatrix}1\\0\end{bmatrix}+\begin{bmatrix}0\\1\end{bmatrix}\otimes\begin{bmatrix}0\\1\end{bmatrix}\otimes\begin{bmatrix}1\\0\end{bmatrix}+\begin{bmatrix}0\\1\end{bmatrix}\otimes\begin{bmatrix}1\\0\end{bmatrix}\otimes\begin{bmatrix}0\\1\end{bmatrix}=\left[\begin{array}{cc|cc}1&0&0&0\\0&1&1&0\end{array}\right]$$

is called the *independent tensor*. Any $2\times2\times2$ tensor that can be transformed to the independent tensor by swapping slices is called an *independent tensor*.

Proposition 2.35. *Proposition 5.8, page 19 of [2]*

The independent tensor has rank-3.

Proof. Suppose that the independent tensor $\boldsymbol{X}$ can be written as a sum of two rank-1 tensors, $\boldsymbol{A}$ and $\boldsymbol{B}$:

$$\left[\begin{array}{cc|cc}1&0&0&0\\0&1&1&0\end{array}\right]=\left[\begin{array}{cc|cc}a_{111}&a_{121}&a_{112}&a_{122}\\a_{211}&a_{221}&a_{212}&a_{222}\end{array}\right]+\left[\begin{array}{cc|cc}b_{111}&b_{121}&b_{112}&b_{122}\\b_{211}&b_{221}&b_{212}&b_{222}\end{array}\right].$$

To make the computations easier, let's scale both sides of the equation so that $a_{111}=1$. The rank-1 tensors $\boldsymbol{A}$ and $\boldsymbol{B}$ must both contain only non-zero entries, otherwise it would not be possible to create the independent tensor. Since both $\boldsymbol{A}$ and $\boldsymbol{B}$ are rank-1 with only non-zero entries, there are scalars $\alpha_1, \alpha_2, \alpha_3$, and $\beta_1, \beta_2, \beta_3$ such that

$$\left[\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{array}\right] = \left[\begin{array}{cc|cc} 1 & \alpha_1 & \alpha_3 & a_{122} \\ \alpha_2 & a_{221} & a_{212} & a_{222} \end{array}\right] + \left[\begin{array}{cc|cc} b_{111} & \beta_1 b_{111} & \beta_3 b_{111} & b_{122} \\ \beta_2 b_{111} & b_{221} & b_{212} & b_{222} \end{array}\right].$$

Since $x_{121} = x_{211} = x_{112} = 0$ the following equations holds:

$$-\alpha_1 = \beta_1 b_{111}, \quad -\alpha_2 = \beta_2 b_{111}, \quad -\alpha_3 = \beta_3 b_{111}.$$

Now, since $x_{122} = 0$, we get $a_{122} = -b_{122}$. Then again by the fact that $\boldsymbol{A}$ and $\boldsymbol{B}$ are rank-1, we get $\alpha_1\alpha_2 = -\beta_1\beta_2 b_{111}^2$. By substituting, for example, $-\alpha_1 = \beta_1 b_{111}$ we obtain $\alpha_2 = \beta_2 b_{111}$, which implies $b_{111} = 0$, a contradiction to the assumption that the entries of $\boldsymbol{B}$ be non-zero. □

Theorem 2.36. *Theorem 5.13, page 26 of [2]*
The transformations in $GL_{2,2,2}(\mathbb{C})$ on the set of all tensors in $\mathbb{C}^{2\times 2\times 2}$ induce seven orbits: four in rank-2 and one in each of the other ranks (0,1, and 3).

Proof. Case 1. Clearly, for rank-0 tensors the only orbit is the rank-0 tensor. Hence

$$\left[\begin{array}{cc|cc} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

is the representative of the rank-0 orbit.

Case 2. For rank-1 tensors there is also one orbit. Remember that every rank-1 tensor can be reduced to a tensor with only one non-zero entry. This entry can be translated to any position in the tensor and can be scaled to any non-zero number. In particular, scaling it to 1 gives the representative of the rank-1 orbit,

$$\left[\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Case 3. Recall that a rank-2 tensor can be transformed to a tensor with exactly two non-zero elements. Without loss of generality, assume one non-zero entry is in the (1,1,1)-th position. We're interested in where the second non-zero entry is located with respect to the first. In this case there are exactly four possibilities:

$$\boldsymbol{X}_1 = \left[\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right], \quad \boldsymbol{X}_2 = \left[\begin{array}{cc|cc} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right],$$

$$\boldsymbol{X}_3 = \left[\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{array}\right], \quad \boldsymbol{X}_4 = \left[\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right].$$

To see why $\boldsymbol{X}_1$ is not in the same orbit as the other three choices, we apply the hyperdeterminant. Since the hyperdeterminant of $\boldsymbol{X}_1$ vanishes (while the hyperdeterminant of the other three do not), and since the hyperdeterminant of a $2 \times 2 \times 2$ tensor is invariant under the action of $\mathrm{GL}_{2,2,2}(\mathbb{C})$, we see that $\boldsymbol{X}_1$ is in a different orbit and hence is a representative of that orbit. For the other three choices, $\boldsymbol{X}_3$ is not in the orbits of the other two because to transform $\boldsymbol{X}_3$ into either of the other two choices we must first create a non-zero entry in slice $\boldsymbol{X}_{\bullet\bullet 2}$. The only way to do this is to perform $(\boldsymbol{X}_{\bullet\bullet 1} + \boldsymbol{X}_{\bullet\bullet 2})$. But then no matter what operation follows, since the front slice was copied to the back slice, the back slice remains a scalar multiple of the first. Finally, we consider $\boldsymbol{X}_2$ and $\boldsymbol{X}_4$. A similar argument shows that by performing $(\boldsymbol{X}_{1\bullet\bullet} + \boldsymbol{X}_{2\bullet\bullet})$ on $\boldsymbol{X}_2$ it is not possible to create $\boldsymbol{X}_4$.

Case 4. Recall the standard reduction of a $2 \times 2 \times 2$ tensor,

$$\left[\begin{array}{cc|cc} a & 0 & c & d \\ 0 & b & e & 0 \end{array}\right]$$

We know that the maximum rank of any $2 \times 2 \times 2$ tensor is 3, and hence depending on how many of these five entries are non-zero tells us how to classify all types of $2 \times 2 \times 2$ tensors. We omit the details and state the most important points: (i) the independent tensor occurs in every case in which the rank is three, and (ii) when there are four non-zero entries (i.e. if either $a = 0, c = 0$, or $e = 0$) a decomposition into two rank-1 tensors can be given only if there is a non-real number as one of the entries in both tensors. That is, if we can write (say, for $a = 0$),

$$\left[\begin{array}{cc|cc} 0 & 0 & c & d \\ 0 & b & e & 0 \end{array}\right] = \left[\begin{array}{cc|cc} 0 & 0 & 0 & 0 \\ 0 & b & 0 & -x \end{array}\right] + \left[\begin{array}{cc|cc} 0 & 0 & c & d \\ 0 & 0 & e & x \end{array}\right]$$

where $x \in \mathbb{C}$ is the unique complex number satisfying the decomposition. We see that over $\mathbb{R}$ this decomposition must occur as a sum of three rank-1 tensors, which gives the extra orbit. A natural concern may be whether the position of the non-zero entries matters. In fact, it does not since transformations (for example, switching slices) non-trivially permutes a direction, and hence every possible configuration of the independent matrix can be reached. This follows from the fact that the independent tensor is symmetric in the three dimensions. We conclude that the independent tensor is the representative of this orbit,

$$\left[\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{array}\right].$$

□

Table 2.2 summarizes the canonical forms over $\mathbb{C}$. Recall that over $\mathbb{R}$ there were two rank-3 canonical forms: G_2 and G_3. Over $\mathbb{C}$, these combine into a single orbit class since there is no longer any distinction between real and complex eigenvalues. To see this explicitly, take $x_i, y_i \in \mathbb{R}^2$ for $i = 1, 2$ (not necessarily linearly independent). Then the tensor

$$x_1 \otimes x_2 \otimes x_3 + y_1 \otimes x_2 \otimes y_3 - x_1 \otimes y_2 \otimes y_3 + y_1 \otimes x_2 \otimes y_3$$

is in the $\mathrm{GL}_{2,2,2}(\mathbb{R})$-orbit class of G_3 and rank-3 over $\mathbb{R}$. Let $z_k = x_k + iy_k$ and $\bar{z_k} = x_k - iy_k$ for $k = 1, 2, 3$. Then the tensor above can be written as

$$\frac{1}{2}(\bar{z_1} \otimes z_2 \otimes \bar{z_3} + z_1 \otimes \bar{z_2} \otimes z_3).$$

This is in the $\mathrm{GL}_{2,2,2}(\mathbb{C})$-orbit class of G_2 and has rank-2 over $\mathbb{C}$ [9].

rank	canonical form	$\operatorname{rank}_{\boxplus}$
0	$\left[\begin{array}{cc\|cc} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$	$(0,0,0)$
1	$\left[\begin{array}{cc\|cc} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$	$(1,1,1)$
2	$\left[\begin{array}{cc\|cc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{array}\right]$	$(2,2,1)$
2	$\left[\begin{array}{cc\|cc} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right]$	$(1,2,2)$
2	$\left[\begin{array}{cc\|cc} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right]$	$(2,1,2)$
2	$\left[\begin{array}{cc\|cc} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right]$	$(2,2,2)$
3	$\left[\begin{array}{cc\|cc} 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{array}\right]$	$(2,2,2)$

Table 2.2: Canonical forms of $2 \times 2 \times 2$ tensors over $\mathbb{C}$

2.10 Canonical Forms over Finite Fields

Consider a general n-dimensional tensor $\boldsymbol{X} = [x_{i_1 \dots i_n}]$ of format $2 \times \cdots \times 2$ (n factors) over $\mathbb{F}$. The flattening of $\boldsymbol{X}$ is the vector $\text{flat}(\boldsymbol{X}) = [x_{1\dots 1}, \dots, x_{i_1 \dots i_n}, \dots x_{2\dots 2}]$. The entries are in lex order of the n-tuples of subscripts: $i_1 \dots i_n$ precedes $i'_1 \dots i'_n$ if and only if $i_j < i'_j$ where j is the least index such that $i_j \neq i'_j$. The tensor $\boldsymbol{X}$ precedes the tensor $\boldsymbol{Y}$ if $\text{flat}(\boldsymbol{X})$ precedes $\text{flat}(\boldsymbol{Y})$ in lex order: $x_{i_1 \dots i_n} < y_{i_1 \dots i_n}$ where $i_1 \dots i_n$ is the least n-tuple with $x_{i_1 \dots i_n} \neq y_{i_1 \dots i_n}$. For $\mathbb{F}_2, \mathbb{F}_3$ and $\mathbb{F}_5$ we use the standard total order $0 < 1$, $0 < 1 < 2$, and $0 < 1 < 2 < 3 < 4$, respectively. This order allows us to define the minimal element for a set of tensors.

Using computer algebra, we classify the canonical forms with respect to the action of the symmetry groups $\text{GL}_{2,\dots,2}(\mathbb{F}) \rtimes S_n$, for $n = 3, 4$ and $\mathbb{F} = \mathbb{F}_p$ for $p = 2, 3, 5$. For each canonical form, we determine the size of its orbit and the rank of the tensors in that orbit. Just as we did over $\mathbb{R}$ and $\mathbb{C}$, we consider the action of the direct product $\text{GL}_{2,\dots,2}(\mathbb{F})$ acting by simultaneous changes of basis along the n directions. We will refer to this as the small symmetry group. Additionally, the semi-direct product $\text{GL}_{2,\dots,2}(\mathbb{F}) \rtimes S_n$, where S_n acts by permuting the n factors is the large symmetry group. The actions of the symmetry groups decompose the set of $2 \times \cdots \times 2$ tensors into a disjoint union of orbits, where the tensors in each orbit are equivalent under the group action. The canonical form is the minimal element of each orbit.

For $2 \times \cdots \times 2$ tensors over $\mathbb{F}_p$, lower bounds for the number of canonical forms with respect to the small and large symmetry groups can be computed. The order of the small symmetry group is $(p^2 - 1)^n (p^2 - p)^n$. Let's see why this is. The order of $\text{GL}_2(\mathbb{F}_p)$ is the number of linearly independent vectors in a two-dimensional vector space over $\mathbb{F}_p$. For the first vector, there are $p^2 - 1$ possible non-zero choices. For the second vector, there are $p^2 - p$ choices: all vectors that aren't a multiple of the first vector. Thus, the order of $\text{GL}_2(\mathbb{F}_p)$ is $(p^2 - 1)(p^2 - p)$. Since there are n factors in the small symmetry group $GL_{2,\ldots,2}(\mathbb{F}_p)$, the order is $(p^2-1)^n(p^2-p)^n$. For the large symmetry group, we multiply the order of the small symmetry group by $n!$, the order of S_n. Thus, lower bounds of the orbits with respect to the small and large symmetry groups are given by

$$s(n,p) = \left\lceil \frac{p^{2^n}}{(p^2-1)^n(p^2-p)^n} \right\rceil, \qquad l(n,p) = \left\lceil \frac{p^{2^n}}{(p^2-1)^n(p^2-p)^n n!} \right\rceil.$$

These formulas show how quickly the number of canonical forms grows into an unmanageable number:

n	$s(n,2)$	$l(n,2)$	$s(n,3)$	$l(n,3)$
3	2	1	1	1
4	51	3	9	1
5	552337	4603	7272337	60603

We consider $2\times 2\times 2$ tensors over $\mathbb{F}_p$, where $p = 2,3,5$. We begin by describing how to classify tensors by rank.

• The zero tensor is the only rank-0 tensor. Let $a,b,c \in \mathbb{F}_p^2$ be non-zero vectors. Then the rank-1 tensors belong to $\left\{ a\otimes b\otimes c \mid a,b,c \in \mathbb{F}_p^2 \right\}$.

• Suppose we have computed the set of rank-r tensors. To compute the tensors of rank-$(r+1)$, we consider all possible sums $\boldsymbol{X}+\boldsymbol{Y}$ where $\text{rank}(\boldsymbol{X}) = r$ and $\text{rank}(\boldsymbol{Y}) = 1$. Clearly, $\text{rank}(\boldsymbol{X}+\boldsymbol{Y}) \le r+1$. However, it is possible that $\text{rank}(\boldsymbol{X}+\boldsymbol{Y}) \le r$ and hence we only keep the tensors which have rank exactly $r+1$.

• After the tensors in each rank have been computed, we sort the tensors (within each rank) in lex order. For each rank, we choose the minimal tensor and compute the orbit of this tensor under the group action. Then we remove the tensors of this orbit from the set. This process is complete once there are no more tensors of the current rank.

In subsection 2.10.1 we present the Maple programs used for our calculations, and in subsections 2.10.2 to 2.10.4 we provide summaries of the output.

2.10.1 Maple Code for Canonical Forms

The following Maple procedure computes the tensors in a given rank and sorts them in lex order. The minimal element is the canonical form, and we apply the small symmetry group to compute the orbit of this element. We continue until there are no more tensors in a given rank. At each step below we provide an explanation of our methods.

```
# the outerproduct procedure takes three vectors and computes the
# outer product


outerproduct := proc( a, b, c )
  global PRIME:
  local i, j, k, result:
  result := table():
  for i to 2 do for j to 2 do for k to 2 do result[i,j,k]:= 0 od od od:
  for i to 2 do for j to 2 do for k to 2 do
    result[i,j,k] := ( result[i,j,k] + a[i]*b[j]*c[k] ) mod PRIME
  od od od:
  RETURN( eval(result) )
end:


# convert a tensor in table format to list (vector) format


flatten := proc( x )
  local i, j, k, xflat:
  xflat := []:
  for i to 2 do for j to 2 do for k to 2 do
    xflat := [ op(xflat), x[i,j,k] ]
```

```
  od od od:
  RETURN( xflat )
end:

# convert a tensor in list (vector) format to table format

unflatten := proc( x )
  local i, j, k, result, t:
  result := table():
  t := 0:
  for i to 2 do for j to 2 do for k to 2 do
    t := t + 1:
    result[i,j,k] := x[t]
  od od od:
  RETURN( eval(result) )
end:

# convert a tensor in list (vector) format to a row vector of
# two 2x2 matrices

matrixform := proc( x )
  local i, j, k, result, t:
  result := Vector[row]( 2, i -> Matrix(2,2) ):
  t := 0:
  for i to 2 do for j to 2 do for k to 2 do
    t := t + 1:
    result[i][j,k] := x[t]
```

```
  od od od:
  RETURN( result )
end:

# compare two tensors as lists in lexicographical order

compare := proc( x, y )
  local i, ii, result:
  ii := 0:
  for i from nops(x) to 1 by -1 do
    if x[i] <> y[i] then ii := i fi
  od:
  if ii = 0 then
    result := false
  else
    result := evalb( x[ii] < y[ii] )
  fi:
  RETURN( result )
end:

# compute all tensors of given rank starting with rank 0

tensorrank := table():
tensorrank[0] := { [0,0,0,0,0,0,0,0] }:

VECTORS := sort( VECTORS, compare );
```

```
# compute the outer product of three non-zero vectors.

tensorrank[1] := {}:
for a in VECTORS do for b in VECTORS do for c in VECTORS do
  abc := flatten( outerproduct( a, b, c ) ):
  if not member( abc, tensorrank[0] ) and
     not member( abc, tensorrank[1] ) then
    tensorrank[1] := tensorrank[1] union { abc }
  fi
od od od:

r := 1:
while tensorrank[r] <> {} do
  printf( "  \n" ):
  printf( "  starting rank %d \n", r+1 ):
  start := time():
  tensorrank[r+1] := {}:
  xcount := 0:
  for x in tensorrank[r] do
    xcount := xcount + 1:
    if xcount mod 100 = 0 then print(xcount, time()-start) fi:
    for y in tensorrank[1] do
      z := [ seq( ( x[i]+y[i] ) mod PRIME, i=1..NUMBER ) ]:
      new := true:
      for s from 0 to r+1 do
        new := new and not member( z, tensorrank[s] )
      od:
```

```
      if new then
        tensorrank[r+1] := tensorrank[r+1] union {z}
      fi
    od
  od:
  printf( "  \n" ):
  printf("total %3d percent %a time %a \n", nops(tensorrank[r+1]),
    evalf(100*nops(tensorrank[r+1])/(PRIME^NUMBER)),time()-start ):
  r := r + 1
od:

maximumrank := r - 1;

add( nops(tensorrank[i]), i=0..maximumrank );

for r from 0 to maximumrank do
  tensorrank[r] := sort( convert(tensorrank[r],list), compare )
od:

printf( "  \n" ):
for r from 0 to maximumrank do
  printf( "  rank %d   number %4d   %a \n",
    r, nops(tensorrank[r]), tensorrank[r][1] ):
  print( matrixform(tensorrank[r][1]) )
od:

# compute orbits of tensors within each rank
```

```
# group action does not change rank of tensor

# the group of invertible 2 x 2 matrices over the field F_p

GROUP := []:
for a from 0 to PRIME-1 do
for b from 0 to PRIME-1 do
for c from 0 to PRIME-1 do
for d from 0 to PRIME-1 do
  if ( a*d - b*c ) mod PRIME <> 0 then
    GROUP := [ op(GROUP), Matrix([[a,b],[c,d]]) ]
  fi
od od od od:

# act on a tensor by a group element

groupaction := proc( g, xlist, m )
  global PRIME:
  local gv, i, j, k, result, v, xtable:
  xtable := unflatten( xlist ):
  if m = 1 then
    for j to 2 do for k to 2 do
      v := Vector( [ xtable[1,j,k], xtable[2,j,k] ] ):
      gv := [ ( g[1,1]*v[1] + g[1,2]*v[2] ) mod PRIME,
              ( g[2,1]*v[1] + g[2,2]*v[2] ) mod PRIME ]:
      xtable[1,j,k] := gv[1]:
      xtable[2,j,k] := gv[2]
```

```
    od od
  fi:
  if m = 2 then
    for i to 2 do for k to 2 do
      v := Vector( [ xtable[i,1,k], xtable[i,2,k] ] ):
      gv := [ ( g[1,1]*v[1] + g[1,2]*v[2] ) mod PRIME,
              ( g[2,1]*v[1] + g[2,2]*v[2] ) mod PRIME ]:
      xtable[i,1,k] := gv[1]:
      xtable[i,2,k] := gv[2]
    od od
  fi:
  if m = 3 then
    for i to 2 do for j to 2 do
      v := Vector( [ xtable[i,j,1], xtable[i,j,2] ] ):
      gv := [ ( g[1,1]*v[1] + g[1,2]*v[2] ) mod PRIME,
              ( g[2,1]*v[1] + g[2,2]*v[2] ) mod PRIME ]:
      xtable[i,j,1] := gv[1]:
      xtable[i,j,2] := gv[2]
    od od
  fi:
  result := flatten( xtable ):
  RETURN( result )
end:

# compute the orbit of a given tensor under the direct product group

smallorbit := proc( x )
```

```
  global GROUP:
  local a, ax, bax, cbax, b, c, result:
  result := {}:
  for a in GROUP do
    ax := groupaction( a, x, 1 ):
    for b in GROUP do
      bax := groupaction( b, ax, 2 ):
      for c in GROUP do
        cbax := groupaction( c, bax, 3 ):
        result := result union { cbax }
      od
    od
  od:
  RETURN( result )
end:

representatives := table():
for r from 0 to maximumrank do
  printf( "  \n" ):
  printf( "  small orbits of rank %d \n", r ):
  representatives[r] := []:
  printf( "  \n" ):
  remaining := copy( tensorrank[r] ):
  while remaining <> [] do
    remaining := sort( remaining, compare ):
    x := remaining[1]:
    start := time():
```

```
    xorbit := sort( convert( smallorbit(x), list ), compare ):
    representatives[r] := [ op(representatives[r]), xorbit[1] ]:
    printf( "  small orbit size %4d    element %a    time %a \n",
    nops(xorbit), xorbit[1], time()-start ):
    remaining := convert(remaining,set) minus convert(xorbit,set):
    remaining := convert(remaining,list)
  od
od:

# total number of orbits (representatives are the canonical forms)

add( nops(representatives[i]), i=0..maximumrank );

for i from 0 to maximumrank do
  printf( "  \n" ):
  printf( "  small canonical forms of rank %d \n", i ):
  printf( "  \n" ):
  for x in representatives[i] do print( x, matrixform(x) ) od
od:

# compute the orbit of a given tensor under the semidirect product

largeorbit := proc( x )
  global S3:
  local i, j, k, ijk, p, px, py, y, result:
  y := unflatten( x ):
  result := {}:
```

```
  for p in S3 do
    py := table():
    for i to 2 do for j to 2 do for k to 2 do
      ijk := [i,j,k]:
      py[i,j,k] := y[ ijk[p[1]], ijk[p[2]], ijk[p[3]] ]:
      px := flatten( py )
    od od od:
    result := result union smallorbit( px )
  od:
  RETURN( result )
end:

representatives := table():
for r from 0 to maximumrank do
  printf( "  \n" ):
  printf( "  large orbits of rank %d \n", r ):
  representatives[r] := []:
  printf( "  \n" ):
  remaining := copy( tensorrank[r] ):
  while remaining <> [] do
    remaining := sort( remaining, compare ):
    x := remaining[1]:
    start := time():
    xorbit := sort( convert( largeorbit(x), list ), compare ):
    representatives[r] := [ op(representatives[r]), xorbit[1] ]:
    printf( "  large orbit size %4d    element %a    time %a \n",
    nops(xorbit), xorbit[1], time()-start ):
```

```
    remaining := convert(remaining,set) minus convert(xorbit,set):
    remaining := convert(remaining,list)
  od
od:

# total number of orbits (representatives are the canonical forms)

add( nops(representatives[i]), i=0..maximumrank );

for i from 0 to maximumrank do
  printf( "  \n" ):
  printf( "  large canonical forms of rank %d \n", i ):
  printf( "  \n" ):
  for x in representatives[i] do print( x, matrixform(x) ) od
od:
```

2.10.2 Canonical Forms over $\mathbb{F}_2$

For the case $p = 2$, the maximum rank is 3. The number of tensors in each rank and the approximate percentages are listed below.

rank	0	1	2	3
number	1	27	162	66
$\approx$ %	0	11	63	26

For the large symmetry group $\mathrm{GL}_{2,2,2}(\mathbb{F}_2) \rtimes S_3$ we present the ranks, orbit sizes, and canonical forms in Table 2.3. For the small symmetry group $\mathrm{GL}_{2,2,2}(\mathbb{F}_2)$, the first

rank	orbit size	canonical form
0	1	$\left[\begin{array}{cc\|cc} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$
1	27	$\left[\begin{array}{cc\|cc} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right]$
2	54	$\left[\begin{array}{cc\|cc} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{array}\right]$
2	108	$\left[\begin{array}{cc\|cc} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{array}\right]$
3	54	$\left[\begin{array}{cc\|cc} 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{array}\right]$
3	12	$\left[\begin{array}{cc\|cc} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \end{array}\right]$

Table 2.3: Large orbits of $2 \times 2 \times 2$ tensors over $\mathbb{F}_2$

orbit in rank 2 splits into three orbits each of size 18 with canonical forms

$$\left[\begin{array}{cc|cc} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{array}\right], \qquad \left[\begin{array}{cc|cc} 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{array}\right], \qquad \left[\begin{array}{cc|cc} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{array}\right].$$

See Table 2.4.

rank	orbit size	canonical form
0	1	$\left[\begin{array}{cc\|cc} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$
1	27	$\left[\begin{array}{cc\|cc} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right]$
2	18	$\left[\begin{array}{cc\|cc} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{array}\right]$
2	18	$\left[\begin{array}{cc\|cc} 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{array}\right]$
2	18	$\left[\begin{array}{cc\|cc} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{array}\right]$
2	108	$\left[\begin{array}{cc\|cc} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{array}\right]$
3	54	$\left[\begin{array}{cc\|cc} 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{array}\right]$
3	12	$\left[\begin{array}{cc\|cc} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \end{array}\right]$

Table 2.4: Small orbits of $2 \times 2 \times 2$ tensors over $\mathbb{F}_2$

2.10.3 Canonical Forms over $\mathbb{F}_3$

For $p = 3$ the maximum rank is 3. The number of tensors in each rank and the approximate percentages are listed below.

rank	0	1	2	3
number	1	128	4032	2400
$\approx$ %	0	2	61	37

For the large symmetry group $\mathrm{GL}_{2,2,2}(\mathbb{F}_3) \rtimes S_3$ we present the ranks, orbit sizes, and canonical forms in Table 2.5. The tables differ only in the size of the orbits and the last canonical form. For the small symmetry group $\mathrm{GL}_{2,2,2}(\mathbb{F}_3)$, the first orbit in rank 2 splits into three orbits each of size 192 with the same canonical forms as before,

$$\left[\begin{array}{cc|cc} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{array}\right], \qquad \left[\begin{array}{cc|cc} 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{array}\right], \qquad \left[\begin{array}{cc|cc} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{array}\right].$$

See Table 2.6.

2.10.4 Canonical Forms over $\mathbb{F}_5$

The Maple procedure used to determine the ranks and canonical forms of $2 \times 2 \times 2$ tensors is not very efficient over larger finite fields. In order to compute the ranks and canonical forms, more sophisticated algorithms and programming techniques are used.

The first improvement uses the fact that over finite fields $\mathbb{F}_p$, the general linear group $\mathrm{GL}_2(\mathbb{F}_p)$ is generated by two matrices. This was proved by Waterhouse in [37]. For $p = 5$, $\mathrm{GL}_2(\mathbb{F}_{|})$ has order 480 and is generated by

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

rank	orbit size	canonical form
0	1	$\left[\begin{array}{cc\|cc} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$
1	128	$\left[\begin{array}{cc\|cc} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right]$
2	576	$\left[\begin{array}{cc\|cc} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{array}\right]$
2	3456	$\left[\begin{array}{cc\|cc} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{array}\right]$
3	1536	$\left[\begin{array}{cc\|cc} 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{array}\right]$
3	864	$\left[\begin{array}{cc\|cc} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 2 \end{array}\right]$

Table 2.5: Large orbits of $2 \times 2 \times 2$ tensors over $\mathbb{F}_3$

rank	orbit size	canonical form
0	1	$\left[\begin{array}{cc\|cc} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$
1	128	$\left[\begin{array}{cc\|cc} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right]$
2	192	$\left[\begin{array}{cc\|cc} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{array}\right]$
2	192	$\left[\begin{array}{cc\|cc} 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{array}\right]$
2	192	$\left[\begin{array}{cc\|cc} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{array}\right]$
2	3456	$\left[\begin{array}{cc\|cc} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{array}\right]$
3	1536	$\left[\begin{array}{cc\|cc} 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{array}\right]$
3	864	$\left[\begin{array}{cc\|cc} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 2 \end{array}\right]$

Table 2.6: Small orbits of $2 \times 2 \times 2$ tensors over $\mathbb{F}_3$

and any one of the following matrices

$$\begin{bmatrix} 1 & 0 \\ 1 & 3 \end{bmatrix}, \qquad \begin{bmatrix} 1 & 0 \\ 2 & 2 \end{bmatrix}, \qquad \begin{bmatrix} 1 & 0 \\ 3 & 2 \end{bmatrix}, \qquad \begin{bmatrix} 1 & 0 \\ 4 & 3 \end{bmatrix}$$

To compute the action of the large symmetry group, we use a more efficient algorithm called the **spinning algorithm**.

Algorithm 2.37. To compute the orbit of an element x under the group action:

- Set $\mathcal{O} \leftarrow \{\,\}$ and set $\mathcal{N} \leftarrow \{x\}$.
- Perform the following iteration until $\mathcal{N} = \{\,\}$:
  - Set $\mathcal{O} \leftarrow \mathcal{N}$.
  - Apply the generators of the group to every element of $\mathcal{N}$, obtaining a set $\mathcal{M}$ of new elements in the orbit of x.
  - Set $\mathcal{N} \leftarrow \mathcal{M} \setminus \mathcal{O}$ and set $\mathcal{O} \leftarrow \mathcal{O} \cup \mathcal{M}$.

At termination, the set $\mathcal{O}$ is stable under the action of the generators, and hence also under the action of the group. Therefore $\mathcal{O}$ is the orbit of x under the group action, since every element of $\mathcal{O}$ was obtained by applying a group element to x.

The next improvement is based on the use of **signal arrays**, with encoding and decoding. We reverse the approach used for $2 \times 2 \times 2$ tensors: we first compute the orbits under the action of the symmetry group, and then we determine the rank corresponding to each orbit. This approach was suggested in a discussion with Jiaxiong Hu, a former M.Sc. student of Murray Bremner. We thank Jiaxiong for his helpful programming tips.

There are $p^8 - 1$ nonzero arrays of size $2 \times 2 \times 2$ over $\mathbb{F}_p$. We create a signal array with this number of entries, allocate one byte for each entry, and initialize every entry to 0. It is easy to go back and forth between the indices of this signal array and the corresponding arrays over $\mathbb{F}_p$ by using the **flatten** and **unflatten** procedures defined in the Maple program for $2 \times 2 \times 2$ tensors: the flattened form of an array over $\mathbb{F}_p$ is the base p numeral for the corresponding index in the signal array.

Set the orbit counter c to 0, and then perform the following iteration:

- Increment the orbit counter c.
- Find the leftmost 0 in the signal array (i.e., the least index with entry 0). Call this index i; at the first iteration, $i = 1$.
- Decode the index i to obtain the corresponding array x over $\mathbb{F}_p$. Compute the orbit $\mathcal{O}$ of x.
- Encode the arrays in $\mathcal{O}$ to find the corresponding indices of the signal array.
- Store c in each entry of the signal array corresponding to an array in $\mathcal{O}$.

This iteration terminates when every entry of the signal array is nonzero. Upon termination, c equals the total number of orbits.

To compute the rank corresponding to each orbit, we create another signal array of the same size and set every entry to 0. Every entry of this new signal array will eventually contain the rank of the corresponding array over $\mathbb{F}_p$. We start by noting that the first orbit found by the previous iteration consists exactly of the tensors over $\mathbb{F}_p$ which have rank 1; we therefore set the corresponding entries of the new signal array to 1. Assume that the tensors of rank r have already been marked in the new signal array. To each tensor of rank r, we add the minimal tensor of rank 1 (its flattened form has seven 0s and one 1), and obtain a tensor of rank $r + 1$. We find the orbit counter corresponding to this tensor, and for each tensor in this orbit, we store $r + 1$ in the corresponding entry of the new signal array.

For the case $p = 5$, the maximum rank is 3. The number of tensors in each rank and the approximate percentages are listed below.

rank	0	1	2	3
number	1	864	224,640	165,120
≈ %	0	0.22%	57.51%	42.27%

For the large symmetry group $\mathrm{GL}_{2,2,2}(\mathbb{F}_2) \rtimes S_3$ we present the ranks, orbit sizes, and canonical forms in Table 2.7. For the small symmetry group $\mathrm{GL}_{2,2,2}(\mathbb{F}_5)$, the first orbit in rank 2 splits into three orbits each of size 2880 with canonical forms

$$\left[\begin{array}{cc|cc} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{array}\right], \qquad \left[\begin{array}{cc|cc} 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{array}\right], \qquad \left[\begin{array}{cc|cc} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{array}\right].$$

See Table 2.8. We point out that for primes greater than 5, the amount of memory required for the program was too large.

rank	orbit size	canonical form
0	1	$\left[\begin{array}{cc\|cc} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$
1	864	$\left[\begin{array}{cc\|cc} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right]$
2	8640	$\left[\begin{array}{cc\|cc} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{array}\right]$
2	216000	$\left[\begin{array}{cc\|cc} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{array}\right]$
3	69120	$\left[\begin{array}{cc\|cc} 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{array}\right]$
3	96000	$\left[\begin{array}{cc\|cc} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 2 \end{array}\right]$

Table 2.7: Large orbits of $2 \times 2 \times 2$ tensors over $\mathbb{F}_5$

rank	orbit size	canonical form
0	1	$\left[\begin{array}{cc\|cc} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$
1	864	$\left[\begin{array}{cc\|cc} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right]$
2	2880	$\left[\begin{array}{cc\|cc} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{array}\right]$
2	2880	$\left[\begin{array}{cc\|cc} 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{array}\right]$
2	2880	$\left[\begin{array}{cc\|cc} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{array}\right]$
2	216000	$\left[\begin{array}{cc\|cc} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{array}\right]$
3	69120	$\left[\begin{array}{cc\|cc} 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{array}\right]$
3	96000	$\left[\begin{array}{cc\|cc} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 2 \end{array}\right]$

Table 2.8: Small orbits of $2 \times 2 \times 2$ tensors over $\mathbb{F}_5$

Chapter 3

On $2 \times 2 \times 2 \times 2$ Tensors

In the previous chapter we displayed a $2 \times 2 \times 2$ tensor in terms of its two 2×2 frontal slices: $\boldsymbol{X}_{\bullet\bullet 1}$ and $\boldsymbol{X}_{\bullet\bullet 2}$. We represent a $2 \times 2 \times 2 \times 2$ tensor $\boldsymbol{X} = [x_{ijkl}]$ with $i, j, k, l \in \{1, 2\}$, in terms of its two $2 \times 2 \times 2$ frontal slices: $\boldsymbol{X}_{\bullet\bullet\bullet 1}$ and $\boldsymbol{X}_{\bullet\bullet\bullet 2}$. As usual, the indices i, j, k refer to the row, column, and 2×2 slice, respectively. The last index refers to the top or bottom $2 \times 2 \times 2$ slice:

$$\boldsymbol{X} = \left[\begin{array}{c} \boldsymbol{X}_{\bullet\bullet\bullet 1} \\ \hline \boldsymbol{X}_{\bullet\bullet\bullet 2} \end{array}\right] = \left[\begin{array}{c|c} \boldsymbol{X}_{\bullet\bullet 11} & \boldsymbol{X}_{\bullet\bullet 21} \\ \hline \boldsymbol{X}_{\bullet\bullet 12} & \boldsymbol{X}_{\bullet\bullet 22} \end{array}\right] = \left[\begin{array}{cc|cc} x_{1111} & x_{1211} & x_{1121} & x_{1221} \\ x_{2111} & x_{2211} & x_{2121} & x_{2221} \\ \hline x_{1112} & x_{1212} & x_{1122} & x_{1222} \\ x_{2112} & x_{2212} & x_{2122} & x_{2222} \end{array}\right]$$

3.1 Rank

3.1.1 Outer Product Rank

Last chapter, we reproved that the maximum rank of a $2 \times 2 \times 2$ tensor is 3. As a corollary, we gave a formula for the upper bound of the rank of a tensor of format $2 \times \cdots \times 2$ (with $n \geq 3$ factors): the rank is less than or equal to $3 \cdot 2^{n-3}, n \geq 3$. In particular, for $n = 4$, the rank is bounded by 6. When we compute the canonical forms of these tensors over the finite fields $\mathbb{F}_2$ and $\mathbb{F}_3$, we will see that the maximum ranks are 6 and 5, respectively.

3.2 Polynomial Invariants of $2 \times 2 \times 2 \times 2$ Tensors

Huggins et al. [21] determined that the hyperdeterminant (as defined in [13] [14]) of a $2 \times 2 \times 2 \times 2$ tensor is a polynomial of degree 24 in 16 unknowns which has 2,894,276 terms. The methods used are difficult so we do not go into detail here. In [26], Thibon and Luque construct a set of polynomial invariants which allow them to obtain a closed formula for the hyperdeterminant of a $2 \times 2 \times 2 \times 2$ tensor. We will summarize their results below. This section has applications in Quantum Information Theory [26]. Let $V = \mathbb{C}^2$, $\mathcal{H} = V^{\otimes 4}$ and $G = \mathrm{SL}_{2,2,2,2}(\mathbb{C})$. The authors refer to the following as states:

$$\Phi = \sum_{i,j,k,l=0}^{1} \boldsymbol{X}_{ijkl} e_i \otimes e_j \otimes e_k \otimes e_l$$

For convenience, we have digressed from our usual convention of indexing i, j, k, l from $\{1, 2\}$ and are instead indexing from $\{0, 1\}$. To reduce the size of the expressions encountered, we will set the elements $\boldsymbol{X}_{ijkl} = x_r$, for $r = 0, \ldots, 15$, where r is the integer whose binary expression is $ijkl$, that is, $r = 8i + 4j + 2k + l$. For example, x_{15} is the last entry of the tensor: x_{1111} (or by our usual convention, x_{2222}). They are the coefficients of a quadrilinear form

$$\boldsymbol{X}(u, v, w, t) = \sum_{i,j,k,l=0}^{1} X_{ijkl} u_i v_j w_k t_l$$

on $V \times V \times V \times V$. This form allows us to compute Cayley's hyperdeterminant of 1843:

$$H = x_0 x_{15} - x_1 x_{14} - x_2 x_{13} + x_3 x_{12} - x_4 x_{11} + x_5 x_{10} + x_6 x_9 - x_7 x_8.$$

There are two linearly independent invariants of degree 4, given by any two of the three determinants, which can be formed by interpreting $\boldsymbol{X}$ as a linear map $\mathbb{C}^4 \to \mathbb{C}^4$:

$$L = \begin{vmatrix} x_0 & x_4 & x_8 & x_{12} \\ x_1 & x_5 & x_9 & x_{13} \\ x_2 & x_6 & x_{10} & x_{14} \\ x_3 & x_7 & x_{11} & x_{15} \end{vmatrix}, \quad M = \begin{vmatrix} x_0 & x_8 & x_2 & x_{10} \\ x_1 & x_9 & x_3 & x_{11} \\ x_4 & x_{12} & x_6 & x_{14} \\ x_5 & x_{13} & x_7 & x_{15} \end{vmatrix}, \quad N = \begin{vmatrix} x_0 & x_1 & x_8 & x_9 \\ x_2 & x_3 & x_{10} & x_{11} \\ x_4 & x_5 & x_{12} & x_{13} \\ x_6 & x_7 & x_{14} & x_{15} \end{vmatrix}.$$

Up to an appropriate choice of signs these three invariants sum to zero. Also, any pair of L, M, N are linearly independent, and H^2 cannot be expressed as a linear combination of these polynomials. $\boldsymbol{X}$ has six degree 2 biquadratic forms in all possible pairs of variables. These are constructed by taking the determinants of order 2 of the partial derivatives of $\boldsymbol{X}$ with respect to the complementary pair of variables. Let us explain this in more detail by demonstration. First, we expand the quadrilinear form

$$\begin{aligned} \boldsymbol{X}(u, v, w, t) = \sum_{i,j,k,l=0}^{1} X_{ijkl} u_i v_j w_k t_l \\ &= X_{0000} u_0 v_0 w_0 t_0 + X_{0001} u_0 v_0 w_0 t_1 + X_{0010} u_0 v_0 w_1 t_0 + X_{0011} u_0 v_0 w_1 t_1 \\ &+ X_{0100} u_0 v_1 w_0 t_0 + X_{0101} u_0 v_1 w_0 t_1 + X_{0110} u_0 v_1 w_1 t_0 + X_{0111} u_0 v_1 w_1 t_1 \\ &+ X_{1000} u_1 v_0 w_0 t_0 + X_{1001} u_1 v_0 w_0 t_1 + X_{1010} u_1 v_0 w_1 t_0 + X_{1011} u_1 v_0 w_1 t_1 \\ &+ X_{1100} u_1 v_1 w_0 t_0 + X_{1101} u_1 v_1 w_0 t_1 + X_{1110} u_1 v_1 w_1 t_0 + X_{1111} u_1 v_1 w_1 t_1 \end{aligned}$$

Now we will compute $\frac{\partial \boldsymbol{X}}{\partial w_i}$, for $i = 0, 1$:

$$\begin{aligned} \frac{\partial \boldsymbol{X}}{\partial w_0} &= X_{0000} u_0 v_0 t_0 + X_{0001} u_0 v_0 t_1 + X_{0100} u_0 v_1 t_0 + X_{0101} u_0 v_1 t_1 \\ &+ X_{1000} u_1 v_0 t_0 + X_{1001} u_1 v_0 t_1 + X_{1100} u_1 v_1 t_0 + X_{1101} u_1 v_1 t_1 \end{aligned}$$

$$\begin{aligned} \frac{\partial \boldsymbol{X}}{\partial w_1} &= X_{0010} u_0 v_0 t_0 + X_{0011} u_0 v_0 t_1 + X_{0110} u_0 v_1 t_0 + X_{0111} u_0 v_1 t_1 \\ &+ X_{1010} u_1 v_0 t_0 + X_{1011} u_1 v_0 t_1 + X_{1110} u_1 v_1 t_0 + X_{1111} u_1 v_1 t_1 \end{aligned}$$

Then we compute the second partial derivative of each equation above with respect to t_i for $i = 0, 1$:

$$\frac{\partial^2 \boldsymbol{X}}{\partial w_0 \partial t_0} = X_{0000}u_0v_0 + X_{0100}u_0v_1 + X_{1000}u_1v_0 + X_{1100}u_1v_1$$

$$\frac{\partial^2 \boldsymbol{X}}{\partial w_0 \partial t_1} = X_{0001}u_0v_0 + X_{0101}u_0v_1 + X_{1001}u_1v_0 + X_{1101}u_1v_1$$

$$\frac{\partial^2 \boldsymbol{X}}{\partial w_1 \partial t_0} = X_{0010}u_0v_0 + X_{0110}u_0v_1 + X_{1010}u_1v_0 + X_{1110}u_1v_1$$

$$\frac{\partial^2 \boldsymbol{X}}{\partial w_1 \partial t_1} = X_{0011}u_0v_0 + X_{0111}u_0v_1 + X_{1011}u_1v_0 + X_{1111}u_1v_1$$

We can write these four equations in a 2×2 matrix, where the (i, j)th entry is the second partial derivative with respect to w_i, t_j, for $i, j \in \{0, 1\}$. We'll call the determinant of this matrix b_{uv},

$$b_{uv} = \det\left(\frac{\partial^2 \boldsymbol{X}}{\partial w_i \partial t_j}\right).$$

The polynomial b_{uv} has 32 terms, which we will not produce here. We combine like terms with respect to the nine variables

$$u_0^2v_0^2,\quad u_0^2v_0v_1,\quad u_0^2v_1^2,\quad u_0u_1v_0^2,\quad u_0u_1v_0v_1,\quad u_0u_1v_1^2,\quad u_1^2v_0^2,\quad u_1^2v_0v_1,\quad u_1^2v_1^2,$$

and write the coefficients of these variables in a 3×3 matrix in the following way:

$$B_{uv} = \begin{bmatrix} u_0^2v_0^2 & u_0^2v_0v_1 & u_0^2v_1^2 \\ u_0u_1v_0^2 & u_0u_1v_0v_1 & u_0u_1v_1^2 \\ u_1^2v_0^2 & u_1^2v_0v_1 & u_1^2v_1^2, \end{bmatrix}$$

Finally, $D_{uv} = \det(B_{uv})$ gives a sextic invariant of $\boldsymbol{X}$ with 144 terms. This process yields six such invariants in total:

$$D_{uv},\qquad D_{uw},\qquad D_{ut},\qquad D_{wt},\qquad D_{vt},\qquad D_{vw}.$$

These determinants are degree-6 invariants of $\boldsymbol{X}$, and since this space is four dimensional, these determinants must be linearly dependent [26]. We have that D_{uv}, D_{uw}, and D_{ut} are linearly independent and hence

$$D_{uv} = D_{wt}, \qquad D_{uw} = D_{vt}, \qquad D_{ut} = D_{vw}.$$

And we have the dependence

$$HL = D_{uw} - D_{ut}, \quad HM = D_{ut} - D_{uv}, \quad HN = D_{uv} - D_{uw}$$

Thibon and Luque do not go into detail, and conclude that any of the D_{qs} can be taken as the generator of degree 6 [26].

The hyperdeterminant is of degree 24 and admits an expression in terms of the fundamental invariants. To find the explicitation of the hyperdeterminant, we need the invariants b_{qs}. Take, for example, b_{ut}. We can consider $\boldsymbol{X}$ as a trilinear form T in u, v, w, where t is a parameter. The hyperdeterminant $\mathrm{Det}(T)$ is homogeneous of degree 4 in its coefficients, which are themselves linear forms in the variables t_0, t_1. Thus, $R(t_0, t_1) = \mathrm{Det}(T)$ is a binary quartic in the components t_0, t_1 of t, and we can form its discriminant Δ, which will be an invariant of $\boldsymbol{X}$. In this case $\Delta = \mathrm{Det}(\boldsymbol{X})$. The hyperdeterminant of a $2 \times 2 \times 2 \times 2$ tensor takes the form of a homogeneous quartic in two variables t_0 and t_1:

$$R(t_0, t_1) = c_0 t_0^4 + 4c_1 t_0^3 t_1 + 6c_2 t_0^2 t_1^2 + 4t_0 t_1^3 + c_4 t_1^4,$$

where the c_i for $i = 1, 2, 3, 4$ are degree 4 polynomials in the components of the tensor [15]. The discriminant of this quartic polynomial is a degree 6 expression in the coefficients given by

$$\Delta = S^3 - 27T^2$$

where the invariants S and T are given by

$$S = c_0 c_4 - 4c_1 c_3 + 3c_2^2, \quad T = c_0 c_2 c_4 - c_0 c_3^2 + 2c_1 c_2 c_3 - c_1^2 c_4 - c_2^3,$$

Since the invariants S and T of R are invariants of $\boldsymbol{X}$, the problem of expressing $\mathrm{Det}(\boldsymbol{X})$ in terms of the fundamental invariants of $\boldsymbol{X}$ is reduced to the problem of finding the expressions of S and T. Thibon and Luque omit the details and conclude that a computer algebra system gives the following results:

$$S = \frac{1}{12}H^4 - \frac{2}{3}H^2L + \frac{2}{3}H^2M - 2HD_{ut} + \frac{4}{3}(L^2 + LM + M^2)$$

$$T = \frac{1}{216}H^6 - \frac{1}{18}H^4(L-M) - \frac{1}{6}H^3D_{ut} + \frac{1}{9}H^2(2L^2 - LM + 2M^2)$$

$$+ \frac{2}{3}H(L-M)D_{ut} - \frac{8}{27}(L^3 - M^3) - \frac{4}{9}LM(L-M) + D_{ut}^2$$

Setting $D = D_{ut}$, $U = H^2 - 4(L-M)$, and $V = 12(HD - 2LM)$, the expressions can be written as

$$S = \frac{1}{12}\left(U^2 - 2V\right), \quad T = \frac{1}{216}\left(U^3 - 3UV + 216D^2\right).$$

3.3 Canonical Forms over Finite Fields

Recall that for $p = 5$ the Maple procedure used to determine the ranks and canonical forms of $2\times2\times2$ tensors was not very efficient. This remains true for the generalization to $2\times2\times2\times2$ tensors. Indeed, a large amount of computer time and memory is used over $\mathbb{F}_2$, and the original procedure breaks down completely for $\mathbb{F}_3$. For this case, we only consider the action of the large symmetry group $\mathrm{GL}_{2,2,2,2}(\mathbb{F}_p) \rtimes S_4$. At this point we also mention the difficulty of determining the canonical forms of $2\times2\times2\times2$ tensors over $\mathbb{F} = \mathbb{R}$ and $\mathbb{F} = \mathbb{C}$. In general, a finite classification for tensors of arbitrary size and order is not possible because the dimensions of $\mathbb{F}^{d_1\times\cdots\times d_k}$ exceeds the dimension of $\mathrm{GL}_{d_1,\ldots,d_k}(\mathbb{F})$ when $d_1\cdots d_k > d_1^2 + \cdots + d_k^2$ [9]. Consequently, any explicit classification must include continuous parameters (entries that must satisfy certain conditions).

Over finite fields the general linear group $\mathrm{GL}_2(\mathbb{F}_p)$ is generated by two matrices [37]. For $p = 2$ the matrices are

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \qquad \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

For $p = 3$, the group is generated by

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$$

and either one of

$$\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \qquad \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}.$$

The symmetric group S_4 is generated by the transposition (12) and the 4-cycle (1234).

Once again we will use the spinning algorithm to compute the action of the large symmetry group. We do not consider the small symmetry group because the number of orbits is too large to tabulate. Just as we did for $2 \times 2 \times 2$ tensors over $\mathbb{F}_5$, we will use signal arrays (with encoding and decoding) to compute the orbits under the action of the symmetry group, then determine the rank corresponding to each orbit. There are $p^{16} - 1$ nonzero arrays of size $2 \times 2 \times 2 \times 2$ over $\mathbb{F}_p$. We create a signal array with this number of entries, allocate one byte for each entry, and initialize every entry to 0.

3.3.1 Canonical Forms over $\mathbb{F}_2$

Consider the case $p = 2$. The set of $2 \times 2 \times 2 \times 2$ tensors with entries in $\mathbb{F}_2$ contains 65536 elements, and the maximum rank is 6. The number of tensors of each rank and the approximate percentages are summarized below.

rank	0	1	2	3	4	5	6
number	1	81	2268	21744	37530	3888	24
$\approx$ %	0.0015	0.1236	3.4607	33.1787	57.2662	5.9326	0.0366

For the large symmetry group $GL_{2,2,2,2}(\mathbb{F}_2) \rtimes S_4$, there are 30 orbits. The ranks, orbit sizes, and canonical forms are given in Table 3.1 and Table 3.2. (We write dot for zero to increase legibility.) Within each rank, the tensors are listed in lex order.

3.3.2 Canonical Forms over $\mathbb{F}_3$

For the case $p = 3$ the set of $2 \times 2 \times 2 \times 2$ tensors with entries in $\mathbb{F}_3$ contains 43046721 elements. This time we find that the maximum rank is 5, a perfect example of how properties of tensors like rank are field-dependent. The number of tensors of each rank and the approximate percentages are summarized below.

rank	0	1	2	3	4	5
number	1	512	101376	6397441	36018623	528768
$\approx$ %	0.000002	0.001189	0.235502	14.861620	83.673328	1.228358

For the large symmetry group $GL_{2,2,2,2}(\mathbb{F}_3) \rtimes S_4$, there are 49 orbits. The ranks, orbit sizes, and canonical forms are given in Table 3.3, Table 3.4 and Table 3.4.

#	rank	orbit size	canonical form (flattened)															
1	0	1	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
2	1	81	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	1
3	2	324	·	·	·	·	·	·	·	·	·	·	·	·	·	1	1	·
4	2	1296	·	·	·	·	·	·	·	·	·	·	·	1	1	·	·	·
5	2	648	·	·	·	·	·	·	·	1	1	·	·	·	·	·	·	·
6	3	648	·	·	·	·	·	·	·	·	·	·	·	1	·	1	1	·
7	3	144	·	·	·	·	·	·	·	·	·	1	1	·	1	·	1	1
8	3	3888	·	·	·	·	·	·	·	1	·	·	·	1	1	·	·	·
9	3	2592	·	·	·	·	·	·	·	1	·	·	1	·	1	1	·	·
10	3	2592	·	·	·	·	·	·	·	1	·	1	1	·	1	·	1	·
11	3	3888	·	·	·	·	·	·	·	1	1	·	·	·	·	·	1	·
12	3	7776	·	·	·	·	·	·	·	1	1	·	·	·	·	1	1	·
13	3	216	·	·	·	1	1	·	·	·	1	1	1	·	1	1	1	1
14	4	162	·	·	·	·	·	·	·	1	·	·	·	1	·	1	1	·
15	4	2592	·	·	·	·	·	·	·	1	·	1	1	·	1	·	·	·
16	4	5184	·	·	·	·	·	·	·	1	1	·	·	1	·	1	1	·
17	4	108	·	·	·	·	·	1	1	·	·	1	1	·	·	·	·	·
18	4	972	·	·	·	·	·	1	1	·	·	1	1	·	·	·	·	1
19	4	1944	·	·	·	·	·	1	1	·	·	1	1	·	·	·	1	·
20	4	1944	·	·	·	·	·	1	1	·	·	1	1	1	·	·	1	·

Table 3.1: Large orbits of $2 \times 2 \times 2 \times 2$ tensors over $\mathbb{F}_2$

#	rank	orbit size	canonical form (flattened)															
21	4	7776	·	·	·	·	·	1	1	·	·	1	1	1	1	·	·	·
22	4	1296	·	·	·	·	·	1	1	·	1	·	1	1	·	·	·	·
23	4	7776	·	·	·	·	·	1	1	·	1	·	1	1	·	·	·	1
24	4	3888	·	·	·	1	·	1	1	·	1	·	·	·	·	·	1	1
25	4	3888	·	·	·	1	·	1	1	·	1	·	·	·	1	·	1	1
26	5	648	·	·	·	·	·	1	1	·	·	1	1	·	1	·	1	1
27	5	648	·	·	·	1	·	1	1	·	·	1	1	·	1	·	·	·
28	5	1296	·	·	·	1	·	1	1	·	·	1	1	·	1	·	·	1
29	5	1296	·	·	·	1	·	1	1	·	1	·	·	·	·	·	·	1
30	6	24	·	1	1	·	1	·	1	1	1	·	1	1	1	1	·	1

Table 3.2: Large orbits of $2 \times 2 \times 2 \times 2$ tensors over $\mathbb{F}_2$ (continued)

#	rank	orbit size	canonical form (flattened)															
1	0	1	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
2	1	512	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	1
3	2	4608	·	·	·	·	·	·	·	·	·	·	·	·	·	1	1	·
4	2	55296	·	·	·	·	·	·	·	·	·	·	·	1	1	·	·	·
5	2	41472	·	·	·	·	·	·	·	1	1	·	·	·	·	·	·	·
6	3	24576	·	·	·	·	·	·	·	·	·	·	·	1	·	1	1	·
7	3	13824	·	·	·	·	·	·	·	·	·	1	1	·	1	·	·	2
8	3	331776	·	·	·	·	·	·	·	1	·	·	·	1	1	·	·	·
9	3	442368	·	·	·	·	·	·	·	1	·	·	1	·	1	1	·	·
10	3	1327104	·	·	·	·	·	·	·	1	·	1	1	·	1	·	·	·
11	3	497664	·	·	·	·	·	·	·	1	1	·	·	·	·	·	1	·
12	3	1990656	·	·	·	·	·	·	·	1	1	·	·	·	·	1	1	·
13	3	1327104	·	·	·	·	·	·	·	1	1	·	·	1	·	1	1	·
14	3	110592	·	·	·	1	·	1	1	·	·	1	1	·	1	·	·	1
15	3	331776	·	·	·	1	1	·	·	·	1	1	1	·	1	1	1	1
16	4	12288	·	·	·	·	·	·	·	1	·	·	·	1	·	1	1	·
17	4	1327104	·	·	·	·	·	·	·	1	1	·	·	1	·	1	2	·
18	4	3456	·	·	·	·	·	1	1	·	·	1	1	·	·	·	·	·
19	4	110592	·	·	·	·	·	1	1	·	·	1	1	·	·	·	·	1
20	4	331776	·	·	·	·	·	1	1	·	·	1	1	·	·	·	1	·

Table 3.3: Large orbits of $2 \times 2 \times 2 \times 2$ tensors over $\mathbb{F}_3$

#	rank	orbit size	canonical form (flattened)															
21	4	165888	·	·	·	·	·	1	1	·	·	1	1	·	1	·	·	2
22	4	663552	·	·	·	·	·	1	1	·	·	1	1	1	·	·	1	·
23	4	3981312	·	·	·	·	·	1	1	·	·	1	1	1	1	·	·	·
24	4	124416	·	·	·	·	·	1	1	·	·	1	2	·	·	·	·	·
25	4	995328	·	·	·	·	·	1	1	·	·	1	2	·	·	·	·	1
26	4	1990656	·	·	·	·	·	1	1	·	·	1	2	·	1	·	·	1
27	4	124416	·	·	·	·	·	1	1	·	1	·	·	2	·	·	·	·
28	4	995328	·	·	·	·	·	1	1	·	1	·	·	2	·	·	·	1
29	4	995328	·	·	·	·	·	1	1	·	1	·	·	2	·	·	1	1
30	4	248832	·	·	·	·	·	1	1	·	1	·	1	1	·	·	·	·
31	4	3981312	·	·	·	·	·	1	1	·	1	·	1	1	·	·	·	1
32	4	27648	·	·	·	1	·	1	1	·	·	1	1	·	1	·	·	·
33	4	110592	·	·	·	1	·	1	1	·	·	1	1	·	1	·	·	2
34	4	663552	·	·	·	1	·	1	1	·	·	1	1	·	1	·	1	·
35	4	1990656	·	·	·	1	·	1	1	·	·	1	1	·	2	·	·	1
36	4	1990656	·	·	·	1	·	1	1	·	1	·	·	·	·	·	1	1
37	4	1990656	·	·	·	1	·	1	1	·	1	·	·	·	·	·	2	1
38	4	663552	·	·	·	1	·	1	1	·	1	·	·	·	·	1	2	·
39	4	1327104	·	·	·	1	·	1	1	·	1	·	·	·	·	1	2	1
40	4	331776	·	·	·	1	·	1	1	·	1	·	·	·	1	·	·	1

Table 3.4: Large orbits of $2 \times 2 \times 2 \times 2$ tensors over $\mathbb{F}_3$ (continued)

#	rank	orbit size	canonical form (flattened)															
41	4	663552	·	·	·	1	·	1	1	·	1	·	·	·	1	·	·	2
42	4	1990656	·	·	·	1	·	1	1	·	1	·	·	·	1	·	1	1
43	4	1990656	·	·	·	1	·	1	1	·	1	·	·	·	1	·	2	1
44	4	1990656	·	·	·	1	·	1	1	·	1	·	·	·	1	·	2	2
45	4	3981312	·	·	·	1	·	1	1	·	1	·	·	·	1	1	2	·
46	4	248832	·	·	·	1	1	·	·	·	1	1	1	·	1	1	1	2
47	4	5184	·	1	1	·	1	·	·	2	1	·	·	2	·	2	2	·
48	5	497664	·	·	·	1	·	1	1	·	·	1	1	·	2	·	·	·
49	5	31104	·	1	1	·	1	·	·	2	1	·	·	2	1	1	1	2

Table 3.5: Large orbits of $2 \times 2 \times 2 \times 2$ tensors over $\mathbb{F}_3$ (continued)

Chapter 4

Conclusion

This chapter serves to provide a summary of our original contributions, as well as other areas we investigated based on our results, and a brief discussion on tensor applications.

4.1 Our Results

Our primary focus was understanding canonical forms of $2 \times 2 \times 2$ and $2 \times 2 \times 2 \times 2$ tensors. For $2 \times 2 \times 2$ tensors, there were 8 over $\mathbb{R}$, and 7 over $\mathbb{C}$, both were results that have been independently rediscovered many times in the literature. We provided original results by extending this classification to prime fields, $\mathbb{F}_p$ for $p = 2, 3, 5$. We used computer algebra to first compute the tensors in each rank, then we considered the action of the small symmetry group (the direct product of the general linear groups) and the large symmetry group (the semi-direct product of the general linear groups with the symmetric group), which decomposed the set of tensors into a disjoint union of orbits. The tensors in each orbit were equivalent under the group action and thus the canonical form was taken to be the minimal element of each orbit. We were not able to consider the same problem for larger primes because our computer did not have sufficient memory. Moreover, when we considered the same problem for the larger tensor format $2\times2\times2\times2$, our algorithms were not efficient. By employing more sophisticated programming techniques we were able to classify $2 \times 2 \times 2 \times 2$ tensors over $\mathbb{F}_p$, for $p = 2, 3$, using only the large symmetry group (the action of the small symmetry group provided too many orbits). Even with more efficient techniques, we did not have sufficient memory to consider larger prime fields.

4.2 Canonical Forms of Symmetric $2 \times 2 \times 2$ Tensors

In this section, we discuss how our results of canonical forms over prime fields can be extended to $2 \times 2 \times 2$ symmetric tensors. This problem has already been considered by Weinberg [38] over $\mathbb{R}$ and $\mathbb{C}$. We will summarize these results. To our knowledge, no one has obtained a classification of symmetric tensors over finite fields. We show examples of tensors that cannot be decomposed as a sum of symmetric simple tensors over $\mathbb{F}_2$. We also define a new notion of tensor rank called symmetric tensor rank and show that the maximum symmetric tensor rank is larger over some prime fields $\mathbb{F}_p$ than over $\mathbb{R}$ and $\mathbb{C}$.

We begin with some terminology and theorems. First recall Definition 2.9: an order-k tensor $[x_{i_1 \dots i_k}] \in \mathbb{F}^{n \times \cdots \times n}$ is called **symmetric** if

$$x_{i_{\pi(1)} \cdots i_{\pi(k)}} = x_{i_1 \dots i_k}, \qquad i_1, \dots, i_k \in \{1, \dots, n\},$$

for all permutations $\pi \in S_k$.

Proposition 4.1. *Proposition 3.7 [7]*
Let $\mathbf{X} = [x_{i_1 \dots i_k}] \in \mathbb{F}^{n \times \cdots \times n}$ be an order-k tensor. Then $\pi(\mathbf{X}) = \mathbf{X}$ for all permutations $\pi \in S_k$ if and only if $x_{i_{\pi(1)} \cdots i_{\pi(k)}} = x_{i_1 \dots i_k}$, where $i_1, \dots, i_k \in \{1, \dots, n\}$ for all permutations $\pi \in S_k$.

Consider the $2 \times 2 \times 2$ symmetric tensor

$$\boldsymbol{X} = \left[\begin{array}{cc|cc} x_{111} & x_{121} & x_{112} & x_{122} \\ x_{211} & x_{221} & x_{212} & x_{222} \end{array}\right].$$

Then $x_{112} = x_{121} = x_{211}$ and $x_{122} = x_{212} = x_{221}$, and so we can label $\boldsymbol{X}$ as:

$$\boldsymbol{X} = \left[\begin{array}{cc|cc} a & b & b & c \\ b & c & c & d \end{array}\right].$$

Recall that a simple order-k tensor is the outer product of k non-zero vectors. A simple order-k symmetric tensor is the outer product of a non-zero vector k times. From this we can define a new notion of tensor rank.

Definition 4.2. [7] The **symmetric tensor rank** of a symmetric tensor $\boldsymbol{X}$ is

$$\text{rank}_S(\boldsymbol{X}) = \min\left\{ s \in \mathbb{N} \mid \boldsymbol{X} = \sum_{i=1}^{s} u_i \otimes \cdots \otimes u_i \right\},$$

where the u_i are non-zero for all i.

Given a symmetric tensor $\boldsymbol{X}$, we can decompose it in the usual way (as a sum of simple tensors), or as a sum of symmetric simple tensors. Since a symmetric decomposition is constrained, we have the inequality $\text{rank}(\boldsymbol{X}) \leq \text{rank}_S(\boldsymbol{X})$ for all $\boldsymbol{X}$. See Comon et al. [7] for a discussion on typical and generic ranks of symmetric tensors.

4.2.1 Canonical Forms over $\mathbb{R}$ and $\mathbb{C}$

Weinberg [38] computes the canonical forms of symmetric $2 \times 2 \times 2$ tensors over $\mathbb{R}$ and $\mathbb{C}$, which we will display without proof. We summarize the results over $\mathbb{R}$ in Table 4.1. Over $\mathbb{C}$, there are four canonical forms (the same ones as in the real case) after we exclude the tensor

$$\left[\begin{array}{cc|cc} 0 & 1 & 1 & -1 \\ 1 & -1 & -1 & 0 \end{array}\right].$$

That is, there is only one orbit in symmetric rank 3. The maximum symmetric rank is 3 over $\mathbb{R}$ and $\mathbb{C}$. We will see shortly that the maximum rank is larger over $\mathbb{F}_p$ for $p = 3, 7$. We also mention that Weinberg [38] lists the canonical forms of $2 \times 2 \times 2 \times 2$ symmetric tensors over $\mathbb{R}$ and $\mathbb{C}$, which we do not summarize here.

symmetric rank	canonical form
0	$\left[\begin{array}{cc\|cc} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$
1	$\left[\begin{array}{cc\|cc} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$
2	$\left[\begin{array}{cc\|cc} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right]$
3	$\left[\begin{array}{cc\|cc} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{array}\right]$
3	$\left[\begin{array}{cc\|cc} 0 & 1 & 1 & -1 \\ 1 & -1 & -1 & 0 \end{array}\right]$

Table 4.1: Canonical forms of $2 \times 2 \times 2$ symmetric tensors over $\mathbb{R}$

4.2.2 Canonical Forms over $\mathbb{F}_2$

We use computer algebra to consider $2 \times 2 \times 2$ symmetric tensors over prime fields $\mathbb{F}_p$. The number of symmetric tensors over $\mathbb{F}_p$ is equal to p^4. To see why this is, consider a general $2 \times 2 \times 2$ tensor

$$\boldsymbol{X} = \left[\begin{array}{cc|cc} a & b & b & c \\ b & c & c & d \end{array}\right].$$

We can place p different elements for entry $x_{111} = a$. We can also place p different elements for entry $x_{121} = b$. Once an entry is fixed for x_{121}, that same entry must be used for x_{211} and x_{112}, otherwise the tensor would not be symmetric. This argument is also true for the entries we labelled c. Finally, d can have p different choices. Thus we arrive at p^4.

Over $\mathbb{F}_2$, there are 256 tensors, 16 of which are symmetric. The set of rank-1 tensors is $\{a \otimes a \otimes a \mid a \in \mathbb{F}_2^2\}$ where a is a non-zero vector. There are three symmetric simple tensors,

$$\left[\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right], \quad \left[\begin{array}{cc|cc} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right], \quad \left[\begin{array}{cc|cc} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{array}\right].$$

We emphasize that not all 16 symmetric tensors can be written as a sum of symmetric simple tensors. The following 8 tensors do not have a symmetric decomposition:

$$\left[\begin{array}{cc|cc} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{array}\right], \quad \left[\begin{array}{cc|cc} 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{array}\right], \quad \left[\begin{array}{cc|cc} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{array}\right], \quad \left[\begin{array}{cc|cc} 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{array}\right],$$
$$\left[\begin{array}{cc|cc} 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{array}\right], \quad \left[\begin{array}{cc|cc} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{array}\right], \quad \left[\begin{array}{cc|cc} 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 \end{array}\right], \quad \left[\begin{array}{cc|cc} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 \end{array}\right].$$

Recall that the large symmetry group is the semi-direct product of general linear groups and the symmetric group, S_n. Since symmetric tensors are unchanged under the action of S_n, we only consider the small symmetry group. We can use our original computer program to compute orbits of the group action, but we need to make one

symmetric rank	orbit size	canonical form
0	1	$\left[\begin{array}{cc\|cc} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$
1	3	$\left[\begin{array}{cc\|cc} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right]$
2	3	$\left[\begin{array}{cc\|cc} 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{array}\right]$
3	1	$\left[\begin{array}{cc\|cc} 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 \end{array}\right]$

Table 4.2: Canonical forms of $2 \times 2 \times 2$ symmetric tensors over $\mathbb{F}_2$

modification: the same group element $g \in \mathrm{GL}_{2,2,2}(\mathbb{F}_2)$ must act on all three sides of the tensor. Otherwise, a symmetric tensor will not necessarily be symmetric after a transformation. Therefore, the orbit of $\boldsymbol{X}$ is

$$\mathcal{O}_{\boldsymbol{X}} := \{(g, g, g) \cdot \boldsymbol{X} \mid g \in \mathrm{GL}_{2,2,2}(\mathbb{F}_p)\}.$$

The maximum symmetric rank over $\mathbb{F}_2$ is 3. The symmetric ranks, orders of each orbit, and the minimal representatives of each orbit are given in Table 4.2.

4.2.3 Canonical Forms over $\mathbb{F}_3$

There are 6561 tensors over $\mathbb{F}_3$, 81 of which are symmetric. Unlike in the previous case, all 81 symmetric tensors can be generated as a sum of simple symmetric tensors. The maximum symmetric rank in this case is 4. The symmetric ranks, orders of each orbit, and the minimal representatives of each orbit are given in Table 4.3.

symmetric rank	orbit size	canonical form
0	1	$\left[\begin{array}{cc\|cc} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$
1	8	$\left[\begin{array}{cc\|cc} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right]$
2	24	$\left[\begin{array}{cc\|cc} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{array}\right]$
3	8	$\left[\begin{array}{cc\|cc} 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{array}\right]$
3	24	$\left[\begin{array}{cc\|cc} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 2 \end{array}\right]$
4	16	$\left[\begin{array}{cc\|cc} 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 \end{array}\right]$

Table 4.3: Canonical forms of $2 \times 2 \times 2$ symmetric tensors over $\mathbb{F}_3$

symmetric rank	orbit size	canonical form
0	1	$\left[\begin{array}{cc\|cc} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$
1	24	$\left[\begin{array}{cc\|cc} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right]$
2	240	$\left[\begin{array}{cc\|cc} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{array}\right]$
3	120	$\left[\begin{array}{cc\|cc} 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{array}\right]$
3	80	$\left[\begin{array}{cc\|cc} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 2 \end{array}\right]$
3	160	$\left[\begin{array}{cc\|cc} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 2 \end{array}\right]$

Table 4.4: Canonical forms of $2 \times 2 \times 2$ symmetric tensors over $\mathbb{F}_5$

4.2.4 Canonical Forms over $\mathbb{F}_5$

There are 390625 tensors over $\mathbb{F}_3$, 625 of which are symmetric. Again, all 625 symmetric tensors can be generated as a sum of simple symmetric tensors. The maximum symmetric rank in this case is 3. The symmetric ranks, orders of each orbit, and the minimal representatives of each orbit are given in Table 4.4.

4.2.5 Canonical Forms over $\mathbb{F}_7$

There are 5764801 tensors over $\mathbb{F}_7$, 2401 of which are symmetric. Since the number of symmetric tensors is low enough, our computer has enough memory to consider larger primes fields. Again, all 2401 symmetric tensors can be generated as a sum of simple symmetric tensors, which allows us to conjecture that $\mathbb{F}_2$ is a special case that did not allow every tensor to be written as a sum of symmetric simple tensors. The maximum symmetric rank in this case is 5, and we also see that the number of orbits has jumped to 14. The symmetric ranks, orders of each orbit, and the minimal representatives of each orbit are given in Table 4.5. We write the canonical form in flattened form to save space.

We believe classifying the orbits of symmetric tensors over larger prime fields $\mathbb{F}_p$ is still computationally feasible. Indeed, for $p = 11$ there are 14641 symmetric tensors, and for $p = 67$ there are 20151121, which is still less than half of the number of $2 \times 2 \times 2 \times 2$ tensors over $\mathbb{F}_2$ (43046721 tensors) that we considered in Chapter 3. We do not consider any more cases here.

symmetric rank	orbit size	canonical form (flattened)
0	1	0 0 0 0 0 0 0 0
1	16	0 0 0 0 0 0 0 1
2	16	0 0 0 0 0 0 0 2
2	112	0 1 1 0 1 0 0 2
3	16	0 0 0 0 0 0 0 3
3	112	0 1 1 0 1 0 0 1
3	336	0 1 1 0 1 0 0 3
3	224	1 0 0 0 0 0 0 2
4	336	0 0 0 1 0 1 1 0
4	112	0 1 1 0 1 0 0 6
4	336	0 1 1 0 1 0 0 6
4	224	1 0 0 0 0 0 0 3
4	224	1 0 0 1 0 1 1 2
5	336	0 1 1 0 1 0 0 5

Table 4.5: Canonical forms of $2 \times 2 \times 2$ symmetric tensors over $\mathbb{F}_7$

4.3 Applications

In this section we briefly discuss some applications of tensors. A recent book (2009) by Cichocki et al. [6] provides a wide survey of models and algorithmic aspects of various types of matrix and tensor decompositions. They also discuss applications in data clustering, text mining, economics, email surveillance, musical instrument classification, face recognition, handwritten digit recognition, gene expression classification, and many others.

We will briefly mention two applications that we find interesting. In Game Theory, a branch of economics, one learns about $n \times n$ matrix games where the entries represent payoffs in different scenarios for a certain number of players. Gonzalez-Alcon et al. [17] have recently introduced a class of $2 \times 2 \times 2$ matrix games and provided a classification of these games.

Schulman [32] has discussed how tensor problems like maximum rank are used in cryptographic algorithms and security protocols.

Bibliography

[1] G. Bergqvist: Exact probabilities for typical ranks of $2 \times 2 \times 2$ and $3 \times 3 \times 2$ tensors. *Linear Algebra and its Applications*, to appear.

[2] F. Block, P. Schmid: Rank and orbits of $2 \times 2 \times 2$ matrices. Preprint. http://homepages.warwick.ac.uk/staff/F.S.Block/final_Quantum.pdf.

[3] M. R. Bremner: Unpublished lecture notes on $2 \times 2 \times 2$ tensors based on ten Berge [36]. Nonassociative Algebra Seminar, Department of Mathematics and Statistics, University of Saskatchewan, February 2011.

[4] M. R. Bremner, S. G. Stavrou: Canonical forms of $2{\times}2{\times}2$ and $2{\times}2{\times}2{\times}2$ arrays over $\mathbb{F}_2$ and $\mathbb{F}_3$. *Linear and Multilinear Algebra* (to appear).

[5] A. Cayley: On the theory of linear transformations. *Cambridge Mathematical Journal* 4 (1845) 193–209.

[6] A. Cichocki, R. Zdunek, A. H. Phan, S.-I. Amari: *Nonnegative matrix and tensor factorizations* John Wiley & Sons. Inc., 2009.

[7] P. Comon, G. Golub, L.-H. Lim, B. Mourrain: Symmetric tensors and symmetric tensor rank. *SIAM Journal on Matrix Analysis and Applications* 30 (2008) 1084 – 1127.

[8] P. Comon, J.M.F. ten Berge, L. De Lathauwer, J. Castaing: Generic and typical ranks of multi-way arrays. *Linear Algebra and its Applications* 430 (2009) 2997–3007.

[9] V. de Silva, L.-H. Lim, : Tensor rank and the ill-posedness of the best low-rank approximation problem. *SIAM Journal on Matrix Analysis and Applications* 30, 3 (2008) 1084 - 1127.

[10] D. Z. Djokovic, A. Osterloh: On polynomial invariants of several qubits. *Journal of Mathematical Physics* 50 (2009) 033509, 23 pages.

[11] R. Ehrenborg: Canonical forms of two by two by two matrices. *Journal of Algebra* 213:1 (1999) 195–224.

[12] S. Friedland: On the generic and typical ranks of 3-tensors. *Linear Algebra and its Applications* 436 (2012) 478 - 497.

[13] I.M. Gelfand, M.M. Kapranov, A.V. Zelevinsky: *Discriminants, resultants and multidimensional determinants.* Birkhauser Boston, Boston, MA, 1994.

[14] I.M. Gelfand, M.M. Kapranov, A.V. Zelevinsky: Hyperdeterminants. *Advances in Mathematics* 96 (1992) 226 - 263.

[15] P. Gibbs: Elliptic curves and hyperdeterminants in quantum gravity. *Prespacetime Journal* 1, 8 (2010) 1218 - 1224.

[16] P. Gibbs: Schlafli's hyperdeterminant. http://hyperdeterminant.wordpress.com/2008/10/11/schlaflis-hyperdeterminant/

[17] C. Gonzalez-Alcon, P. Borm, R. Hendrickx: Nash equilibria in $2 \times 2 \times 2$ trimatrix games with identical anonymous best-replies. http://econpapers.repec.org/paper/dgrkubcen/2011138.htm

[18] J. Hastad: Tensor rank is NP-complete. *Journal of Algorithms* 11, 4 (1990) 644 - 654.

[19] C. Hillar, L.-H. Lim: Most tensor problems are NP hard. http://arxiv.org/abs/0911.1393.

[20] F. L. Hitchcock: The expression of a tensor or a polyadic as a sum of products. *J. Math. Phys.* 6 (1927) 164–189

[21] P. Huggins, B. Sturmfels, J. Yu, D. Yuster: The hyperdeterminant and triangulations of the 4-cube. *Mathematics of Computation* 77, 263 (2008) 1653 - 1679.

[22] T. Kolda, B. Bader: Tensor decompositions and applications. *SIAM Review* 51, 3 (2009) 455-500.

[23] E. Kreyszig: *Introductory functional analysis with applications*. John Wiley & Sons. Inc., 1978.

[24] J. B. Kruskal: Rank, decomposition, and uniqueness for 3-way and n-way arrays. *Multiway Data Analysis*. Elsevier Science, Amsterdam (1989) 7–18.

[25] C. Le Paige: Sur les formes trilinéaires. *Comptes rendus de l'Académie des Sciences* 92 (1881) 1103-1105.

[26] J.-G. Luque, J.-Y. Thibon: Polynomial invariants of four qubits. *Physical Review* A (3) 67 (2003), no.4, 042303, 5 pages.

[27] J.-G. Luque, J.-Y. Thibon: Algebraic invariants of five qubits. *Journal of Physics A: Mathematical and Theoretical* 39 (2006), no. 2. 371 - 377.

[28] C.D. Martin: The rank of a $2 \times 2 \times 2$ tensor. *Linear and Multilinear Algebra (to appear)* Published online 1 January 2011. DOI:10.1080/03081087.2010.538923

[29] R. Oldenburger: On canonical binary trilinear forms. *Bulletin of the American Mathematical Society* 38 (1932) 385–387.

[30] R. Oldenburger: Real canonical binary trilinear forms. *American Journal of Mathematics* 59 (1937) 427–435.

[31] A. Rovi: Analysis of 2 x 2 x 2 tensors. Master's thesis, 2010. http://www.maths.ox.ac.uk/system/files/Analysis_of_2_x_2_x_2_Tensors_0.pdf

[32] L. J. Schulman: Cryptography from tensor problems. *IACR Cryptology ePrint Archive* Vol. 2012 (2012) 244.

[33] E. Schwartz: Über binäre trilineare Formen. *Mathematische Zeitschrift* 12 (1922) 18–35.

[34] A. Stegeman, N. Sidiropoulos: On Kruskal's uniqueness condition for the candecomp/parafac decompositions. *Linear Algebra and its Applications* 420, (2007) 540 – 552.

[35] A. Stegeman, P. Comon: Subtracting a best rank-1 approximation may increase tensor rank. *Linear Algebra and its Applications* 433, (2010) 1276 - 1300.

[36] J. M. F. ten Berge: Kruskal's polynomial for $2 \times 2 \times 2$ arrays and a generalization to $2 \times n \times n$ arrays. *Psychometrika* 56, 4 (1991) 631 - 636.

[37] W.C. Waterhouse: Two generators for the general linear groups over finite fields. *Linear and Multilinear Algebra* 24, 4 (1989) 227 - 230.

[38] D. D. A. Weinberg: Canonical forms for symmetric tensors. *Linear Algebra and its Applications* 57 (1984) 271–282.