UNIVERSITAT DE BARCELONA
Facultat de Matemàtiques i Informàtica

GRAU DE MATEMÀTIQUES

Treball final de grau

SEARCH OF OPTIMAL LOW-RANK APPROXIMATIONS USING TENSOR NETWORKS

Autor: Aran Roig

Director: Dr. Nahuel Statuto
Realitzat a: Departament de Matemàtiques i Informàtica

Barcelona, June 5, 2025

Abstract

Tensor network structure search has been interesting research topic since the raise on complexity of deep learning models and quantum mechanics. This Bachelor Thesis main goal is to give an automated search of an optimal tensor network structure for representing a given tensor with some fixed error.

For these purpose we first give an introduction to some basic tensor algebra, we present tensor networks and tensor network states. Then we present algorithms that find the cores of a tensor network that better represent an objective tensor. We also present an algorithm for finding optimal tensor network structures that can guarantee that we will find the most optimal cores.

Finally, we prove that these algorithms converge under certain assumptions and then we perform some practical experiments that empirically proves that it is possible to find other more optimized low-rank decompositions of tensors without significant losses on performance and accuracy.

Resum

La recerca de l'estructura òptima de xarxes de tensors ha estat un tema d'interès des de l'augment en la complexitat dels models d'aprenentatge profund i de la mecànica quàntica. Aquest treball de final de grau té com a objectiu principal oferir una cerca automatitzada d'una estructura òptima de xarxa de tensors per representar un tensor donat amb un error fixat.

Per aconseguir aquest propòsit, primer oferim una introducció als tensors, a les xarxes tensorials. Aleshores presentem algoritmes que troben els nuclis de la xarxa tensorial que millor representen un tensor objectiu. També presentem un algoritme per trobar quina estructura de xarxa tensorial ens pot garantir que poguem trobar els nuclis més optims.

Finalment demostrem que aquests algoritmes convergeixen sota algunes suposicions i després fem uns quants experiments que empíricament mostren que és possible trobar altres descomposicions de baix rang de tensors sense que suposin una pèrdua significant en rendiment i precissió.

2020 Mathematics Subject Classification. 11G05, 11G10, 14G10

Agraïments

Vull agrair a ...

Contents

Chapter 1

Introduction

On the last decades, neural networks have emerged as one of the most influential topics inside the fields of artificial inteligence and machine learning. Neural networks are inspired on how the human brain works, they are made as a simplification about how our networks interact and in some way they try to emulate the way we think.

In the last years, neural networks have gained a lot of importance, positioning themselves at the center of impactful technological advances. Some models that use neural networks at its core are for example large language models (LLMs) which serve as advanced virtual agents, diffusers which generate images and other media. Other models based on neural networks have been used to make important advances in a lot of different fields such as in medicine [1], transportation [2], education [3], and others.

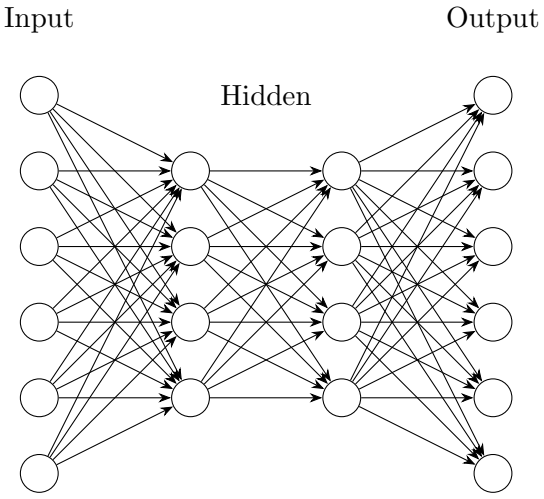


Figure 1.1: A representation of a fully connected neural network. Each dot is a neuron and each edge represents a connection. Source: own elaboration

Neural networks are modeled after how real neurons work, but in a simplified manner: each neural network contains a set of neurons which are split in different layers. In a fully connected neural network, each neuron has an output that is connected to each neuron of the next layer, and all neurons of a layer have as inputs the outputs of all the neurons of the previous layer. The first layer is called the input layer and the last layer is called the output layer. All the layers in between are called hidden layers.

As real neurons work, each connection will have a distinct role on when one neuron should fire or not. So, the outputs of each neuron will depend on its inputs. The output of each neuron in a neural network is modeled using usually a non-linear function $f : \mathbb{K}^N \to \mathbb{K}$, with N being the number of neurons on the previous layer.

Each neuron of the neural network also has assigned some weights w_{ji} to each input connection, with the idea that some connections will influence the firing of the network more than others. The key to making a neural network "learn" is to properly adjust these weights w_{ji} that are also called the *parameters* of the neural network.

So, more formally, the output of a neuron j in some layer will be given by the formula:

$$y_j(x) = f\left(\sum_{i=1}^{N} w_{ji}x_i + b_j\right)$$

Where x_i is the output of the neuron i of the previous layer, and b_j is some bias that is also a parameter of the neural network.

We can represent the weights of a fully connected neural network as a matrix, that we call the **weight matrix**. Since neural networks have been incresing in complexity during the recent years, these weight matrices can be very large and can contain a lot of parameters. So naturally there has been a rising interest on compressing weight matrices without sacrificing accuracy or performance.

The main goal of this thesis will be to compress the parameters of neural networks using tensor networks, a concept that originates from the study of many-body quantum systems [4] and recently has attracted significant interest on machine learning. We will also give another application by compressing images.

Tensor networks are a structure that is aimed to represent and efficiently manipulate large tensors by breaking them into smaller ones, called core tensors, which are connected into a specific pattern. Through the thesis we will explain how these smaller tensors are connected between them, but the general idea is that after contracting these connections, we get a representation of the original tensor.

So, what we aim to do is that given some tensor $T \in \mathbb{K}^{N_1 \times N_2 \times \cdots \times N_n}$ that will be our weight matrix reshaped onto a tensor, we will find a tensor network structure (TNS) that represents a good aproximation of T. This structure will minimize the relative error of the representation of T, the size of the TNS, since we are interested in compressing T to fewer parameters, and we will also care about the computational complexity of recovering T. We will also present the alternating least square method adapted on finding the cores of the best tensor network state. We will also find the cores of the tensor network using backpropagation, and we will compare both algorithms.

We will see that finding the best structure is an integer programming problem and we will describe the TnALE algorithm that finds a locally best structure using gradient-less optimization.

1.1 Thesis structure

The second chapter of the thesis will focus on the preliminaries about tensor algebra. We will introduce tensors from a mathematical standpoint presenting the tensor product space. We will see that the tensor product space forms vector space and then we will describe some operations between tensors.

After that, we will also describe some reshaping operations that can be applied to tensors to transform its orders i.e their number of dimensions. These reshaping operations will be useful on the following chapters since we will need them for applying the TN-ALS algorithm and also for the last chapter of applications, when we transform matrices into tensors and then apply our low-rank approximation that will use tensor networks.

On the third chapter we will present the basics of the Penrose Notation and we will give a formal definition to tensor networks (TNs) and tensor network states (TNS). We will give some common examples of tensor networks and after that we will define the notion of the rank of a tensor related to a TNS, called G-rank. We will see that the G-rank can be a lot smaller than the traditional tensor rank as it is discussed in [5], proving that there exists better compression methods than the canonical polyadic decomposition, which we will see on the same chapter.

On the fourth chapter we will describe the main algorithms of this thesis. First, we will give two algorithms that fixed a TN, finds a low-rank decomposition of smaller tensors that can represent our original tensor with a certain error. These algorithms are the tensor network alternating least squares (TN-ALE) and we will also solve this problem using backpropagation.

Then, we will unfix the tensor structure and we will present the TnALE algorithm which finds the most optimal tensor network structure for representing a tensor T by evaluating a loss function that depends on the structure itself. TnALE also saves doing the evaluation of some tensor network structures and lastly we will prove that TnALE converges to a certain structure and that we can save these structure evaluations, under some assumptions.

Finally on the fifth chapter we will apply all of these algorithms for practical cases such as compressing images and the most important, compressing the weight matrices of fully connected neural networks.

Chapter 2

Tensors

In this chapter we will construct tensors in a formal way and we will lay down the basics of tensor algebra. We will define the notion of rank of a tensor and then, we will describe some basic tensor reshaping operations. All of this will be crucial for then presenting tensor contractions and tensor networks on the following chapter.

2.1 The tensor product space

We will denote $\mathbb{V}_1, \ldots, \mathbb{V}_n$ as finite vector spaces over a field $\mathbb{K}$ ($\mathbb{R}$ if unspecified) of dimension $\dim \mathbb{V}_i = N_i \; \forall i = 1, \ldots, n$.

We will present the notion of a tensor. From a mathematical standpoint, it can be defined as a multilinear map:

Definition 2.1. *A **multilinear map** or a tensor is an application* $T : \mathbb{V}_1 \times \cdots \times \mathbb{V}_n \to \mathbb{K}$ *which satisfies:*

1. $T(v_1, \ldots, \lambda v_i, \ldots, v_n) = \lambda \cdot T(v_1, \ldots, v_i, \ldots, v_n)$
2. $T(v_1, \ldots, v_i + u, \ldots, v_n) = T(v_1, \ldots, v_i, \ldots, v_n) + T(v_1, \ldots, u, \ldots, v_n)$
   $\forall i = 1, \ldots, n, \; u \in \mathbb{V}_i, \lambda \in \mathbb{K}$

Linear maps and bilinear maps are specific cases of multilinear maps with $n = 1$ and $n = 2$ respectively.

Now we will present formally the tensor space and its elements by defining a relation between all the vectors in the free vector space over the cartesian product $\mathbb{V}_1 \times \cdots \times \mathbb{V}_n$. First of all, we will need to define what the free vector space of a set is:

Definition 2.2. *Let S be a set and $\mathbb{K}$ a field. We denote the free vector space over S as $\mathbb{K}[S]$, which is the vector space containing all linear combinations of elements from S with coefficients of $\mathbb{K}$, i.e:*

$$\mathbb{K}[S] = \left\{ \sum_{i=1}^{n} a_i s_i \mid s_i \in S, n \in \mathbb{Z}_+, a_i \in \mathbb{K} \right\}$$

Example 2.3. If we take $\mathbb{K} = \mathbb{R}$ and $S = (x, y)$, then the elements of $\mathbb{R}[S]$ have the form $ax + by$ with $a, b \in \mathbb{R}$. We can see that $\mathbb{R}[S]$ is the vector space $\mathbb{R}^2$ with some fixed basis (x, y).

Let $\mathbb{L} = \mathbb{K}[\mathbb{V}_1 \times \cdots \times \mathbb{V}_n]$. We define relation R as the smallest equivalence relation that satisfies:

$$(v_1, \ldots, \alpha v_i, \ldots, v_n) \sim \alpha(v_1, \ldots, v_n) \; \forall i = 1, \ldots, n, \forall \alpha \in \mathbb{K}$$

$$(v_1, \ldots, v_i + u_i, \ldots, v_n) \sim (v_1, \ldots, v_i \ldots, v_n) + (v_1, \ldots, u_i, \ldots, v_n) \; \forall i = 1, \ldots, n$$

Proposition 2.4. *$[0]_R$ is a subspace of $\mathbb{K}[\mathbb{V}_1 \times \cdots \mathbb{V}_n]$*

Proof. $0 \in [0]_R$. It is sufficient to see that for all $\lambda \in \mathbb{K}$ and $u, v \in [0]_R$ then $u + \lambda v \in [0]_R$. We can write $u + \lambda v \sim_R (u_1 + v_1, \ldots, u_i + v_i, \ldots, u_n + v_n) \sim_R 0 + (v_1, \ldots, \lambda v_i, \ldots, v_n) \sim_R 0 + \lambda 0 \sim_R 0$ and therefore $u + \lambda v \sim_R 0$ since R is an equivalence relation. □

Definition 2.5. *The **tensor product space** $\mathbb{V}_1 \otimes \cdots \otimes \mathbb{V}_n$ is defined as the quotient $\mathbb{L}/R_0$. We denote the equivalence class of $(v_1, \ldots, v_n)$ as $v_1 \otimes \cdots \otimes v_n$.*

We will see with the universal property of the tensor product that since the elements of $\mathbb{L}/R$ satisfy the properties that define a multilineal map, each vector of $\mathbb{L}/R$ is a **tensor**.

We will give an example of a tensor product space:

Example 2.6. Given $\mathbb{R}^2$ and $\mathbb{R}^3$ with its canoncial vector space structures, $\mathbb{R}^2 \otimes \mathbb{R}^3$ is defined by the equivalence classes that follow the relations

$$(\alpha u, v) \sim \alpha(u, v) \sim (u, \alpha v) \quad \forall \alpha \in \mathbb{K}, \; u \in \mathbb{R}^2, v \in \mathbb{R}^3$$

$$(u + u', v) \sim (u, v) + (u' + v) \quad \forall u, u' \in \mathbb{R}^2, v \in \mathbb{R}^3$$

$$(u, v + v') \sim (u, v) + (u + v') \quad \forall u \in \mathbb{R}^2, v, v' \in \mathbb{R}^3$$

And an element of $\mathbb{R}^2 \otimes \mathbb{R}^3$ would be the representant of $[((1,2,3),(0,4))]$ in which for example other representants of the same equivalence class would be $2 \cdot ((1,2,3),(0,2))$ or $((1,0,3),(0,2)) + ((0,2,0),(0,2))$

Now we can state and prove the universal property of the tensor product:

Theorem 2.7 (Universal property of the tensor product)**.** *Let φ be the quotient mapping from $\mathbb{V}_1 \times \cdots \times \mathbb{V}_n$ to $\mathbb{V}_1 \otimes \cdots \otimes \mathbb{V}_n$. For every multilinear map $h : \mathbb{V}_1 \times \cdots \times \mathbb{V}_n \to X$ where X is any vector space there exists an unique linear map $\tilde{h} : \mathbb{V}_1 \otimes \cdots \otimes \mathbb{V}_n \to X$ such that the following diagram commutes:*

$$\begin{array}{ccc} \mathbb{V}_1 \times \cdots \times \mathbb{V}_1 & \xrightarrow{\varphi} & \mathbb{V}_1 \otimes \cdots \otimes \mathbb{V}_n \\ & \searrow^{h} & \downarrow \tilde{h} \\ & & X \end{array}$$

Proof. Let $\varphi(v_1, \ldots, v_n) := [(v_1, \ldots, v_n)] \in \mathbb{V}_1 \otimes \cdots \otimes \mathbb{V}_n$. Let $h : \mathbb{V}_1 \times \cdots \mathbb{V}_n \to X$ be a multilinear map. We define $\tilde{H} : \mathbb{K}[\mathbb{V}_1 \times \cdots \times \mathbb{V}_n] \to X$ by:

$$\tilde{H}\left(\sum_{i=1}^{p} a_i (v_i^1, \ldots, v_i^n)\right) := \sum_{i=1}^{p} a_i h(v_i^1, \ldots, v_i^n)$$

Consider now the vector subspace $[0]_R$ of $\mathbb{K}[\mathbb{V}_1 \times \cdots \times \mathbb{V}_n]$. Since h is multilinear, we can see that $\tilde{H}$ sends every element of $[0]_R$ to $0 \in X$, therefore $W \subseteq \ker \tilde{H}$ and hence

$\tilde{H}$ induces a well-defined linear map $\tilde{h} : \mathbb{K}[\mathbb{V}_1 \times \cdots \times \mathbb{V}_n]/R = \mathbb{V}_1 \otimes \cdots \otimes \mathbb{V}_n \to X$ that satisfies $\tilde{h}(v_1 \otimes \cdots \otimes v_n) = h(v_1, \ldots, v_n)$

Suppose that exists another mapping $f : \mathbb{V}_1 \otimes \cdots \otimes \mathbb{V}_n \to X$ such that $f(v_1 \otimes \cdots \otimes v_n) = h(v_1, \ldots, v_n)$, then we would have

$$f\left(\sum_{i=1}^{p} a_i v_i^1 \otimes \cdots \otimes v_i^n\right) = \sum_{i=1}^{p} a_i h(v_i^1, \ldots, v_i^n) = \tilde{h}\left(\sum_{i=1}^{n} a_i v_i^1 \otimes \cdots \otimes v_i^n\right)$$

therefore, the linear mapping $\tilde{h}$ is unique □

We can now construct explicitly the corresponding vector space (and a basis) of $\mathbb{V}_1 \otimes \cdots \otimes \mathbb{V}_n$.

Proposition 2.8. *Let $\{e_1^i, e_2^i, \ldots, e_{N_i}^i\}$ be basis for each $\mathbb{V}_i$ and $N_i = \dim \mathbb{V}_i$. Then the set*

$$\mathcal{B}_\otimes = \{e_{i_1}^1 \otimes \cdots \otimes e_{i_n}^n = [(e_{i_1}^1, \ldots, e_{i_n}^n)]_R : 1 \leqslant i_j \leqslant N_j, 1 \leqslant j \leqslant n\}$$

is a basis of $\mathbb{V}_1 \otimes \cdots \otimes \mathbb{V}_n$.

Proof. Let $v_1 \otimes \cdots \otimes v_n \in \mathbb{V}_1 \otimes \cdots \otimes \mathbb{V}_n$, if we write $v_i = \sum_{j=1}^{N_i} \lambda_j^i e_j^i$ then, because of the multilinearity of R we get that

$$v_1 \otimes \cdots \otimes v_n = \left(\sum_{i=1}^{N_1} \lambda_1^i e_1^i\right) \otimes \cdots \otimes \left(\sum_{i=1}^{N_n} \lambda_n^i e_n^i\right) = \sum_{s_1,\ldots,s_n}^{N_1,\ldots,N_n} \lambda_1^{s_1} \cdots \lambda_n^{s_n} (e_1^{s_1} \otimes \cdots \otimes e_n^{s_n})$$

And since all elements of $\mathbb{V}_1 \otimes \mathbb{V}_n$ can be written in this form, $\mathcal{B}_\otimes$ spans the entire space. For proving the independance of each element of $\mathcal{B}_\otimes$, suppose that we have a linear combination such that

$$\sum_{s_1,\ldots,s_n}^{N_1,\ldots,N_n} \lambda_{s_1,\ldots,s_n} e_{s_1}^1 \otimes \cdots \otimes e_{s_n}^n = 0 \tag{2.1.1}$$

Let $\{{}^*e_1^i, \ldots, {}^*e_n^i\}$ be the dual basis for $\mathbb{V}^i$ such that ${}^*e_j^i(e_k^i) = \delta_{jk}$. We define the multilinear map

$$\begin{aligned} f_{(k_1,\ldots,k_n)} : \mathbb{V}_1 \times \cdots \times \mathbb{V}_n &\longrightarrow \mathbb{K} \\ (v_1, \ldots, v_n) &\longmapsto {}^*e_{k_1}^1(v_1) \cdot {}^*e_{k_2}^2(v_2) \cdots {}^*e_{k_n}^n(v_n) \end{aligned}$$

With $1 \leqslant k_i \leqslant N_i$. The image of $f_{(k_1,\ldots,k_n)}$ extracts the coefficient $\lambda_{k_1,\ldots,k_n}$ of any tensor $v_1 \otimes \cdots \otimes v_n$.

Applying the universal property of the tensor product, there exists an unique linear map $\tilde{f}_{(s_1,\ldots,s_n)} : \mathbb{V}_1 \otimes \cdots \otimes \mathbb{V}_n \to \mathbb{K}$ such that

$$\tilde{f}_{(k_1,\ldots,k_n)}(e_{j_1}^1 \otimes \cdots \otimes e_{j_n}^n) = f_{(k_1,\ldots,k_n)}(e_{j_1}^1, \ldots, e_{j_n}^n) = \delta_{j_1 k_1} \cdot \delta_{j_2 k_2} \cdots \delta_{j_n k_n}$$

If we now apply the linerar combination to $\tilde{f}_{(s_1,\ldots,s_n)}$ we get

$$\tilde{f}_{(k_1,\ldots,k_n)}\left(\sum_{s_1,\ldots,s_n}^{N_1,\ldots,N_n} \lambda_{s_1,\ldots,s_n} e_{s_1}^1 \otimes \cdots \otimes e_{s_n}^n\right) = \sum_{s_1,\ldots,s_n}^{N_1,\ldots,N_n} \lambda_{s_1,\ldots,s_n} \cdot \delta_{k_1 s_1} \cdot \delta_{k_2 s_2} \cdots \delta_{k_n s_n} = \lambda_{k_1,\ldots,k_n}$$

But from (2.1.1) $\lambda_{k_1,\ldots,k_n} = 0$ for all $1 \leqslant k_i \leqslant N_i$. Therefore, the elements of $\mathcal{B}_\otimes$ are independent and form a basis. □

Now we can directly work with the tensor space since we have a well defined basis. The dimension of $\mathbb{V}_1 \otimes \cdots \otimes \mathbb{V}_n$ is $N_1 \cdot N_2 \cdots N_n$ and its elements can be expressed as

$$T = \sum_{s_1,\ldots,s_n}^{N_1,\ldots,N_n} T_{s_1,\ldots,s_n} \cdot e^1_{s_1} \otimes \cdots \otimes e^n_{s_n} \tag{2.1.2}$$

If we wanted to store a tensor in a computer program, it would be enough to save all the $T_{s_1,\ldots,s_n}$ entries. We will define the size of a tensor as the number of entries of $\mathbb{K}$ that we would need to store the tensor.

Definition 2.9. *We will define the* ***size*** *of the tensor $T \in \mathbb{V}_1 \otimes \cdots \otimes \mathbb{V}_n$ as* $\text{Size}(T) = N_1 N_2 \cdots N_n$. *We will also say that the* ***order*** *of T is n.*

Now, we would want to define a product between two tensors of different tensor product spaces. In other words, suppose that $T \in \mathbb{V}_1 \otimes \cdots \mathbb{V}_n$ and $U \in \mathbb{W}_1 \otimes \cdots \otimes \mathbb{W}_m$. We want to define a tensor product operation that its output gives tensor $T \otimes U \in \mathbb{V}_1 \otimes \cdots \otimes \mathbb{V}_n \otimes \mathbb{W}_1 \otimes \cdots \otimes \mathbb{W}_m$. We will define this operation in the following way:

Definition 2.10 (Tensor product)**.** *With the above notation, let* $\dim \mathbb{V}_i = N_i$, $\dim \mathbb{W}_j = M_j$ *and some basis* $\{e^i_1, \ldots, e^i_{N_i}\}$ *of each* $\mathbb{V}_i$ *and* $\{p^j_1, \ldots, p^j_{M_j}\}$ *of each* $\mathbb{W}_j$. *We define the tensor product $T \otimes U$ as*

$$T \otimes U = \sum_{i_1,\ldots,i_n}^{N_1,\ldots,N_n} \sum_{j_1,\ldots,j_m}^{M_1,\ldots,M_m} T_{i_1,\ldots,i_n} U_{j_1,\ldots,j_m} \cdot e^1_{i_1} \otimes \cdots \otimes e^n_{i_n} \otimes p^1_{j_1} \otimes \cdots \otimes p^m_{j_m} \tag{2.1.3}$$

Which naturally is an element of $\mathbb{V}_1 \otimes \cdots \otimes \mathbb{V}_n \otimes \mathbb{W}_1 \otimes \cdots \otimes \mathbb{W}_m$. We call the tensor product for two tensors of order 2 as the **outer product**, so we can say that the tensor product is a generalization of it.

Example 2.11. Let $\{a_1, a_2\} \subset \mathbb{V}_1$, $\{b_1, b_2\} \subset \mathbb{V}_2$ and $\{c_1, c_2\} \subset \mathbb{W}_1$, $\{d_1, d_2\} \subset \mathbb{W}_2$ be basis of their corresponding vector spaces. Let $T \in \mathbb{V}_1 \otimes \mathbb{V}_2$ and $U \in \mathbb{W}_1 \otimes \mathbb{W}_2$ defined as:

$$T = 2(a_1 \otimes b_1) + 3(a_2 \otimes b_1) \qquad U = c_1 \otimes d_1 + c_2 \otimes d_2$$

Then, the tensor product $T \otimes U$ would be:

$$T \otimes U = (2(a_1 \otimes b_1) + 3(a_2 \otimes b_1)) \otimes (c_1 \otimes d_1 + c_2 \otimes d_2)$$

Applying multilinearity we get:

$$T \otimes U = 2(a_1 \otimes b_1 \otimes c_1 \otimes d_1) + 2(a_1 \otimes b_1 \otimes c_2 \otimes d_2) +$$

$$3(a_2 \otimes b_1 \otimes c_1 \otimes d_1) + 3(a_2 \otimes b_1 \otimes c_2 \otimes d_2)$$

Now, we already know how to add and subtract two tensors of the same tensor product space since its sum is already defined by its vector space. We will introduce a tensor norm since in the following chapters we will want to know if a tensor is "small" or "big", and also but not less important, we want to know if two tensors are near each other to test convergence for algorithms involving tensor networks. In our case we will stick to the Frobenius norm:

Definition 2.12. *We define the frobenius norm as:*

$$\|\cdot\|_F : \mathbb{V}_1 \otimes \cdots \otimes \mathbb{V}_n \longrightarrow \mathbb{R}_+$$
$$\left(\sum_{s_1,\ldots,s_n}^{N_1,\ldots,N_n} T_{s_1,\ldots,s_n} \cdot e^1_{s_1} \otimes \cdots \otimes e^n_{s_n}\right) \longmapsto \sqrt{\sum_{s_1,\ldots,s_n}^{N_1,\ldots,N_n} T^2_{s_1,\ldots,s_n}}$$

2.2 Tensor ranks

In this section we will present the rank of a tensor. It will serve as the extension of the matrix rank, which is defined as the dimension of the vector space spanned by the vectors on its columns. For the matrix rank, we say that if it is spanned by a single vector, the rank is 1, so we could somewhat say that a vector has rank 1. We will do something similar, defining that a tensor t is a rank-1 tensor if it can be written as

$$t = v^1 \otimes \cdots \otimes v^n$$

with $v^i \in \mathbb{V}_i$. Therefore, the rank of a tensor will be r if it can be spanned from r rank-1 tensors:

Definition 2.13. *We say that a tensor T has rank r as* $\operatorname{rank} T = r$ *with $r \in \mathbb{Z}_+$ if r is the minimum value such that we can write T as the following form*

$$T = \sum_{p=1}^{r} \lambda_p v^1_p \otimes \cdots \otimes v^n_p \tag{2.2.1}$$

where $v^i_1, \ldots, v^i_r \in \mathbb{V}_i, i = 1, \ldots, n$ and $\lambda_p \in \mathbb{K}$

The rank of a tensor is bounded by $\prod_{i=1}^n N_i$ since we can decompose every tensor as the sum of the elements of a basis of the tensor product space, and so the rank of the tensor does not exceed the number of elements of the basis.

Unlike matrices, determining the rank of a tensor is an NP-hard problem (See Section 8 of [6]). Finding the maximum rank, i.e determining $\max_{T \in \mathbb{V}_1 \otimes \cdots \otimes \mathbb{V}_n} \operatorname{rank} T$ still remains an unresolved problem. [7]

By manipulating the vectors $v^1_p \otimes \cdots \otimes v^n_p$ to represent each value of the tensor $T_{s_1,\ldots,s_n}$, we can find an slightly better upper bound for the tensor rank:

Proposition 2.14. *Let* $\dim \mathbb{V}_i = N_i$*, then*

$$\operatorname{rank} T \leqslant \left\lfloor \frac{\prod_{i=1}^n N_i}{\sum_{i=1}^n N_i} \right\rfloor \tag{2.2.2}$$

Proof. Let $r = \operatorname{rank} T$. We can write $T = \sum_{p=1}^r v^1_p \otimes \cdots \otimes v^n_p$. Now, each term of this sum has $\sum_{i=1}^n N_i$ adjustable parameters, since each $v^{(i)}_p$ is a vector of $\mathbb{V}_i$ with its dimension being N_i. So, in total we will have $r \sum_{i=1}^n N_i$ adjustable parameters in our decomposition. Since our tensor T is completly determined by $\prod_{i=1}^n N_i$ parameters, we can impose $r \sum_{i=1}^n N_i \leqslant \prod_{i=1}^n N_i$ □

Decomposing a tensor T in rank-1 tensors as in eq. (2.2.1) is known as **tensor rank decomposition**. One could ask if given a tensor T and fixed r, can we construct a tensor T' of some fixed rank r' such that $\|T - T'\|_F$ is minimum. The decomposition T' is called **canonical polyadic decomposition** and it is an special case of a tensor network that we will see on the following chapter.

We will now present some essential reshaping operations that will help us manipulating tensors. Our main goal by presenting the reshaping operations is the unfolding and folding operations, that somehow "transforms" a tensor onto a matrix and the other way around respectively.

2.3 Reshaping operations

Any tensor $T \in \mathbb{V}_1 \otimes \cdots \otimes \mathbb{V}_n$ can be identified as an n-dimensional array. In other words, for each tensor T we can define a discrete function $\mathcal{T}$ that encodes the representation in a basis of the tensors T as:

$$\begin{aligned} \mathcal{T} : \prod_{i=1}^{n} \{1, \dots, N_i\} &\longrightarrow \mathbb{K} \\ (i_1, \dots, i_n) &\longmapsto T_{i_1,\dots,i_n} \end{aligned}$$

From now on we will identify the set of all images of $\mathcal{T}$ as an element of $\mathbb{K}^{N_1 \times \cdots \times N_n}$. Therefore, a lot of times we will write a tensor $T \in \mathbb{V}_1 \otimes \cdots \otimes \mathbb{V}_n$ as an element of $\mathbb{K}^{N_1 \times \cdots \times N_n}$ with $\dim \mathbb{V}_i = N_i$, since this correspondance is already well defined. We will also write $T(i_1, \dots, i_n)$ as the image of $\mathcal{T}$ of $(i_1, \dots, i_n)$. Since now we can see a tensor as an n-dimensional array thanks to the mapping $\mathcal{T}$, we can start reshaping tensors. But before that, we will define what is a mode of a tensor:

Definition 2.15. *We define the j-th of a tensor as its j-th dimension. A tensor of order n has n different modes.*

For example, having a 3-order tensor $T \in \mathbb{R}^{N_1 \times N_2 \times N_3}$. We can write each entry of the tensor as $T(i_1, i_2, i_3)$. The first mode of T is i_1, the second one, i_2 and the third i_3. In other words, when we say the j-mode of a tensor we are referring at one argument of the discrete function $\mathcal{T}$

Now we will introduce the linearization operation which simplify the notation a lot when we define reshaping operations.

Definition 2.16 (Linearization)**.** *Fixed $N_1, \dots, N_n \in \mathbb{Z}_+$, given $i_1, \dots, i_n \in \mathbb{Z}_+$ such that $1 \leqslant i_1, \leqslant N_1, \dots, 1 \leqslant i_n \leqslant N_n$, we define the **linearization** of the indices $i_1, \dots, i_n$ as the mapping $\prod_{i=1}^{n} \{1, \dots, N_i\} \to \{1, \dots, \prod_{i=1}^{n} N_i\}$ and with its images defined as:*

$$\overline{i_1, i_2, \dots, i_n} = \sum_{j=2}^{n} \left((i_j - 1) \prod_{k=1}^{j} N_k \right) + i_1$$

The purpose of the linearization mapping is to give a bijection within each element of the form $(i_1, \dots, i_n) \in \prod_{i=1}^{n} \{1, \dots, N_i\}$ to a positive natural number. For example, consider $n = 3$ and $N_1 = N_2 = N_3 = 3$. The encoding of the tuple $(1, 1, 1)$ corresponds to 1, the tuple $(2, 1, 1)$ to 2, the tuple $(1, 2, 3)$ to $1 + (2-1) \cdot 3 + (3-1) \cdot 3 = 10$. One should keep in

mind that when defining an array in computer science, usually the first element starts at the index 0, and our linearization operation starts with the indices at 1. Throughout the thesis we will stick to array indices starting at 1 for consistency, but by doing a change of variables $i'_j = i_j - 1$ before applying the linearization and then subtracting 1 also on the image, we will get the same operation but for arrays starting at 0.

We will now define the vectorization operation, which reshapes a tensor $T \in \mathbb{K}^{N_1 \times N_2 \times \cdots \times N_n}$ to a vector of $\mathbb{K}^{N_1 \cdot N_2 \cdots N_n}$

Definition 2.17. *We define the* ***vectorization*** *of T as the first order tensor (or vector) $\mathcal{V} \in \mathbb{K}^{N_1 N_2 \cdots N_n}$ defined entrywise as*

$$\mathcal{V}(\overline{i_1 i_2 \dots i_n}) = T(i_1, i_2, \dots, i_n)$$

We will write the vectorization of T as $\operatorname{vec} T$

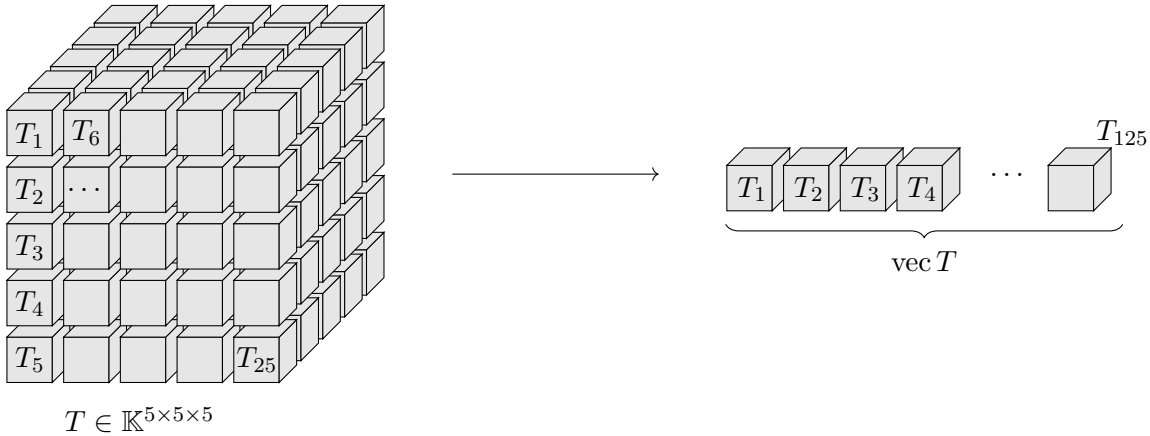


Figure 2.1: Representation of tensor vectorization. Source: own elaboration

We will also define the permutation of the modes of a tensor. Permutting modes can be seen as swapping arguments of the function $\mathcal{T}$ following some permutation $\sigma \in S_n$.

Definition 2.18 (Tensor permutation)**.** *Given a permutation $\sigma \in S_n$, and a tensor $T \in \mathbb{K}^{N_1 \times N_2 \times \cdots \times N_n}$ we define $T_\sigma \in \mathbb{K}^{N_{\sigma(1)} \times N_{\sigma(2)} \times \cdots \times N_{\sigma(n)}}$ entrywise as*

$$T_\sigma(i_1, \dots, i_n) = T(i_{\sigma^{-1}(1)}, i_{\sigma^{-1}(2)}, \dots, i_{\sigma^{-1}(n)})$$

Example 2.19. Let $N_1, N_2 \in \mathbb{Z}_+$ and $M \in \mathbb{K}^{N_1 \times N_2}$. The matrix transposition M^T can be written as $M_{(2,1)}$

Example 2.20. Suppose that we have $X \in \mathbb{R}^{3\times2\times2}$ defined as

$$X = \left[\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right]$$

Then the tensor $X_{(3,1,2)} \in \mathbb{R}^{2\times2\times3}$ would be written entrywise as

$$X_{(3,1,2)}(i_1, i_2, i_3) = X(i_3, i_1, i_2)$$

And therefore,

$$X_{(3,1,2)} = \left[\begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 \\ 1 & 1 & -1 \end{pmatrix} \right]$$

Reshaping tensors onto matrices will also be very useful since it will let us treat high order tensors as matrices and then apply numerical algorithms there.

Definition 2.21 (Tensor unfolding)**.** *Let T be a tensor of order n with $n \geqslant 2$. Let $\sigma \in S_n$ be a permutation of $(1, 2, \ldots, n)$. We define the **unfolding** of the tensor T as the $2nd$-order tensor or matrix $\mathcal{U} \in \mathbb{R}^{\prod_{i=1}^{d} N_{\sigma(i)} \times \prod_{i=d+1}^{n} N_{\sigma(i)}}$ entrywise as*

$$\mathcal{U}(\overline{i_{\sigma(1)}, \ldots, i_{\sigma(d)}}, \overline{i_{\sigma(d+1)}, \ldots, i_{\sigma(n)}}) = T(i_1, \ldots, i_n)$$

We will write $\mathcal{U} = \text{unfold}\,(\mathcal{T}, (\sigma(1), \ldots, \sigma(d)), (\sigma(d+1), \ldots, \sigma(n)))$

Example 2.22. Suppose that we have X defined as in Example 2.20. Then unfold $(X, (1, 2), (3))$ would result in a matrix $\mathbb{R}^{3\times 4}$ with its elements defined as $\mathcal{U}(\overline{i_1 i_2}, \overline{i_3}) = X(i_1, i_2, i_3)$. Computing each entry gives

$$\text{unfold}\,(X, (1, 2), (3)) = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & -1 \end{pmatrix}$$

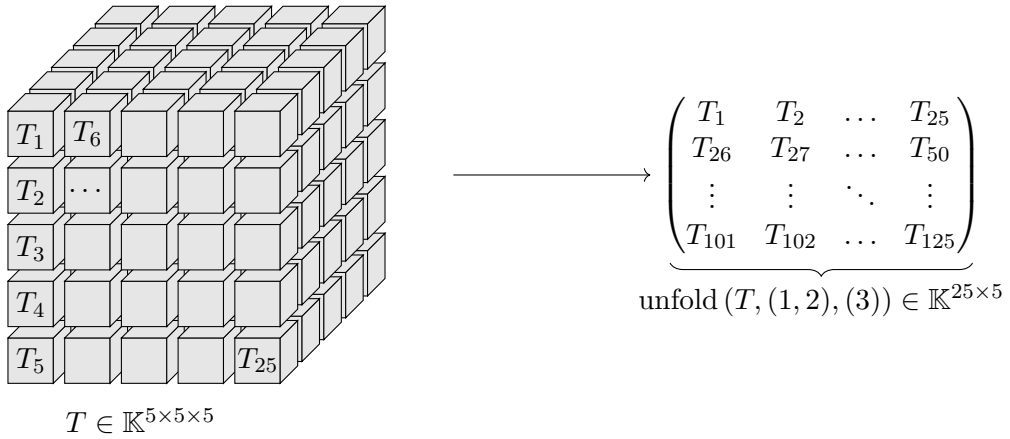


Figure 2.2: Representation of tensor unfolding. Source: own elaboration

Now we will introduce the reverse operation of the tensor unfolding: tensor folding.

Definition 2.23 (Tensor folding)**.** *Let $\sigma \in S_n$, $1 \leqslant d \leqslant n$. Let $N = \prod_{i=1}^{d}$, $\mathcal{U} = \prod i = d + 1^n$ and M be a matrix of $\mathbb{K}^{N\times M}$. We define the **folding** of M following σ as the tensor $T \in \mathbb{K}^{N_{\sigma(1)}\times\cdots\times N_{\sigma(n)}}$ defined entrywise as:*

$$T(i_1, \ldots, i_n) = \mathcal{U}(\overline{i_{\sigma^{-1}(1)}, \ldots, i_{\sigma^{-1}(d)}}, \overline{i_{\sigma^{-1}(d+1)}, \ldots, i_{\sigma^{-1}(n)}})$$

We will write $T = (\mathcal{U}, (\sigma(1), \ldots, \sigma(d)), (\sigma(d+1), \ldots, \sigma(n)))$

Now that we have defined tensors and all the basic operations of tensor algebra, in the following chapter we will construct tensor networks using the contraction operation that we have seen in this chapter. On chapter 4, the reshaping operations will reappear when we present the algorithms that find tensor network cores.

Chapter 3

Tensor networks

In this chapter we will start by presenting tensor networks from a mathematically formal standpoint. Then, we will present the notion of the tensor G-rank that will be useful for determining if a tensor network can represent a tensor. We will present some common tensor network structures. After that, we will discuss over the ordering in which its the most optimal to contract a tensor network and we will give an algorithm that finds an optimal order of contraction, and finally, we will present the alternating least squares algorithm applied to generic tensor networks for explicitly finding the tensor cores of a tensor network.

The concept of tensor networks originated from a physics background. Roger Penrose described how its diagrammatic language could be used in various applications of physics [8]. Later, in 1992, Steven R. White developed de Density Matrix Renormalization Group (DRMG) algorithm for quantum lattice systems. It was considered the first successfull tensor network application [9].

3.1 Tensor contraction and the Penrose Notation

In this section we will define the tensor contraction operation, which is one of the most important operations in tensor algebra and is the core concept behind tensor networks. The idea behind tensor contraction is that given two tensors from different spaces, say for example $T \in \mathbb{R}^{N_1 \times \cdots \times N_n}$ and $U \in \mathbb{R}^{M_1 \times \cdots \times M_m}$, we pick one dimension of each tensor N_i and M_j with the same size $N_i = M_j$ and then we obtain a new tensor of order $n + m - 2$ that contains all dimensions except for N_i and M_j, and each element of the resulting tensor is obtained by generalizing the matrix product opertion but along N_i and M_j dimensions.

So, the elements of the tensor contraction $T \times_j^i U$ would be, element-wise:

$$\sum_{k=1}^{N_i} T(s_1, \ldots, s_{i-1}, k, s_{i+1}, \ldots s_n) \cdot U(t_1, \ldots, t_{j-1}, k, t_{j+1}, \ldots, t_m)$$

Where $T \times_j^i U \in \mathbb{K}^{N_1 \times \cdots \times N_{i-1} \times N_{i+1} \times \cdots \times N_n \times M_1 \times \cdots \times M_{j-1} \times M_{j+1} \times \cdots \times M_m}$

Since the notation of tensor contractions is often very tedious, we will introduce the Penrose Notation for representing tensor contractions in a more compact and elegant

way. The Penrose notation dates back from at least the early 1970s and was firstly used by Robert Penrose, to which the name is owed. [8]

Given an nth-order tensor $T \in \mathbb{K}^{N_1 \times \cdots \times N_n}$ we represent it using the Penrose notation is as a circle with as many edges as the order of the tensor, as seen in fig. 3.1

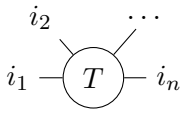


Figure 3.1: Representation of a tensor $T \in \mathbb{K}^{N_1 \times \cdots \times N_n}$ using the Penrose notation. Source: own elaboration

The dimensions of the tensor are not explicitly written in the Penrose notation. Instead, only the order of the indices is preserved. If they are not explicitly set on the labels of the edges, the order of the indexes of the tensor will be determined by their orientation respect to the circle: the order starts from the left and then follows a clockwise rotation. The order in which we encounter the edges will be the order of the indexes. For example, in fig. 3.1, the order would be $i_1, i_2, \ldots, i_n$

Then, we represent a contraction of two tensors on the Penrose notation by joining to edges of different tensors. The edges that are joined will be the edges that the contraction is performed. As seen in fig. 3.2.

Now, we will give a more formal definition of the tensor contraction since there is a natural way in which tensor contraction definition appears from applying a vector space with its dual with different tensors. As before, we take two tensors T and U with $T \in \mathbb{V}_1 \otimes \cdots \otimes \mathbb{V}_n$ and $U \in \mathbb{W}_1 \otimes \cdots \otimes \mathbb{W}_m$. Suppose that there exists some vector space of the dimensions of U that is the dual of some space of the tensor T, in other words, suppose that exists some i and j in which $\mathbb{W}_j = \mathbb{V}_i^*$. Then, the tensor contraction is doing the tensor product of T and U and then applying $\mathbb{V}_i$ with its dual afterwards. Doing this gives the following definition:

Definition 3.1. *Let $P \in \mathbb{V}_1 \otimes \cdots \otimes \mathbb{V}_n$ and suppose that there exists some i and j such that $\mathbb{V}_j = \mathbb{V}_i^*$ The following map*

$$\begin{aligned} \mathcal{C}_j^i : \mathbb{V}_1 \otimes \cdots \otimes \mathbb{V}_n &\longrightarrow \mathbb{V}_1 \otimes \cdots \otimes \mathbb{V}_{i-1} \otimes \mathbb{V}_{i+1} \otimes \cdots \otimes \mathbb{V}_{j-1} \otimes \mathbb{V}_{j+1} \otimes \cdots \otimes \mathbb{V}_n \\ v_1 \otimes \cdots \otimes v_n &\longmapsto (v_1 \otimes \cdots \otimes v_{i-1} \otimes v_{i+1} \otimes \cdots \otimes v_{j-1} \otimes v_{j+1} \otimes \cdots \otimes v_n)\, v_j(v_i) \end{aligned}$$

where $v_j \in \mathbb{V}_i^*$. We define $\mathcal{C}_j^i$ as the tensor contraction mapping of $\mathbb{V}_1 \otimes \cdots \otimes \mathbb{V}_n$ over the indices i and j. We call $\mathcal{C}_i^j(T)$ the contraction of T by indices (i, j).

Now, we define the **contraction of two tensors** T and U defined as before with some space $\mathbb{W}_j$ being the dual of some $\mathbb{V}_i$ as $\mathcal{C}_j^i(T \otimes U)$. We will sometimes write $T \times_j^i U$

We will represent the contraction between two tensors as their representation in the Penrose notation with the edges that represent the indexes that are contracting by joining them, as seen in fig. 3.2.

If we fix basis for $\mathbb{V}_1, \ldots, \mathbb{V}_p, \mathbb{W}_1, \ldots, \mathbb{W}_q$ and we represent X, Y as discrete functions by

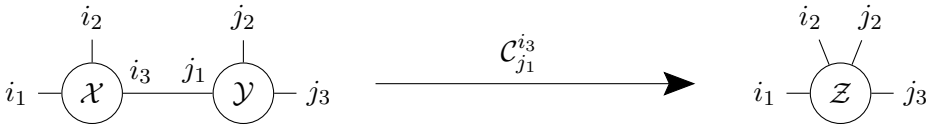


Figure 3.2: Representation in the Penrose notation of the contraction between two tensors $\mathcal{X} \in \mathbb{K}^{N_1 \times N_2 \times N_3}, \mathcal{Y} \in \mathbb{K}^{M_1 \times M_2 \times M_3}$ by their indices i_3 and j_1 with $N_1, N_2, N_3, M_1, M_2, M_3 \in \mathbb{Z}_+$ and $N_3 = M_1$. Source: own elaboration

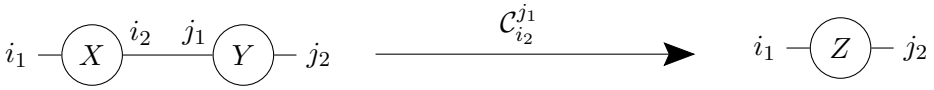
its representations in those basis, we get a way for computing $\mathcal{C}_k^l(X \otimes Y)$ as:

$$\begin{aligned} &\mathcal{C}_k^l(i_1, \ldots, i_{k-1}, i_{k+1}, \ldots, i_p, j_1, \ldots, j_{l-1}, j_{l+1}, \ldots, j_q) \\ &= \sum_{s=1}^{N_k} \mathcal{X}(i_1, \cdots, i_{k-1}, s, i_{k+1}, \cdots, i_p) \mathcal{Y}(j_1, \cdots, j_{l-1}, s, j_{l+1}, \cdots, j_q) \end{aligned} \tag{3.1.1}$$

Example 3.2. The tensor contraction for two tensors $M_1 \in \mathbb{K}^{N_1 \times N_2}$ and $M_2 \in \mathbb{K}^{N_2 \times N_3}$ over one edge of each tensor yields the matrix multiplication. Applying eq. (3.1.1) we get that $Z = C_2^2(X \otimes Y)$ is defined entry-wise as:

$$Z(i_1, j_2) = \sum_{s=1}^{N_2} M_1(i_1, s) M_2(s, j_2)$$

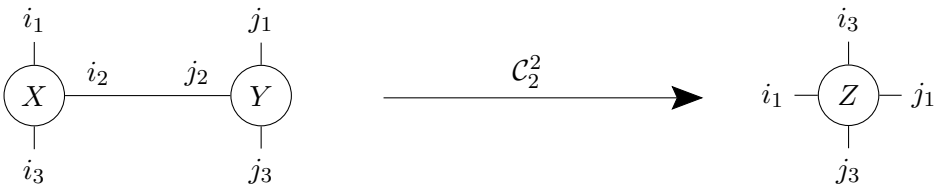
And that is identical to the conventional matrix product. Visually, we can represent it using the Penrsone notation:



Example 3.3. Suppose that we take $X \in \mathbb{R}^{3 \times 2 \times 2}$ from Example 2.22 and $Y \in \mathbb{R}^{2 \times 2 \times 3}$ defined as:

$$X = \left[\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right] \qquad Y = \left[\begin{pmatrix} 1 & 0 & -1 \\ 0 & 2 & -2 \end{pmatrix}, \begin{pmatrix} 1 & 1 & 0 \\ 2 & 0 & 2 \end{pmatrix} \right]$$

We want to compute the contraction of X and Y from the second index of each tensor $Z = C_2^2(X \otimes Y)$. Note that $Z \in \mathbb{R}^{3 \times 2 \times 2 \times 3}$. Using the Penrose Notation, the contraction would look like



And each element of the contracted tensor Z would be defined as:

$$Z(i_1, i_3, j_1, j_3) = \sum_{s=1}^{2} X(i_1, s, i_3) Y(j_1, s, j_3)$$

For example, by fixing $i_1 = i_3 = 1$ we can compute the first "inner matrix" of Z:

$$Z(1,1,j_1,j_3) = \sum_{s=1}^{2} X(1,s,1)Y(j_1,s,j_3)$$

This sum is equal to

$$X(1,1,1)Y(j_1,1,j_3) + X(1,2,1)Y(j_1,2,j_3) = Y(j_1,1,j_3) = \begin{pmatrix} 1 & 0 & -1 \\ 1 & 1 & 0 \end{pmatrix}$$

If we do the same for the rest of indices i_1, i_3, we get

$$Z = \begin{bmatrix} \begin{pmatrix} 1 & 0 & -1 \\ 1 & 1 & 0 \end{pmatrix} & \begin{pmatrix} 0 & 2 & -2 \\ 2 & 0 & 2 \end{pmatrix} \\ \begin{pmatrix} 1 & 2 & -3 \\ 3 & 1 & 2 \end{pmatrix} & \begin{pmatrix} 0 & 2 & -2 \\ 2 & 0 & 2 \end{pmatrix} \\ \begin{pmatrix} 1 & 0 & -1 \\ 1 & 1 & 0 \end{pmatrix} & \begin{pmatrix} 0 & -2 & 2 \\ -2 & 0 & -2 \end{pmatrix} \end{bmatrix}$$

We can that the contraction that we defined on the Penrose notation (collapsing one edge of G) is the same as the formal definition of the contraction of two tensors over the modes that correspond to each edge.

Now we will be extending the Penrose notation for representing some special cases:

Identity tensors

Identity tensors on the Penrose notation are represented as free edges that are not connected to any tensor. We denote by I_n as the tensor identity of order n. I_2 are identity matrices.

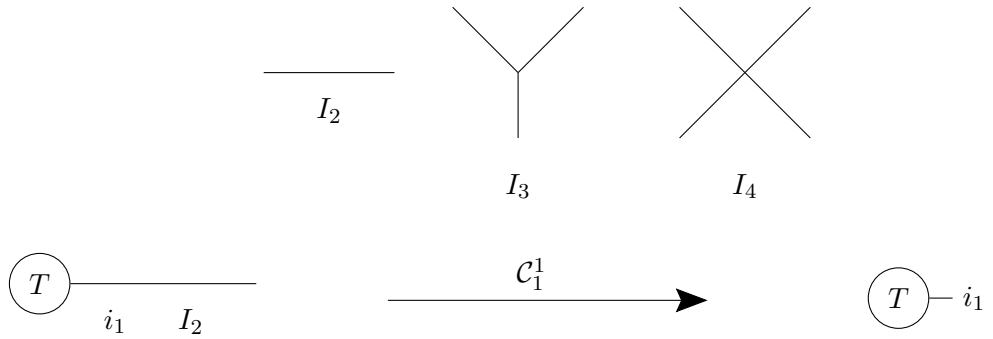


Figure 3.3: The identity matrix represented using the Penrose Notation. Contracting a mode over the identity yields the same tensor. Source: own elaboration.

This makes complete sense since contracting some tensor along some mode over the identity matrix yields the same tensor. This is ilustrated in fig. 3.3

Trace

There may be the case that when contracting a series of tensors, we might end up as what we see as a loop in the Penrose Notation. Contracting over these two indexes we

get the trace of the tensor $\mathcal{T}$ respect the indices i_k and i_p and we denote it as $\mathrm{Tr}_p^k(T)$ (See fig. 3.4). This operation is well defined since contracting over two modes of the same tensor is doing a tensor contraction using Definition 3.1.

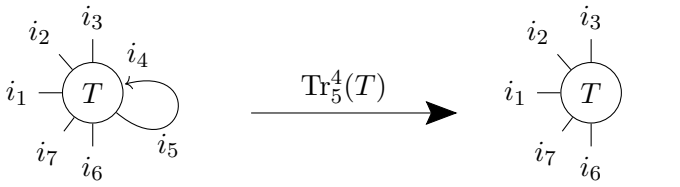


Figure 3.4: Representation of the trace of a tensor using the Penrose notation. Source: own elaboration.

Diagonal tensors

A diagonal tensor $\Lambda \in \mathbb{K}^{N\times\cdots\times N}$ of order n is a tensor which only has entries in his diagonals, i.e, there are $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$ such that

$$\Lambda(i_1, \ldots, i_n) = \begin{cases} \lambda_j & \text{if } i_1 = \ldots i_n = j \\ 0 & \text{otherwise} \end{cases}$$

We will draw diagonal tensors as identity tensors but with an small circle on the middle.

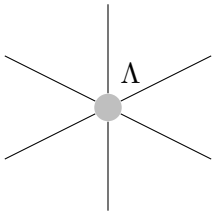


Figure 3.5: Diagonal tensor represented using the Penrose Notation. Source: own elaboration.

Outer product

And if two tensors are unconnected, we can represent it as one tensor but we do not perform any contraction between them since they are not connected. In this case, the outer product of the two tensors is performed.

Figure 3.6: Representation of the outer product using the Penrose notation. Source: own elaboration.

As we previously said, eq. (3.1.1) is equivalent to a matrix product, therefore it is possible to compute a tensor contraction by multiplying two tensors unfolded in a concrete way. the tensors $\mathcal{X}$ and $\mathcal{Y}$, we can compute $\mathcal{C}_l^k(X \otimes Y)$ as a matrix product.

Corollary 3.4. *Let* $X \in \mathbb{K}^{N_1 \times \cdots \times N_k \times \cdots \times N_n}, Y \in \mathbb{K}^{M_1 \times \cdots \times M_l \times \cdots \times M_m}$ *with* $1 \leqslant k \leqslant n$, $1 \leqslant l \leqslant m$, $N_k = M_l$. *The matrix product*

$$\mathcal{X}(\overline{i_1, \ldots, i_{k-1}, i_{k+1}, \ldots, i_n}, \overline{i_k}) \cdot \mathcal{Y}(\overline{j_l}, \overline{j_1, \ldots, j_{l-1}, j_{l+1}, \ldots, j_m})$$

Results in a $\left(\prod_{i=1}^{n} N_i\right)/N_k \times \left(\prod_{i=1}^{m} M_i\right)/M_l$ *matrix, which can be reshaped back onto* $C_l^k(X \otimes Y)$.

We will see an example of this corollary:

Example 3.5. Following from Example 3.3, we will reshape the tensors X and Y for computing Z by only performing a matrix product. We need to first compute the matrices $\mathcal{X} = \text{unfold}\,(X, (1,3), (2)) \in \mathbb{R}^{6\times 2}$ and $\mathcal{Y} = \text{unfold}\,(Y, (2), (1,3)) \in \mathbb{R}^{2\times 6}$. These unfoldings result in:

$$X = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \\ 0 & 1 \\ 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad Y = \begin{pmatrix} 1 & 0 & -1 & 1 & 1 & 0 \\ 0 & 2 & -2 & 2 & 0 & 2 \end{pmatrix}$$

Therefore:

$$Z = XY = \begin{pmatrix} 1 & 0 & -1 & 1 & 1 & 0 \\ 0 & 2 & -2 & 2 & 0 & 2 \\ 1 & 2 & -3 & 3 & 1 & 2 \\ 0 & 2 & -2 & 2 & 0 & 2 \\ 1 & 0 & -1 & 1 & 1 & 0 \\ 0 & -2 & 2 & -2 & 0 & -2 \end{pmatrix}$$

And we can now reshape $Z \in \mathbb{R}^{9\times 9}$ as the tensor in $\mathbb{R}^{3\times 3\times 3\times 3}$ as

$$Z(i_1, i_3, j_1, j_3) = Z(\overline{i_1, i_3}, \overline{j_1, j_3})$$

which in fact, is identical to Z in Example 3.3

3.2 Tensor Network States

In this section we will present the formal definition of tensor network states and the definition of a tensor network. We will use directed graphs to represent tensor networks. Before that we will need to also define what are input and output edges:

Definition 3.6. *Given a directed graph* $G = (V, \bar{E})$ *and a vertex* $i \in V$ *we define*

$$\text{IN}(i) = \{j \in V : (j, i) \in \bar{E}\} \qquad \text{OUT}(i) = \{j \in V : (i, j) \in \bar{E}\}$$

The way tensor networks are constructed consists of picking a connected directed graph $G = (V, \bar{E})$, and for each vertex $i \in V$ we assign a vector space $\mathbb{V}_i$ and for each edge $(i, j) \in \bar{E}$ we assign a vector space $\mathbb{E}_i$ to the tail of the edge and its dual covector space $\mathbb{E}_i^*$ to the head of the edge. In other words, tensor networks can be seen as a series of contractions between a set of tensors $\mathcal{G}_1, \ldots, \mathcal{G}_n$ which once all of them are contracted results in a tensor of $\mathbb{V}_1 \otimes \cdots \otimes \mathbb{V}_n$

More formally, let $\mathbb{V}_1, \dots, \mathbb{V}_d$ be vector spaces with $\dim \mathbb{V}_i = N_i, i = 1, \dots, d$. Let $\mathbb{E}_1, \dots, \mathbb{E}_c$ be finite vector spaces with $\dim \mathbb{E}_i = R_i, i = 1, \dots, c$. For each vertex $i \in V$ we will associate the tensor product space

$$\xi_i := \left(\bigotimes\nolimits_{j \in \mathrm{IN}(i)} \mathbb{E}_j \right) \otimes \mathbb{V}_i \otimes \left(\bigotimes\nolimits_{j \in \mathrm{OUT}(i)} \mathbb{E}_j^* \right)$$

A tensor network is completly defined by the graph G and all the tensor product spaces $\xi_1, \dots, \xi_n$. We will also associate to the tensor network a contraction mapping $\mathcal{C}_G$ defined by contracting factors in $\mathbb{E}_j$ with factors of $\mathbb{E}_j^*$:

$$\mathcal{C}_G : \bigotimes\nolimits_{i=1}^{n} \xi_i \longrightarrow \bigotimes\nolimits_{i=1}^{d} \mathbb{V}_i$$

Note that we have given this shapes to the tensors that we fix onto each vertex $i \in V$ because when we contract the whole graph, we will get a tensor of $\mathbb{V}_1 \otimes \cdots \otimes \mathbb{V}_n$. Since every directed edge (i, j) must point out of a vertex i and point into a vetex j, each copy of $\mathbb{E}_j$ is paired with one copy of $\mathbb{E}^*$, so the contraction $\mathcal{C}_G$ is well defined and it results in a tensor of $\mathbb{V}_1 \otimes \cdots \otimes \mathbb{V}_n$ (See fig. 3.7)

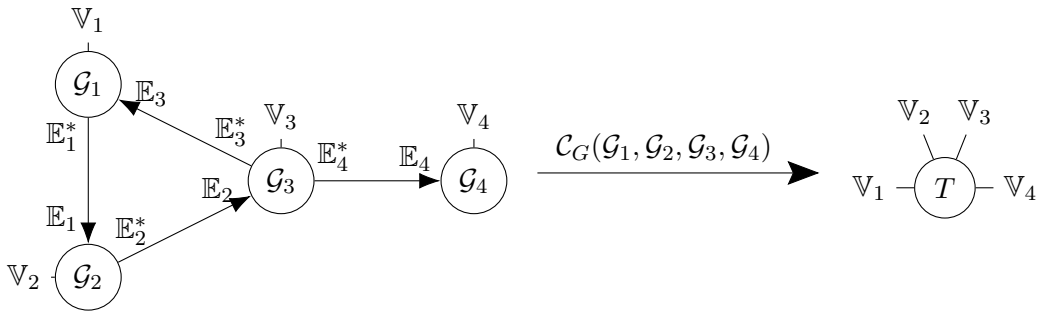


Figure 3.7: Example of a tensor network and a tensor network state evaluation using the Penrose notation. Source: own elaboration.

By picking some tensors $\mathcal{G}_i \in \xi_i$ and evaluating them through $\mathcal{C}_G$ we get a tensor T that we will call **tensor network state**. We will call the tensors $\mathcal{G}_i$ **core tensors**.

Definition 3.7. *A tensor $T \in \mathbb{V}_1 \otimes \cdots \otimes \mathbb{V}_n$ is a **tensor state of a tensor network** if it can be written as $T = \mathcal{C}_G(\mathcal{G}_1 \otimes \cdots \otimes \mathcal{G}_n)$ with $\mathcal{G}_i \in \xi_i$.*

Definition 3.8. *We will define the set of all possible tensor states of a tensor network as the set* $\mathrm{TNS}(G; \mathbb{E}_1, \dots, \mathbb{E}_c, \mathbb{V}_1, \dots, \mathbb{V}_n)$*, i.e*

$$\mathrm{TNS}(G; \mathbb{E}_1, \dots, \mathbb{E}_c, \mathbb{V}_1, \dots, \mathbb{V}_n) := \left\{ \kappa_G(\mathcal{G}_1 \otimes \cdots \otimes \mathcal{G}_n) \in \mathbb{V}_1 \otimes \cdots \otimes \mathbb{V}_n : \mathcal{G}_i \in \xi_i \right\}$$

The vector spaces $\mathbb{E}_1, \dots, \mathbb{E}_c, \mathbb{V}_1, \dots, \mathbb{V}_n$ are not really important on the Penrose representation of the tensor network, and we will usually not write them. Also, since all vector spaces are determined up to isomorphism by its dimension, we will write the tensor network as $\mathrm{TNS}(G; R_1, \dots, R_c, N_1, \dots, N_n)$.

And since n is equal to the number of vertices of G and c is equal to the number of edges of G, we will write $\mathrm{TNS}(G; R)$ for a more compact notation.

Definition 3.9. *Given a tensor network state* $\mathrm{TNS}(G, R)$*, from the graph given by its Penrose Notation, we will call the edges with a dangling end free edges, and the edges that connect two vertex contracted edges*

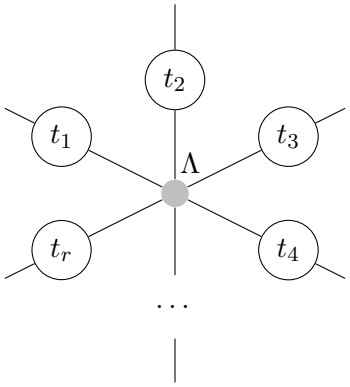
For example, in fig. 3.7, the edges labeled as $\mathbb{V}_1, \mathbb{V}_2, \mathbb{V}_3, \mathbb{V}_4$ are free edges and the rest are contracted edges.

3.3 Common Tensor network structures

Example 3.10. [Canonical polyadic decomposition] As we have seen, canonocal polyadic decomposition consists on decomposing T as a sum of r 1-rank tensors. We can write

$$\sum_{p=1}^{r} \lambda_p v_i^1 \otimes v_i^2 \otimes \cdots \otimes v_i^n$$

As the contraction of a diagonal tensor Λ of order r with its enties being λ_p and the other 1-rank tensors. If we denote $t_i = v_i^1 \otimes v_i^2 \otimes \cdots \otimes v_i^n$ so that the decomposition can be written as $\sum_{p=1}^{r} \lambda_p t_p$, the resulting tensor network has the following star shape:



Example 3.11. A tensor train decomposition or matrix product state of T are a set of 3th-order tensors $\mathcal{G}_1, \mathcal{G}_2, \ldots, \mathcal{G}_n$ with $\mathcal{G}_i \in \mathbb{K}^{R_{i-1} \times N_i \times R_i}$ and $R_0 = R_n = 1$ such that every element of T is written in the form

$$T(i_1, i_2, \ldots, i_n) = \sum_{r_0, \ldots, r_n}^{R_0, \ldots, R_n} \mathcal{G}_1(r_0, i_1, r_1)\mathcal{G}_2(r_1, i_2, r_2) \cdots \mathcal{G}_n(r_{n-1}, i_n, r_n) \tag{3.3.1}$$

We denote $R_0, R_1, \ldots, R_n$ as the ranks of the tensor train decomposition, or TT-ranks.

We can easily see that the tensor train decomposition (or TT) is obtained by our definition of a tensor network when G is a path, also the contraction of the whole network yields eq. (3.3.1)

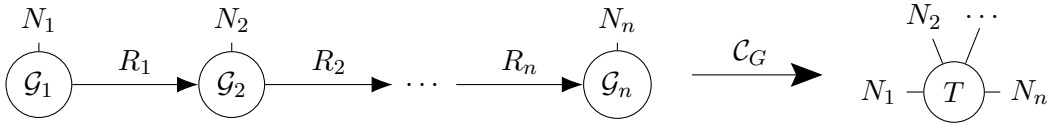


Figure 3.8: Tensor Train decomposition. Source: own elaboration.

Tensor train decompositions are very well studied because they are a chain of products of matrices, there are a lot of mathematical theory that can be applied in this structure,

for example, for finding the cores of the TT decomposition one could fix a core, reshape and unfold onto a matrix the objective tensor T, apply then SVD and update the cores according to the decomposition. A more detailed algorithm can be found in [10]. Since in this thesis is more general and not focused on tensor train decompositions we will skip this algorithm.

Example 3.12. Tensor ring decomposition (or TR) or also known a matrix product state with periodic boundary conditions, is obtained when G is a cycle.

Tensor Ring decomposition is considered generalization of Tensor Train decomposition, it's contraction is the same as eq. (3.3.1) but removing the condition $R_0 = R_1 = 1$. Zhao et. al. explain that Tensor Rings have also been widely studied because of their circular invariance: shifting the cores on the tensor network results in an equivalent tensor network that once its evaluated results in the objective tensor with its dimensions fixed. Thanks to this fact, one can still apply SVD decomposition on Tensor Rings [11].

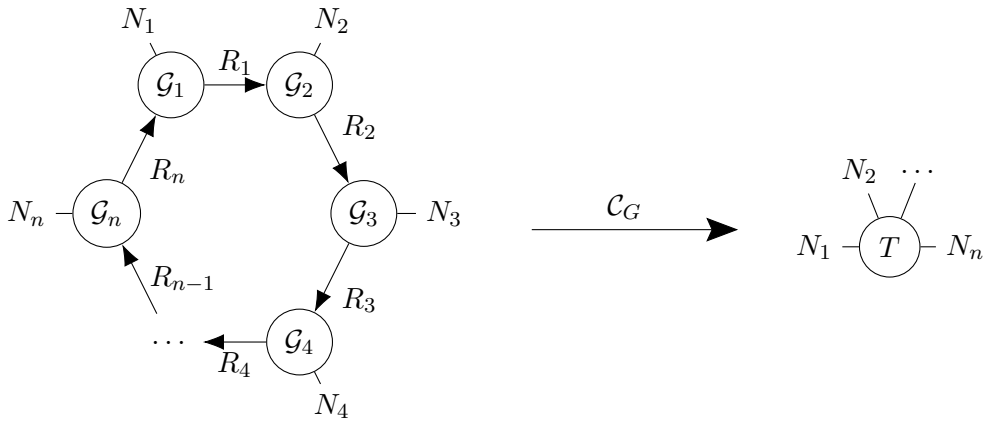


Figure 3.9: Tensor Ring (TR) decomposition. Source: own elaboration.

Example 3.13. The fully connected tensor network decomposition is obtenied when G is a complete graph.

3.4 Tensor Network Ranks

In this section we will present the notion of G-rank of a tensor. The G-rank of a tensor is defined as the minimum ranks that a tensor T can be represented by a tensor state. Knowing the G-rank of a tensor will be useful because we could find cores $\mathcal{G}_1, \ldots, \mathcal{G}_n$ such that we could represent T, and surely enough, with a enough sotisficated algorithm, it could converge to these core tensors. Our main goal will be to try to estimate the G-rank of a tensor, since finding the G-rank explicitly is a very hard task. Approximating the G-rank will be enough for us because in the next chapter we will slightly vary the approximated G-rank for finding a good representation.

As we defined the rank for a tensor T as the minimum number of 1-rank tensors that compose the tensor T, we will define the G-rank as the minimum ranks that a tensor network state needs for representing T. More formally,

Definition 3.14 (Tensor G-rank)**.** *Given a graph G, we define the tensor rank respect to*

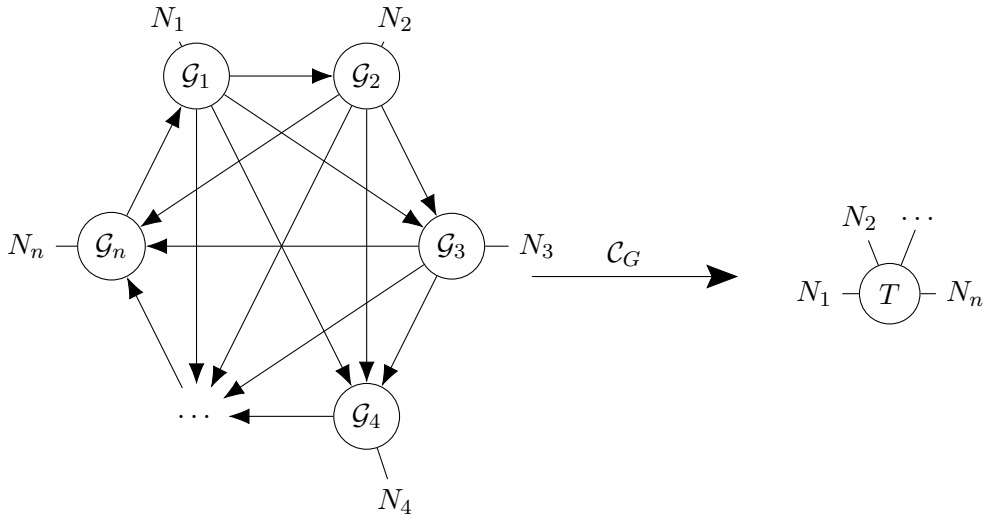


Figure 3.10: Fully connected tensor network decomposition (FCTN). Source: own elaboration.

*a G or G-**rank** as*

$$\operatorname{rank}_G(T) = \min\{(R_1, \ldots, R_c) \in \mathbb{Z}_+^c : T \in \operatorname{TNS}(G; R_1, \ldots, R_c, N_1, \ldots, N_d)\}$$

Where $\min(S)$ *with* $S \subset \mathbb{Z}_+^c$ *denotes the minimal elements of* S*. We treat* $\mathbb{Z}_+^c$ *with its usual partial order:*

$$(a_1, \ldots, a_c) \leqslant (b_1, \ldots, b_c) \iff a_1 \leqslant b_1, a_2 \leqslant b_2, \ldots, a_c \leqslant b_c$$

So for example if $S = \{(3,4,5), (2,1,3), (1,3,2)\}$*, then* $\min(S) = \{(2,1,3), (1,3,2)\}$

Now, we will see that $\operatorname{rank}_G(T)$ is a finite set. The following theorem gives us that if we make $R_1, \ldots, R_c$ big enough, every tensor T can be a state of $\operatorname{TNS}(G; R_1, \ldots, R_c, N_1, \ldots, N_n)$. In fact, these values that guarantee that T is an state are $R_1 = \ldots R_c = \operatorname{rank} T$

Theorem 3.15. *Let* $T \in \mathbb{V}_1 \otimes \cdots \otimes \mathbb{V}_n$ *and let* G *be a connected graph with* n *vertices and* c *edges. There exists* $R_1, \ldots, R_c \in \mathbb{Z}_+$ *such that*

$$T \in \operatorname{TNS}(G; R_1, \ldots, R_c, N_1, \ldots, N_d)$$

in fact, we can choose $R_1 = \cdots = R_c = \operatorname{rank} T$

Proof. Let $r = \operatorname{rank} T$. Then there exist $v_1^{(i)}, \ldots, v_r^{(i)} \in \mathbb{V}_i, i = 1, \ldots, n$ such that

$$T = \sum_{p=1}^{r} v_1^{(p)} \otimes \cdots \otimes v_n^{(p)}$$

We take $R_1 = \cdots = R_c = r$ we take for each $i = 1, \ldots, n$

$$\mathcal{G}_i = \sum_{p=1}^{r} \left(\bigotimes\nolimits_{j \in \operatorname{IN}(i)} e_p^{(j)} \right) \otimes v_p^{(i)} \otimes \left(\bigotimes\nolimits_{j \in \operatorname{OUT}(i)} e_p^{(j)*} \right)$$

Now observe that for each $i = 1, \dots, n$ there exists an unique h such that whenever $j \in \mathrm{IN}(i) \cap \mathrm{OUT}(i)$, $e_p^{(j)}$ and $e_p^{(j)*}$ contract and give δ_{pq}, therefore the summand vanishes except when $p = q$. This together with the assumption that G is a connected graph implies that $\kappa_G(\mathcal{G}_1 \otimes \cdots \otimes \mathcal{G}_n)$ reduces to a sum of terms of the form $v_p^{(1)} \otimes \cdots \otimes v_p^{(d)}$ for $p = 1, \dots, r$, which is of course T □

Now, by picking the ranks as $R_1 = \cdots = R_c = \operatorname{rank} T$ is usually not optimal, since we will end up that our tensor network will use more memory to represent T than the memory we need to write T itself.

We will proceed to show that the G-rank can be a lot more smaller than the tensor rank, and in fact, we will show some examples of tensor networks that can represent tensors in a more efficient way than the canonical polyadic decomposition that we have seen in chapter 2.

The following theorem says that there exists some tensor networks and some tensors that make this claim true:

Theorem 3.16. *For $n \geqslant 3$ there exists a connected simple graph G with n vertices and c edges such that there exists a tensor $T \in \mathbb{V}_1 \otimes \cdots \otimes \mathbb{V}_n$ where its tensor rank* $\operatorname{rank}(T) = r$ *is significantly larger than its G-rank* $\operatorname{rank}_G(T) = (r_1, \dots, r_n)$*. More specifically,*

$$r \gg r_1 + \cdots + r_n$$

The proof of this theorem can be found in [5]. In the paper the authors find that for representing a tensor network that contracts to a tensor $T \in (\mathbb{C}^{n\times n})^* \otimes (\mathbb{C}^{n\times n})^* \otimes \mathbb{C}^{n\times n} \cong \mathbb{C}^{n\times n\times n}$ picking G as C_3 and P_3 the inequality holds for $n = 3$. Then, the proof constructs from T for $n > 3$. This theorem shows that the G-rank can offer a more compact representation of tensors than the classical tensor rank. Nontheless, the theorem is not proven for generic tensor product spaces, in which proving that the G-rank could be smaller than the classical tensor rank are more difficult.

In this chapter we have presented tensor networks and tensor network states, and we have proved that sometimes this representations of a tensor can be more compact than the traditional rank decomposition. What remains now is how to find the cores $\mathcal{G}_1, \dots, \mathcal{G}_n$ of a given tensor network structure that give this compact decomposition. Which we will see in the following chapter.

Chapter 4

Tensor network state search

In this chapter we will study methods for finding optimal TNS such that we can represent an objective tensor T relatively close to a $\mathrm{TNS}(G;R)$, i.e that $T+E \in \mathrm{TNS}(G;R)$ where E is a tensor such that $\|E\|_F \simeq 0$.

Suppose that we have already found some optimal G and R for representing T. We will start the chapter by describing two algorithms for finding the cores $\mathcal{G}_1,\dots,\mathcal{G}_n$ that best represent T inside the state $\mathrm{TNS}(G;R)$. These two algorithms are the **alternated least squares** or **ALS** algorithm and the **backpropagation** algorithm.

Finally, we will focus on the problem of finding the optimal G and R, using the tensor network alternating local enumeration (TnALE) algorithm while proving its convergence.

4.1 Core search algorithms

We know by now that for any tensor T if we choose a graph $G=(V,E)$, we can choose ranks $R_1,\dots,R_c \leqslant \operatorname{rank} T \leqslant \left\lfloor \frac{\prod_{i=1}^n N_i}{\sum_{i=1}^n N_i} \right\rfloor$ such that $T \in \mathrm{TNS}(G,R)$. What we will do is try to reduce $R_1,\dots,R_c$. It is very likely that any reduction of the ranks makes impossible to fully represent T in $\mathrm{TNS}(G;R)$. What we will do instead is represent T as an approximation of a tensor that is inside of $\mathrm{TNS}(G;R)$

Suppose that we have defined a loss function $\pi_D : \mathbb{R}^{N_1\times N_2\times\cdots\times N_n} \to \mathbb{R}_+$ involving D, with D being a dataset. For example, if our objective is compressing an objective tensor $T_o \in \mathbb{R}^{N_1\times N_2\times\cdots\times N_n}$,

$$\pi_D(T) = \frac{\|T-T_o\|_F}{\|T_o\|_F} \tag{4.1.1}$$

With $D=\{T_0\}$. Both algorithms will seek to reduce the compression function (4.1.1). The backpropagation algorithm allows also to set D to a more general set, in a sense that it can be used to "train" our tensor network to represent desired outputs depending on different elements of D as the same way that neural networks do.

4.1.1 The Alternating Least Squares algorithm

In this subsection we will discuss the Alternating Least Squares algorithm that Malik et. al. present on section 5 of [12]

We will denote $T_0 \in \mathbb{K}^{N_1 \times \cdots \times N_n}$ as our objective tensor. Let $G = (V, E)$, $c = \#E$ and $R = (r_1, \ldots, r_c)$. We would want to find some cores $\mathcal{G}_1, \ldots, \mathcal{G}_c$ of $\mathrm{TNS}(G; R)$ such that they minimize

$$\|T_0 - \mathcal{C}_G(\mathcal{G}_1, \ldots, \mathcal{G}_c)\|_F$$

Which is equivalent on minimizing $\pi_{\{T_0\}}$ in (4.1.1).

If we impose that all the tensor cores remain fixed except $\mathcal{G}_m$, our problem would become

$$\arg\min_{\mathcal{G}_m} \|T - \mathcal{C}_G(\mathcal{G}_1, \ldots, \mathcal{G}_n)\|_F$$

Now, we apply the contractions defined by G (using the contraction mapping $\mathcal{C}_G$) for all cores excluding $\mathcal{G}_m$ (figs. 4.1 and 4.2). We will call this contracted tensor $\mathcal{G}^{\neq m}$. In case that we could cannot contract to tensors, i.e, while contracting we get more than one connected component as in fig. 4.2, we conduct the outer product of the contraction of each connected component.

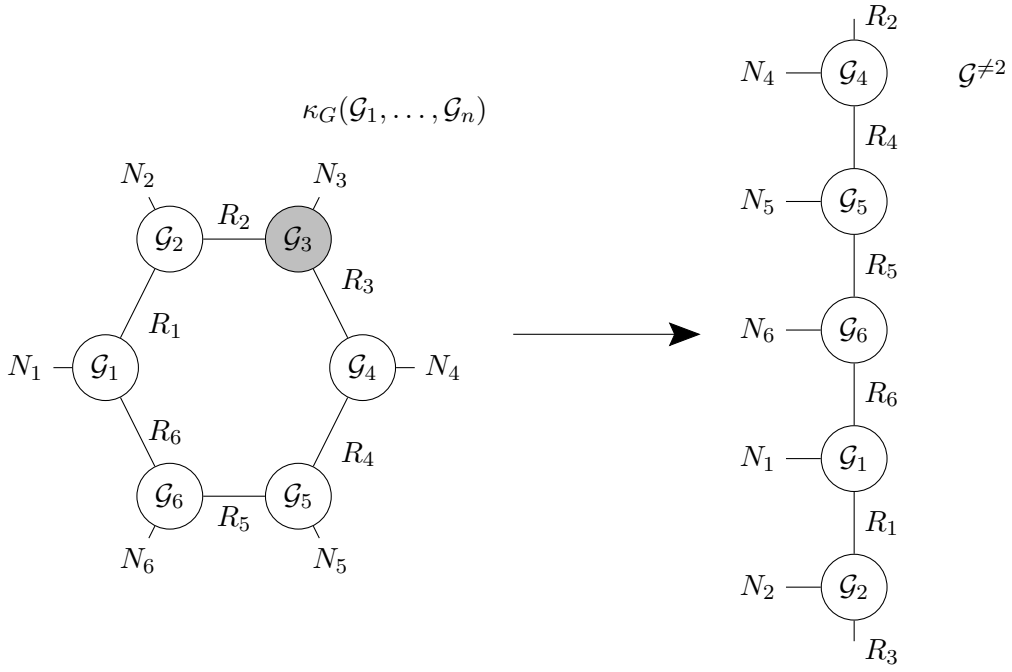


Figure 4.1: The representation of $\mathcal{G}^{\neq m}$ on the TR decomposition, with $m = 2$. Source: own elaboration.

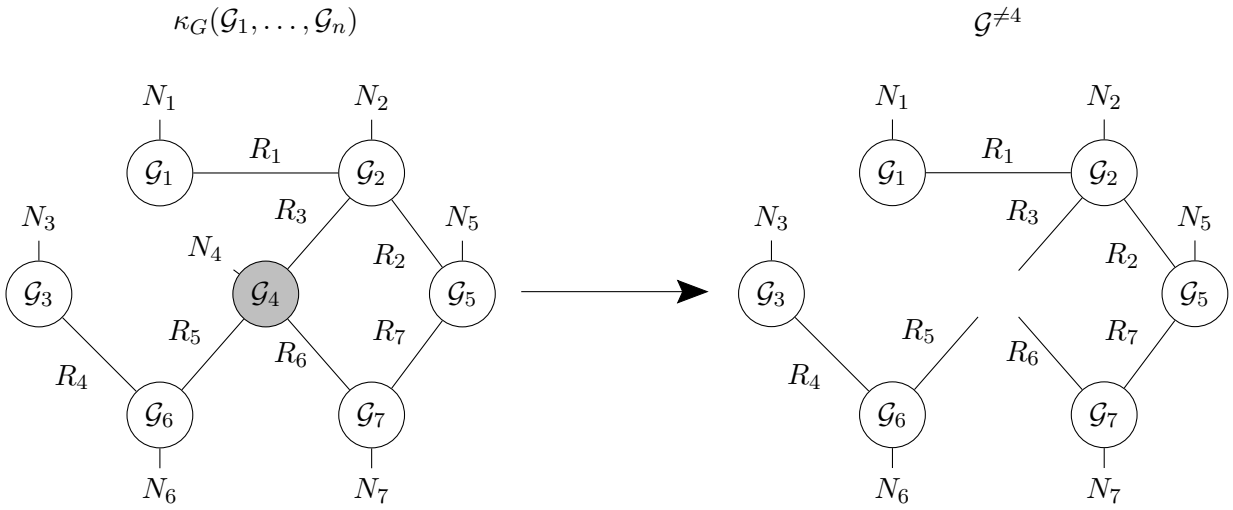


Figure 4.2: The representation of $\mathcal{G}^{\neq 4}$ over some arbitrary TN. Source: own elaboration.

By doing this, the last contractions that remain between $\mathcal{G}^{\neq m}$ and $\mathcal{G}_m$ are equivalent to evaluating the whole network. Now, if we consider appropiate matricizations $T^{(m)}$, $G^{\neq m}$ and G_m, these contractions can be computed calculating the matrix product $G^{\neq m}G_m$, so our minimization problem is now equivalent to solving the least squares problem. That is, given $A \in \mathbb{K}^{m\times n}$ and $B \in \mathbb{K}^{m\times k}$, find $X \in \mathbb{K}^{n\times k}$ such that

$$\arg\min_X \|AX - B\|_2 \tag{4.1.2}$$

In our case, $A = G^{\neq m}, X = G_m$ and $B = T^{(m)}$. Let $x^{(i)}$ be the i-th column of G_m and $y^{(i)}$ the i-th column of $T^{(m)}$. Solving 4.1.2 means solving for each i

$$\arg\min_{x^{(i)}} \|G^{\neq m}x^{(i)} - y^{(i)}\|_2 \tag{4.1.3}$$

Solving the least squares problem by using normal equations

Note that since we cannot guarantee that $(G^{\neq m})^{-1}$ is invertible, we cannot explicitly solve the linear system $G^{\neq m}x^{(i)} = y^{(i)}$. Also, since computing the inverse of a matrix is expensive, we will use the following proposition for solving (4.1.3):

Proposition 4.1. *If x satisfies $A^T(Ax - b) = 0$ then x is a solution to the least-squares problem, i.e, x minimizes $\|Ax - b\|_2$*

Proof. Let $y \in \mathbb{R}^n$ be any vector. Then,

$$\begin{aligned}
\|A(x+y) - b\|_2^2 &= [A(x+y) - b]^T[A(x+y) - b] \\
&= (Ax - b)^T(Ax - b) + 2(Ay)^T(Ax - b) + (Ay)^T(Ay) \\
&= \|Ax - b\|_2^2 + 2y^TA^T(Ax - b) + \|Ay\|_2^2 \\
&= \|Ax - b\|_2^2 + \|Ay\|_2^2 \\
&\geqslant \|Ax - b\|_2^2
\end{aligned}$$

□

Thanks to this proposition, any problem of the form $\arg\min_x \|Ax - b\|_2$ will have as a solution a vector x such that $A^TAx = A^Tb$. In our case, we get that solving eq. (4.1.3) is equivalent to solving the linear system

$$(G^{\neq m})^T G^{\neq m} x^{(i)} = (G^{\neq m})^T y^{(i)} \tag{4.1.4}$$

Nontheless, we can't still gurantee that this system has a solution since we don't know if $(G^{\neq m})^T G^{\neq m}$ is invertible. We could solve this inconvinience applying the Tikhonov regularization. Instead of solving eq. (4.1.4), we will solve the lineal system

$$\left((G^{\neq m})^T G^{\neq m} + \gamma I\right) \hat{x}^{(i)} = (G^{\neq m})^T y^{(i)}$$

where $\gamma > 0$ is called the **Tikhonov factor**. We will pick a small Tikhonov factor (usually $\gamma < 10^{-5}$) to assure that it doesn't smooth too much the solutions of the least squares problem.

Keep in mind that by adding any small perturbance to a linear system will not lead to the exact solution of $x^{(i)}$, but instead to a very close approximation $\hat{x}^{(i)}$. Therefore, the convergence of the algorithm can be affected by adding the Tikhonov factor.

Solving the least squares problem by the pseudoinverse of the SVD

Another approach on solving (4.1.2) that does not depend about if A is singular or not, is to apply the single value decomposition. The following method that we will describe is the same that uses the implementation of the function `torch.linalg.lstsq` of the PyTorch library that is used for solving the least squares problem [13] as a fallback (if the matrix A is well-conditioned, PyTorch uses more efficient methods such as QR factorization).

Given a matrix $A \in \mathbb{K}^{n\times m}$ we denote its conjugate transpose as A^*. If $\mathbb{K} = \mathbb{R}$ then $A^* = A^T$

Theorem 4.2 (SVD Factoriation)**.** *Any matrix $A \in \mathbb{K}^{n\times m}$ can be decomposed as*

$$A = U\Sigma V^*$$

Where $U \in \mathbb{K}^{n\times n}, V \in \mathbb{K}^{m\times m}$ are unitary matrices (i.e, $U^*U = UU^* = I_n$ and $V^*V = VV^* = I_m$) and $\Sigma \in \mathbb{K}^{n\times m}$ is a rectangular diagonal matrix, i.e:

$$\Sigma = \operatorname{Diag}(\sigma_1, \dots, \sigma_r, 0, \dots, 0)$$

*The entries of Σ are called **singular values** of A. They are usually ordered as $\sigma_1 \geqslant \sigma_2 \geqslant \cdots \geqslant \sigma_r \geqslant 0$ where $r = \operatorname{rank}(A)$ We call this factorization the **SVD factorization** of A.*

Proof. We can suppose that $A \neq 0$ because if $A = 0$ we could choose $U = V = I$ and $\Sigma = 0$. We will prove by induction over n. For $n = 1$, consider the matrix $A = (a) \neq 0$ with $a \in \mathbb{K}^m$, i.e A is a $m \times 1$ matrix. If we define $\sigma = \|a\|_2$ then there exists an ortogonal matrix $Q \in \mathbb{K}^{m\times m}$ such that $Q^*a = (\sigma, 0, \dots, 0)^T$, therefore

$$A = Q \begin{pmatrix} \sigma \\ 0 \end{pmatrix}$$

Therefore we can define the SVD of A as $U = Q$, $V = (1, \dots, 1)$ and $\Sigma = \begin{pmatrix} \sigma \\ 0 \end{pmatrix}$. Now, suppose that we can do the SVD factorization for any matrix of $n-1$ rows. Let $A \in \mathbb{K}^{m\times n}$,

$x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$ be two unitary vectors such that $Ax = \sigma y$ with $\sigma = \|A\|_2$. We know that there exists some matrices $V_2 \in \mathbb{K}^{n\times(n-1)}$ and $U_2 \in \mathbb{K}^{m\times(m-1)}$ such that $V = (xV_2) \in \mathbb{K}^{n\times n}$ and $U = (yU_2) \in \mathbb{K}^{m\times m}$ are ortogonals. Then,

$$U^*AV = \begin{pmatrix} y^* \\ U_2^* \end{pmatrix} A \begin{pmatrix} x & V_2 \end{pmatrix} = \begin{pmatrix} y^*Ax & y^*AV_2 \\ U_2^*Ax & U_2^*AV_2 \end{pmatrix}$$

Therefore if we call $B = U_2^*AV_2$ and $w^* = y^*AV_2$ then

$$U^TAV = \begin{pmatrix} \sigma & w^* \\ 0 & B \end{pmatrix} = A_1 \tag{4.1.5}$$

And since

$$\left\| A_1 \begin{pmatrix} \sigma \\ w \end{pmatrix} \right\|_2^2 \geqslant (\sigma^2 + w^*w)^2$$

Then we have that $\|A_1\|_2^2 \geqslant \sigma^2 + w^Tw$ and since $\|A\|_2^2 = \|A_1\|_2^2 =^2$ we get that $w = 0$ and replacing w in (4.1.5) we can find an SVD representation for the matrix A_1. □

Definition 4.3 (Moore-Penrose pseudoinverse)**.** *Let $A \in \mathbb{K}^{n\times m}$. The pseudoinverse of A is a matrix $A^+ \in \mathbb{K}^{m\times n}$ such that satisfies the following four criteria, known as the Moore-Penrose conditions [14]:*

1. $AA^+A = A$
2. $A^+AA^+ = A$
3. $(AA^+)^* = AA^+$
4. $(A^+A)^* = A^+A$

Where A^* denotes the conjugate transposition of A.

Now we make the claim that $\Sigma^+ = \text{Diag}\left(\frac{1}{\sigma_1}, \ldots, \frac{1}{\sigma_r}, 0, \ldots, 0\right) \in \mathbb{K}^{n\times m}$

$$\Sigma = \overbrace{\begin{pmatrix} \sigma_1 & 0 & \ldots & 0 \\ 0 & \sigma_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \sigma_r \\ 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 0 \end{pmatrix}}^{m} \Bigg\} n \qquad \Sigma^+ = \overbrace{\begin{pmatrix} \frac{1}{\sigma_1} & 0 & \ldots & 0 & 0 & \ldots & 0 \\ 0 & \frac{1}{\sigma_2} & \ldots & 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \frac{1}{\sigma_r} & 0 & \ldots & 0 \end{pmatrix}}^{n} \Bigg\} m$$

By plugging each matrix to each Moore-Penrose condition we can see that, in fact, Σ^+ is the pseudoinverse of Σ. The following Lemma will give us a way to compute the Pseudoinverse if we have an SVD factorization:

Lemma 4.4. *If $A = U\Sigma V^T$, then $A^+ = V\Sigma^+U^T$*

Proof. It is sufficient to check if the Moore-Penrose conditions are satisfied. By the properties of transpositions and using that $VV^T = V^TV = I_n$ and $U^TU = UU^T = I_m$:

1. $AA^+A = (U\Sigma V^T)(V\Sigma^+U^T)(U\Sigma V^T) = U\Sigma(V^TV)\Sigma^+(UU^T)\Sigma V^T = U\Sigma I_m\Sigma^+I_n\Sigma V^T = U\Sigma\Sigma^+\Sigma V^T = U\Sigma V^T = A$
2. $A^+AA^+ = (V\Sigma^+U^T)(U\Sigma V^T)(V\Sigma^+U^T) = V\Sigma^+(U^TU)\Sigma(V^TV)\Sigma^+U^T = V\Sigma^+\Sigma\Sigma^+U^T = V\Sigma^+U^T = A^+$
3. $(AA^+)^T = (U\Sigma V^TV\Sigma^+U^T)^T = (U\Sigma\Sigma^+U^T)^T = U\Sigma\Sigma^+U^T = AA^+$
4. $(A^+A)^T = (V\Sigma^+U^TU\Sigma V^T)^T = (V\Sigma^+\Sigma V^T)^T = V\Sigma^+\Sigma V^T = A^+A$

□

Now, for solving (4.1.2) we can use the pseudoinverse of the SVD decomposition of the matrix A:

Proposition 4.5. *$x = V\Sigma^+U^*b$ minimizes $\|Ax - b\|_2$*

Proof. By applying the SVD factorization of the matrix A we get:

$$\|Ax - b\|_2^2 = \|U\Sigma V^*x - b\|_2^2 = \|U^*(U\Sigma V^*x - b)\|_2^2 = \|\Sigma(V^*x) - U^*b\|_2^2$$

In the last equality we used that U is an unitary matrix and $\|Ux\|_2 = \|x\|_2$. Let $y = V^*x$ and $c = U^*b$. Then the problem of minimizing $\|Ax - b\|_2$ reduces to solving

$$\min_y \|\Sigma y - c\|_2^2$$

. Since Σ is a diagonal matrix, if we pick an y with each component y_i of the form

$$y_i = \begin{cases} \dfrac{c_i}{\sigma_i} & \text{if } \sigma_i \neq 0 \\ \lambda_i & \text{if } \sigma_i = 0 \end{cases}$$

With $\lambda_i \in \mathbb{K}$ being any value, then $\|\Sigma y - c\|_2^2 = 0$. We will pick $\lambda_i = 0$ for all i, since this will give us the vector y that minimizes $\|Ax - b\|_2$ with the least norm, thus we have that $y = \Sigma^+c$. By reversing our change of variables, we get that

$$x = Vy = V\Sigma^+c = V\Sigma^+U^*b$$

□

Once we have solved (4.1.2) with any of these two described methods, we can reshape back G_m to $\mathcal{G}_m$. We can iteratively change the varying core tensor $\mathcal{G}_i$ until the algorithm converges.

Algorithm 1 Tensor Network ALS

Input: A tensor network state $T \in \mathbb{K}^{N_1 \times \cdots \times N_n}$ and some fixed error ϵ
Output: Core tensors $\mathcal{G}_1, \ldots, \mathcal{G}_n$

Initialize tensors $\mathcal{G}_1, \ldots, \mathcal{G}_n$
while $\|T - \mathcal{C}_G(\mathcal{G}_1, \ldots, \mathcal{G}_n)\|_F > \epsilon$ **do**
    for $k = 1, \ldots, n$ **do**
        Compute $\mathcal{G}^{\neq k}$
        Compute the proper order of modes δ, σ for unfolding $\mathcal{G}^{\neq k}$
        $G^{\neq k} \leftarrow \text{unfold}(\mathcal{G}^{\neq k}, \sigma, \delta)$
        $T^{(k)} \leftarrow \text{unfold}(T, (k), (1, \ldots, k-1, k+1, \ldots, n))$
        $G_k \leftarrow \arg\min_{G_k} \|G^{\neq k} G_k - T^{(k)}\|_2$
        $\mathcal{G}_k \leftarrow \text{fold}(G_k)$ with a proper mode ordering
    end for
end while
return $\mathcal{G}_1, \ldots, \mathcal{G}_n$

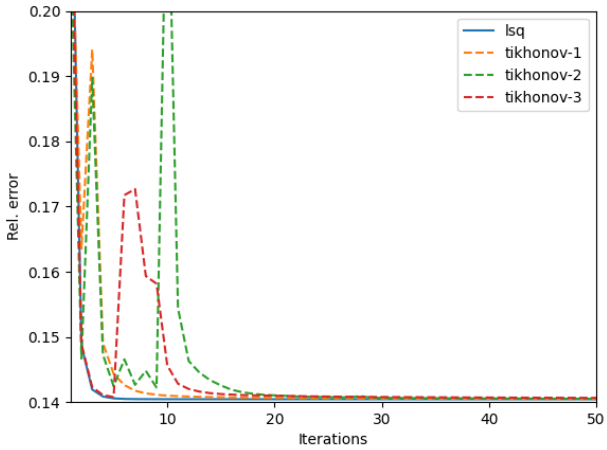
We will not enter in more detail about the explicit order of the modes when folding or unfolding the tensors. On the source code attached to this thesis, there is the complete implementation of this algorithm, detailing all the foldings and unfoldings.

As we can see, the ALS algorithm converges to a certain set of cores $\mathcal{G}_1^*, \ldots, \mathcal{G}_n^*$. If we denote $\Theta_k = \{\mathcal{G}_1^k, \ldots, \mathcal{G}_n^k\}$ as the set of cores $\mathcal{G}_i$ at the k-th iteration of the the algorithm, and if we denote $\mathcal{L}(\Theta_k) = \|G^{\neq k} G_k - T^{(k)}\|$ with $G^{\neq k}$, G_k and $T^{(k)}$ depending on Θ_k. Then, at each step we are minimizing $\mathcal{L}(\Theta_k)$, and so at the next iteration of the algorithm the loss function only can decrease, since by not varying the other core tensors we would at least stay at the same value of $\mathcal{L}(\Theta_k)$, therefore

$$\mathcal{L}(\Theta_{k+1}) \leqslant \mathcal{L}(\Theta_k)$$

Since the sequence $\{\mathcal{L}(\Theta_{(k)})\}_{k=0}^{\infty}$ is bounded by $0 \leqslant \mathcal{L}(\Theta_{(k)})$, therefore the algorithm converges to a fixed point Θ^*

Note that as we have previously said we can not guarantee the convergence of the algorithm if we use the Tikhonov regularization since it doesn't give the exact solution to the system in eq. (4.1.4), and therefore the solution might not solve the least squares problem. Nontheless, the Tikhonov regularization speeds up the computation, as we can see in the following figure:



	Time
lsq	4.47s
tikhonov-1	0.56s
tikhonov-2	0.58s
tikhonov-3	0.60s

Figure 4.3: Comparation between the ALS algorithm when the Tikhonov regularization is used and when not. The dashed lines represent different executions of the algorithm and the continuous line represents an iteration of the algorithm when using the least squares method for solving 4.1.3. The TN used is a TT with an objective tensor of shape $(8, 8, 16, 8, 8)$ and ranks $R = (8, 8, 8, 8)$. We can see that since in some cases when the linear system is near to zero, the Tikhonov regularization makes the relative error skyrocket. With enough iterations the Tikhonov regularization achieves the same relative error as the LSQ method. Source: own elaboration. Source: own elaboration.

4.1.2 Backpropagation

Since our problem of finding tensor cores is defined by minimizing a loss function, a valid idea would be to apply backpropagation.

In this subsection we will give another way of finding the tensor cores $\mathcal{G}_1, \ldots, \mathcal{G}_n$, but with backpropagation. The backpropagation algorithm is one of the core algorithms of deep learning. It consists on trying to find the variation that each parameter does to the loss function by exploiting the chain rule and then do some optimization with the gradient, such as gradient descent. We will now explain briefly the backpropagation algorithm in the context of tensor networks.

Suppose that we have a set D (a dataset), and a loss function $\pi : D \times \mathbb{R}^{N_1 \times N_2 \times \cdots \times N_n} \to \mathbb{R}_+$ The core idea behind backpropagation is to tune the parameters in a way that we minimize π_D. We do this by computing the gradient of all of the parameters of each $\mathcal{G}_i$ and moving the parameters in the same direction as the gradient.

In the case that $D = \{T_0\}$ the backpropagation problem is the same as finding the best TNS that represents T_0, and if for example, dataset has more tensors, it can be applied for example as a deep learning problem, as we will see later in chapter 5.

Let $d = T_0$. We want to compute the gradient of

$$\pi(T_0, \mathcal{C}_G(\mathcal{G}_1, \ldots, \mathcal{G}_n)) \tag{4.1.6}$$

Suppose that we reorder $\mathcal{G}_1, \ldots, \mathcal{G}_n$ in a sense that the contractions for evaluating the tensor network become first contracting $\mathcal{G}_n$ and $\mathcal{G}_{n-1}$, then the resulting tensor with $\mathcal{G}_{n-2}$

and so on. The function in 4.1.6 becomes

$$\pi(T_0, C(\mathcal{G}_1, \mathcal{C}(\mathcal{G}_2, \mathcal{C}(\ldots, \mathcal{C}(\mathcal{G}_{n-1}, \mathcal{G}_n)))) \tag{4.1.7}$$

Let $X_n = \mathcal{C}(\mathcal{G}_{n-1}, \mathcal{G}_n)$ and let $X_k = \mathcal{C}(\mathcal{G}_k, X_{k+1})$ for all $1 \leqslant k \leqslant n-1$. We can rewrite 4.1.7 as $\pi(T_0, X_1)$.

We would want to find the derivative of $\mathcal{G}_k$ respect of the loss function π. Consider the intermidiate contraction

$$X_k = \mathcal{C}(\mathcal{G}_k, X_{k+1})$$

Applying the chain rule we get

$$\frac{\partial \pi}{\partial \mathcal{G}_k} = \frac{\partial \pi}{\partial X_1} \cdot \frac{\partial X_1}{\partial X_2} \cdots \frac{\partial X_{k-1}}{\partial \mathcal{G}_k} \tag{4.1.8}$$

The partial derivative of the core tensor $\mathcal{G}_k$ respect to the loss function tells us in which

If we want to compute the derivative of the loss function respect to an individual element $\mathcal{G}_k^{i_1,\ldots,i_d}$ of the core tensor, using again the chain rule as before we get:

$$\frac{\partial \pi}{\partial \mathcal{G}_k^{i_1,\ldots,i_d}} = \sum_{b_1,\ldots,b_m} \frac{\partial \pi}{\partial X_k^{b_1,\ldots,b_m}} \cdot \frac{\partial X_k^{b_1,\ldots,b_m}}{\partial \mathcal{G}_i^{i_1,\ldots,i_d}} \tag{4.1.9}$$

So, we just have seen that by computing the partials of each intermidiate contraction we can obtain the partial derivative of the loss function respect to each individual parameter of the core tensor.

After knowing each derivative, we can compute the gradient of π in respect to all entries of the core tensors

$$\nabla \pi = \left(\frac{\partial \pi}{\partial \mathcal{G}_1}, \frac{\partial \pi}{\mathcal{G}_2}, \ldots, \frac{\partial \pi}{\partial \mathcal{G}_n} \right) \tag{4.1.10}$$

With each entry $\frac{\partial \pi}{\partial \mathcal{G}_i}$ being a tensor as the same shape of $\mathcal{G}_i$. We can reshape $\nabla \pi$ as a vector of the form

$$\nabla \pi = \left(\frac{\partial \pi}{\partial \mathcal{G}_1^{a_1,\ldots,a_p}}, \ldots, \frac{\partial \pi}{\partial \mathcal{G}_n^{b_1,\ldots,b_m}} \right) \tag{4.1.11}$$

Once we computed $\nabla \pi$, now we have an idea that in a local environment, how adjusting each entry of each core tensor affects on the change of the loss function π. Let $w = (\operatorname{vec} \mathcal{G}_1, \ldots, \operatorname{vec} \mathcal{G}_n)$ be a vector containing all the entries of the core tensors arrenged in order. Then, we can update each core tensor by "descending" by the gradient $\nabla \pi$ i.e updating w by a tiny amount on the direction that minimizes the value of the loss function

$$w^* = w - \eta \cdot \nabla \pi(w) \tag{4.1.12}$$

Where η is a small positive number called the learning rate. If we keep iterating as in 4.1.12, eventually w will reach a local minima of the loss function.

We implemented this algorithm using the PyTorch python library, since it has a built-in auto differentiation engine called `torch.autograd`. For for every operation that we apply to any PyTorch parameter, it remembers the operations that has been applied to the tensor and then it is able to compute the gradient of all of our computations [15].

On (CITE APPENDIX) we can find a comparison between the ALS and the backpropagation algorithms between some tensor network structures.

4.2 Structure search

On the last section we supposed that we had found some optimal G and R for representing T, and then we described some algorithms for finding the cores $\mathcal{G}_1, \ldots, \mathcal{G}_n$. In this section, we will try to find these optimal (G, R).

This problem is known as the *tensor network structure search problem* and its objective is to find the most optimal tensor network that compresses our objective T while maintaining the expressivity of the network, i.e that the network is capable to represent a closer result to the actual objective.

We will define the **loss function** $\mathcal{L} : \mathbb{G} \times \mathbb{F}_G \to \mathbb{R}_+$ as

$$\mathcal{L}(G, R) = \phi(G, R) + \lambda \cdot O(G, R)$$

Where $\mathbb{G}$ is the space of all simple connected graphs of n nodes, $R = (r_1, \ldots, r_K) \in \mathbb{F}_G \subseteq \mathbb{Z}_+^K$ are the ranks of the tensor network, $\phi(G, R)$ represents a function that determines the complexity of the tensor network, $\lambda > 0$ is a tuning parameter and $O : \mathbb{G} \times \mathbb{F}_G \to \mathbb{R}_+$ is the **evaluation function** defined as:

$$O(G, R) = \min_{\mathcal{Z} \in TNS(G,R)} \pi_D(\mathcal{Z}) \tag{4.2.1}$$

We will define the tensor network structure search problem as solving the following discrete optimization problem:

$$\min_{(G,R) \in \mathbb{G} \times \mathbb{F}_G} \mathcal{L}(G, R) \tag{4.2.2}$$

The intuition behind presenting this optimization problem this way is that since the evaluation function depends on the tuning parameter λ we can adjust a balance between giving priority to how expressive is the network or on simplifying its complexity. We will pich ϕ as the **complexity function** as a function $\phi : \mathbb{G} \times \mathbb{F}_G \to \mathbb{R}_+$ In our case, we will pick the inverse of the compression rate. I.e,

$$\phi(G, R) = \frac{\text{Size}(\mathcal{G}_1) \cdot \text{Size}(\mathcal{G}_2) \cdots \text{Size}(\mathcal{G}_n)}{\text{Size}(T)}$$

We call the problem 4.2.2 as Tensor Network Structure Selection (TN-SS). TN-SS is a very generalized problem. There has been a lot of research on solving it under certain conditions:

- Tensor Network Rank Selection (TN-RS), it restricts G to be a fixed graph, and its objective is to find the tensor network ranks $R \in \mathbb{F}_G$.
- Tensor Network Permutation Selection (TN-PS) fixes the ranks R and search over the set of all the simple graphs that are isomorphic to G.
- Lastly, Tensor Network Topology Selection (TN-TS) searches over the set of all simple graphs of N vertices and the tensor ranks R are fixed

In this chapter we will present the tensor network structure search algorithm using alternating local enumeration (TnALE) introduced by Li et. al. on [16].

4.2.1 The TnALE algorithm

The purpose of the TnALE algorithm is to solve 4.2.2 by locally alternating the discrete parameters that define (G, R).

We will start by describing the algorithm itself, and then we will present all the theory for proving its convergence.

The idea behind the TnALE algorithm is very simple: pick a random graph and random ranks G and R and then, for each iteration modifiy the rank R_1 a certain value and then pick the one that minimizes the loss function. Then do the same for R_2 and so on until R_c. Then, update G by adding or removing some edges and stay with the structure that is minimizes the loss function. Then modify $R_c, R_{c_1}, \dots R_2$ as before.

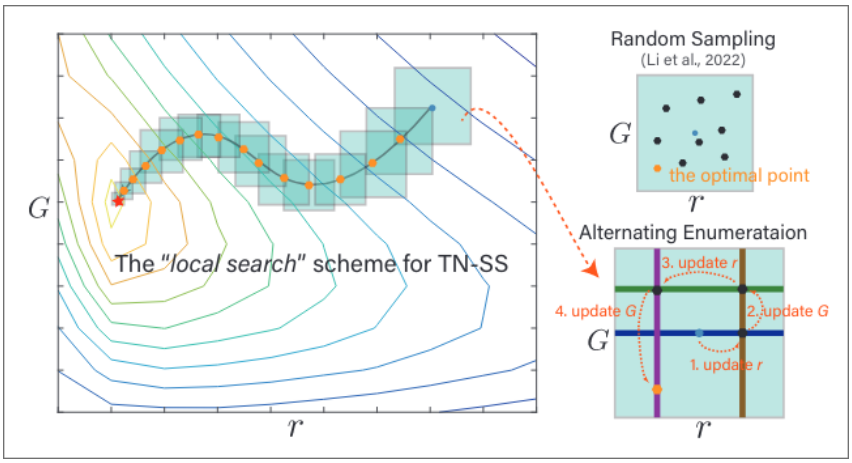


Figure 4.4: A representation of the TnALE algorithm and the difference of its sampling respect to TNLS [17]. r and G denote the structure related variables and each square represents a neighborhood of a given point in the iteration of TnALE. Source: [18]

Once this first pass is finished, the graph G is modified and each "local neighborhood" of the graph G is evaluated. Once again we save the best evaluation overwritting G and R.

Lastly, we do the first pass of varying the ranks $R_1, \dots, R_c$ but in the reverse order $R_c, \dots, R_1$. Doing this "round-trip" over the search of the ranks empirically results in a faster convergence rate to the optimal TNS structure.

$$\begin{array}{ccccccc} R_1 & \longrightarrow & R_2 & \longrightarrow & \dots & \longrightarrow & R_n \\ \uparrow & & & & & & \downarrow \\ R_2 & \longleftarrow & \dots & \longleftarrow & R_n & \longleftarrow & G \end{array}$$

Figure 4.5: The "round-trip" of updating the optimal structure. Source: own elaboration

In Alg. 2 the pseudocode of TnALE is given. During the next subsection, we will need to introducte some theory of gradient-less optimization for proving the convergence of the Tn-ALE algorithm.

Keep in mind that for each time we try to vary some rank R_k or the graph G we will need to evaluate (4.2.1), and it can be computationally very expensive. In the case that

Algorithm 2 Tensor Network Alternating Local Enumeration (Tn-ALE)

Input: A starting point $(G^{(0)}, R^{(0)})$ with $R^{(0)} = (R_1^{(0)}, \dots, R_K^{(0)})^T \in \mathbb{Z}_+^K$ and a rank-related radius $r \in \mathbb{Z}_+$ and the number of "round-trips" D

Output: The optimal (G, R) that minimizes eq. (4.2.2)

1: Initialize $(G, R) = (G^{(0)}.R^{(0)})$ with $R = (R_1, \dots, R_K)^T$
2: **for** $d = 1, \dots D$ **do**
3: **for** $k = 1, \dots K$ **do**
4: **for** $i = -r, \dots, 0, \dots r$ **do**
5: Copy $(\bar{G}, \bar{R}) \leftarrow (G, R)$
6: Update $(\bar{G}, \bar{R})$ by $\bar{R}_k = R_k + r$
7: $h(i) \leftarrow O(G, R)$ `# We can apply linear interpolation`
8: **end for**
9: Update (G, R) by $R_k = \arg\min_i h(i)$
10: **end for**
11: Take the neighborhood of G as $N(G)$
12: **for all** $G' \in N(G)$ **do**
13: Update (G, R) by $G = G'$
14: $h(G') \leftarrow O(G, R)$
15: **end for**
16: Update (G, R) by $G = \arg\min_{G'} h(G')$
17: **for** $k = K, K-1, \dots 2$ **do**
18: **for** $i = -r, \dots, 0, \dots r$ **do**
19: Copy $(\bar{G}, \bar{R}) \leftarrow (G, R)$
20: Update $(\bar{G}, \bar{R})$ by $\bar{R}_k = R_k + r$
21: $h(i) \leftarrow O(G, R)$ `# We can apply linear interpolation`
22: **end for**
23: Update (G, R) by $R_k = \arg\min_i h(i)$
24: **end for**
25: **end for**
26: **return** (G, R)

we are varying a rank (See steps 7 and 21 of Alg. 2), we can approximate $O(G, R)$ by doing a linear interpolation between evaluating at the points $R_k - r, R_k, R_k + r$. Since by increasing ranks empirically yields to a better representation of the objective tensor [18] and therefore we get less relative error (and the other way around is generally also true).

We will now formally prove that the ALS algorithm converges. For that, we will need to introduce the gradientless descent.

4.2.2 Gradientless optimization

What we want to see in this subsection is to extend the notion of the gradient onto discrete functions to then be able to apply a discrete gradient descent over our loss function.

We will start by rewriting the loss function (4.2.2) in a more general form. We define $\mathbb{P}$ as the set of all functions of the form $p : \mathbb{Z}_+^K \to \mathbb{K}_+^L$

$$\min_{x \in \mathbb{Z}_+^K, p \in \mathbb{P}} f_p(x) := f \circ p(x) \tag{4.2.3}$$

Where $f_p : \mathbb{Z}_+^L \to \mathbb{R}_+$ is a generalization of the loss function, and $p : \mathbb{Z}_+^K \to \mathbb{Z}_+^L \in \mathbb{P}$ correspond to the topology-related variable (G, R).

We can show that our tensor network structure search problem corresponds to a concrete case of this general form: since G is completly defined by its adjacency matrix A which is defined as

$$A_{ij} := \begin{cases} 1 & \text{if } (i,j) \in E \\ 0 & \text{otherwise} \end{cases}$$

We can encode the ranks R on the adjacency matrix by setting on each edge its corresponding rank:

$$A_{ij} := \begin{cases} R_{(i,j)} & \text{if } (i,j) \in E \\ 0 & \text{otherwise} \end{cases}$$

Where $R_{(i,j)}$ denotes the rank of the edge $(i.j)$. We will denote this matrix as A_R

Example 4.6. Let $G = P_4$ and $R = (1, 2, 3)$. We associate the edges of G with the ranks R as $R_{(1,2)} = r_1 = 1$, $R_{(2,3)} = r_2 = 2$ and $R_{(3,4)} = r_3 = 3$. We consider $\text{TNS}(G; R)$. The adjacency matrix of G and the encoded rank matrix of $\text{TNS}(G; R)$ is, respectively:

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \qquad A_R = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 2 & 0 \\ 0 & 2 & 0 & 3 \\ 0 & 0 & 3 & 0 \end{pmatrix}$$

Given G and R, we can encode $x \in \mathbb{R}_+^K$ as a vector with the rank of each edge of the graph and $p \in \mathbb{P}$ as a permutation matrix. By varying only p we get all the permutations of the graph G, and by x we get vary the ranks of the edges.

Example 4.7.

$$A_R^1 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 2 & 0 \\ 0 & 2 & 0 & 3 \\ 0 & 0 & 3 & 0 \end{pmatrix} \Leftrightarrow p_1(x) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 2 \\ 0 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 2 \\ 0 \\ 3 \end{pmatrix}$$

$$A_R^2 = \begin{pmatrix} 0 & 0 & 0 & 3 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 2 \\ 3 & 0 & 2 & 0 \end{pmatrix} \Leftrightarrow p_1(x) = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 2 \\ 0 \\ 3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 3 \\ 1 \\ 0 \\ 2 \end{pmatrix}$$

So, our problem is now a particular case of (4.2.3). We can now apply theory of gradientless optimizations in the real domain.

The goal of this section is to try to extend the gradient and its properties into discrete valued functions $f : \mathbb{Z}_+^L \to \mathbb{R}$. It is based on the theory that Golovin et. al. present on [19], and also on the discussion of the TnALE algorithm itself in the appendix of [18]. We

will end this chapter by proving that if we take some assumptions of f like being convex and also smooth we can prove that the TnALE algorithm can converge. For that, we will define what is a discrete gradient, what is α-strong convergence and (β_1, β_2)-smoothness, we will prove some properties about strong convergence and smoothness.

Before we begin, since we will be working with vectors of $\mathbb{Z}_+^L$, we will introduce a norm for them:

Definition 4.8. *The l^2 norm of a vector $x = (x_1, \ldots, x_n)^T$ is defined as*

$$\|x\|_2 = \sqrt{\sum_{i=1}^{n} |x_i|^2}$$

From now on, if x is a vector we will represent $\|x\|$ as $\|x\|_2$. We can now start by defining the concept of finite gradient:

Definition 4.9. *For any function $f : \mathbb{Z}_+^L \to \mathbb{R}$ its finite gradient $\nabla f : \mathbb{Z}_+^L \to \mathbb{R}$ at a point $x \in \mathbb{L}_+^L$ is defined as*

$$\Delta f(x) = [f(x + e_1) - f(x), \ldots, f(x + e_L) - f(x)]^T$$

With e_i being the unit vector of $\mathbb{Z}_+^L$ defined with the i-th entry being 1 and the rest of entries zeros.

Lemma 4.10. *Let $x \in \mathbb{Z}_+^L$. Then $\Delta\|x\|_2^2 = 2x + \mathbf{1}$ where $\mathbf{1}$ denotes the unit vector of $\mathbb{Z}_+^L$ that all of its entries are 1.*

Proof.

$$\Delta\|x\|_2^2 = \Delta\left(\sum_{i=1}^{L} |x_i|^2\right) = \sum_{i=1}^{L} \Delta|x_i|^2 = \sum_{i=1}^{L} 2x_i + e_i = 2x + \mathbf{1}$$

Where in the penultimate equality we used that

$$\Delta|x_i|^2 = |x_i + e_i|^2 - |x_i|^2 = |x_i|^2 + 2\langle x_i, e_i\rangle + |e_i|^2 - |x_i|^2 = 2\langle x_i, e_i\rangle + e_i = 2x_i + e_i$$

□

Now we will extend the strong convexity using the finite gradient. Recalling, a differentiable function $f : X \subseteq \mathbb{R}^n \to \mathbb{R}$ is convex if and only if for all $x, y \in X$

$$f(y) \geqslant f(x) - \langle \Delta f(x), y - x\rangle$$

Where $\Delta f(x)$ denotes the gradient of f. If f also satisfies that

$$f(y) \geqslant f(x) + \langle \Delta f(x) - \frac{\alpha}{2}\mathbf{1}, y - x\rangle + \frac{\alpha}{2}\|y - x\|_2^2 \tag{4.2.4}$$

For some $\alpha \geqslant 0$ e say that f is α-strongly convex.

Definition 4.11. *We say that $f : \mathbb{Z}_+^L \to \mathbb{R}$ is α-strongly convex with $\alpha \geqslant 0$ if it satisfies eq. (4.2.4) with $\Delta f(x)$ being the finite gradient for all $x, y \in \mathbb{Z}_+^L$ if the inequality holds for $\alpha = 0$, we will say that f is convex.*

We will prove now some properties of α-strongly convex functions. The following lemma gives us the equivalence between α-stronglt convex functions and convex functions:

Lemma 4.12. *$f : \mathbb{Z}_+^L$ is an α-strongly convex function iff $g(x) = f(x) - \frac{\alpha}{2}\|x\|_2^2$ is convex $\forall x \in \mathbb{Z}_+^L$*

Proof. ($\Rightarrow$) The first statement is equivalent to proving

$$g(y) \geqslant g(x) + \langle \Delta g(x), y - x\rangle \quad \forall x, y \in \mathbb{Z}_+^L \tag{4.2.5}$$

Applying that $f(x)$ is α-strongly convex we get that

$$\begin{aligned} g(y) - g(x) - \langle \Delta g(x), y - x\rangle &= f(y) - \frac{\alpha}{2}\|y\|_2^2 - f(x) + \frac{\alpha}{2}\|x\|_2^2 - \langle \Delta g(x), y - x\rangle \\ &= f(y) - \frac{\alpha}{2}\|y\|_2^2 - f(x) + \frac{\alpha}{2}\|x\|_2^2 - \langle \Delta f(x) - \alpha(2x + \mathbf{1}), y - x\rangle \\ &= f(y) - f(x) - \langle \Delta f(x) - \frac{\alpha}{2}\mathbf{1}, y - x\rangle - \frac{\alpha}{2}\|y\|_2^2 + \frac{\alpha}{2}\|x\|_2^2 + \frac{\alpha}{2}\langle 2x, y - x\rangle \\ &\geqslant \frac{\alpha}{2}\|y - x\|_2^2 - \frac{\alpha}{2}\|y\|_2^2 + \frac{\alpha}{2}\|x\|_2^2 + \frac{\alpha}{2}\langle 2x, y - x\rangle \\ &= \frac{\alpha}{2}\left(\|y - x\|_2^2 - \|y\|_2^2 - \|x\|_2^2 + 2\langle x, y\rangle\right) = 0 \end{aligned}$$

($\Leftarrow$) From eq. (4.2.5) we get

$$\begin{aligned} & f(y) - \frac{\alpha}{2}\|y\|_2^2 \geqslant f(x) - \frac{\alpha}{2}\|x\|_2^2 + \langle \Delta f(x) - \frac{\alpha}{2}(2x + \mathbf{1}), y - x\rangle \\ \Leftrightarrow\ & f(y) \geqslant f(x) - \frac{\alpha}{2}(\|x\|_2^2 + \|y\|_2^2) + \langle \Delta f(x) - \frac{\alpha}{2}\mathbf{1}, y - x\rangle - \frac{\alpha}{2}\langle 2x, y - x\rangle \\ \Leftrightarrow\ & f(y) \geqslant f(x) + \langle \Delta f(x) - \frac{\alpha}{2}\mathbf{1}, y - x\rangle + \frac{\alpha}{2}\|y - x\|_2^2 \end{aligned}$$

The last inequality gives us that f is α-strongly convex. □

The following lemma tells us that α-strongly convex functions have a more strict inequality of the monotone gradient property (a function is convex iff its gradient is monotone) than convex functions. The second inequality is a result of applying the Cauchy-Swartz inequality to the first one:

Lemma 4.13. *Given f an α-strongly convex function in $\mathbb{Z}_+^L$, then:*

1. $\langle \Delta f(x) - \Delta f(y), x - y\rangle \geqslant \alpha\|x - y\|_2^2 \quad \forall x, y \in \mathbb{Z}_+^L$
2. $\|\Delta f(x) - \Delta f(y)\|_2 \geqslant \alpha\|x - y\|_2 \quad \forall x, y \in \mathbb{Z}_+^L$

Proof. Let $g(x)$ be the same as in Lemma 4.12. To prove (1), we know that

$$\langle \Delta g(x) - \Delta g(y), x - y\rangle \geqslant 0, \forall x, y \in \mathbb{Z}_+^L$$

by the monotone gradient property of the convexity, which is true in both continuous and discrete scenarios. By substituting $\Delta g(x)$ we get that

$$\langle \Delta f(x) - \frac{\alpha}{2}(2x + \mathbf{1}) - \Delta f(y) + \frac{\alpha}{2}(2y + \mathbf{1}), x - y\rangle \geqslant 0$$

Simplifying we obtain that for all $x, y \in \mathbb{Z}_+^L$

$$\langle \Delta f(x) - \Delta f(y), x - y \rangle \geqslant \alpha \|x - y\|_2^2 \tag{4.2.6}$$

For proving (2) Consider the following inequality:

$$\|\Delta f(x) - \Delta f(y)\|_2 \|x - y\|_2 \geqslant \langle \Delta f(x) - \Delta f(y), x - y \rangle \geqslant \alpha \|x - y\|_2^2$$

Where in the first inequality we used the Cauchy-Swartz inequality and the last inequality follows from (4.2.6) □

The following lemma tells us that if we have a β-Lipschitz discrete function, the gradient is bounded. This will be one of the main assumptions that we will make about our loss function.

Lemma 4.14. *If $\|f(x) - f(y)\|_2 \leqslant \beta \|x - y\|$ for all $x, y \in \mathbb{Z}_+^L$ then the norm of the finite gradient is bounded by β i.e $\|\Delta f(x)\|_\infty = \max(|\Delta f(x)_1|, \ldots, |\Delta f(x)_L|) \leqslant \beta$*

Proof. Denote $\Delta f(x)_i$ the i-th entry of $\Delta f(x)$. Then for all $1 \leqslant i \leqslant L$,

$$|\Delta f(x)_i| = |f(x + e_i) - f(x)| \leqslant \beta \|x + e_i - x\| = \beta$$

□

Definition 4.15. *We say f is (β_1, β_2)-smooth for $\beta_1, \beta_2 > 0$ if*

1. *$|f(x) - f(y)| \leqslant \beta \|x - y\|_2$ for all $x, y \in \mathbb{Z}_+^L$*
2. *The function $l(x) := \frac{\beta_2}{2} \|x\|_2^2 - f(x)$ is convex*

The first item makes f to a β_1-Lipschitz function implying the "continuity" of the function, while the second item ensures an upper bound to the change of the gradient, because of the following lemma:

Lemma 4.16. *If $l(x) = \frac{\beta_2}{2} \|x\|_2^2 - f(x)$ is convex, then $\forall x, y \in \mathbb{Z}_+^L$ the following inequality satisfies:*

$$\langle \Delta f(x) - \Delta f(y), x - y \rangle \leqslant \beta \|x - y\|_2^2$$

Proof. By the convexity of $l(x)$ we get that

$$l(y) \geqslant l(x) + \langle \Delta l(x), y - x \rangle$$

Similarly we get

$$l(x) \leqslant l(y) + \langle \Delta l(y), x - y \rangle$$

Summing the two inequalities, we get that

$$l(y) + l(x) \geqslant l(x) + l(y) + \langle \Delta l(y) - \Delta l(x), x - y \rangle$$

And therefore

$$\langle \Delta l(x) - \Delta l(y), x - y \rangle \geqslant 0$$

Applying $l(x)$ we get that

$$\langle \beta(2x + \mathbf{1}) - \Delta f(x) - \beta(2y + \mathbf{1}) + \Delta f(y), x - y \rangle \geqslant 0$$

And by simplifying the inequality we have that

$$\langle \Delta f(x) - \Delta f(y), x - y \rangle \leqslant \beta \|x - y\|_2^2$$

□

4.2.3 Proof of the convergence of TnALE

After presenting α-strongly convex functions and (β_1, β_2)-smooth functions, in this subsection we will now prove that by setting the assumptions that our original problem from 4.2.3 the function f_p is α-strongly convex and (β_1, β_2)-smooth with $0 \leqslant \alpha \leqslant \beta_1 \leqslant \beta_2 \leqslant 1$ and that the minimum point of the discrete gradient which will be a fixed point of the TnALE algorithm does in fact exists and that its gradient is bounded by a very small factor.

We will start by defining the sub-level set of a function at a point. This definition will be useful for later proving that the TnALE algorithm can descend:

Definition 4.17 (Sub-level set)**.** *The level set of f at a point $x \in \mathbb{Z}_+^L$ is defined as the set $\mathbb{L}_x(f) = \{y \in \mathbb{Z}_+^L : f(y) = f(x)\}$ The sub-level set of f at a point x is defined as the set $\mathbb{L}_x^{\downarrow} = \{y \in \mathbb{Z}_+^L : f(y) \leqslant f(x)\}$*

Now, with the following lemma we will see that we can always find a neighbourhood that is contained inside $\mathbb{L}_x^{\downarrow}$, this will mean that we would be able to find a succession of points $(x_n)_{n=1}^{\infty}$ such that each $f(x_i) \leqslant f(x_j)$ if $i > j$

Lemma 4.18. *Let $f : \mathbb{Z}_+^L \to \mathbb{R}$ be a α-strongly convex, (β_1, β_2)-smooth and that its minimum value $f(x^*)$ satisfies that $\|\frac{\beta_2}{2}\boldsymbol{1} - \Delta f(x^*)\| \leqslant \gamma$ where γ is a constant and $0 \leqslant \gamma \leqslant \alpha$. Then $\forall x \in \mathbb{Z}_+^L$ there exists a L-dimensional cube which is of edge length $\frac{2(\alpha-\gamma)}{\beta_2\sqrt{L}}\|x - x^*\|$, tangent at x and inside the sub-level set $\mathbb{L}_x^{\downarrow}(f)$*

The proof of this lemma can be found at lemma B.8 from the appendix of [18]

Lemma 4.19 (Convex combination in the discrete domain)**.** *Suppose that $q = \theta x + (1-\theta)y$ for all $x, y \in \mathbb{Z}_+^L$ and $\theta \in [0, 1]$ and that there is a $\hat{q} \in \mathbb{Z}_+^L$ such that if $\Lambda = q - \hat{q}$ and we suppose that f is α-strongly convex then*

$$\theta f(x) + (1-\theta)f(y) \geqslant f(\hat{q}) + \langle \Delta f(\hat{q}) - \frac{\alpha}{2}\boldsymbol{1}, \Lambda\rangle + \frac{\alpha}{2}\|\Lambda\|_2^2$$

Note that this lemma justifies that we can use the linear interpolation trick when evaluating tensor networks, since by picking $\hat{q}$ as the local minimum when varying one parameter of the rank R_i, since we are moving across one line all of our real evaluations that we are skipping will be bounded below our linear interpolation.

Proof. By the definition of the α-strong convexity we get

$$f(x) \geqslant f(\hat{q}) + \left\langle \Delta f(\hat{q}) - \frac{\alpha}{2}\mathbf{1}, x - \hat{q}\right\rangle + \frac{\alpha}{2}\|x - \hat{q}\|_2^2$$

$$f(y) \geqslant f(\hat{q}) + \left\langle \Delta f(\hat{q}) - \frac{\alpha}{2}\mathbf{1}, y - \hat{q}\right\rangle + \frac{\alpha}{2}\|y - \hat{q}\|_2^2$$

And therefore,

$$\begin{aligned}\theta f(x) + (1-\theta)f(y) &\geqslant f(\hat{q}) + \left\langle \Delta f(\hat{q}) - \frac{\alpha}{2}\mathbf{1}, \Lambda\right\rangle + \frac{\alpha}{2}(\theta\|x\|^2 + (1-\theta)\|y\|^2 + \|\hat{q}\|^2 - 2\langle q, \hat{q}\rangle)\\ &\geqslant f(\hat{q}) + \left\langle \Delta f(\hat{q}) - \frac{\alpha}{2}\mathbf{1}, \Lambda\right\rangle + \frac{\alpha}{2}(\|q\|^2 + \|\hat{q}\|^2 - 2\langle q, \hat{q}\rangle)\\ &= f(\hat{q}) + \left\langle \Delta f(\hat{q}) - \frac{\alpha}{2}\mathbf{1}, \Lambda\right\rangle + \frac{\alpha}{2}\|\Lambda\|^2\end{aligned}$$

□

Theorem 4.20. *Let $f : \mathbb{Z}_+^K \to \mathbb{R}_+$ is α-strongly convex, (β_1, β_2)-smooth and let the minimum of eq. (4.2.3) be (p^*, x^*). Assume that f satisifies $\|\Delta f_{p^*}(x^*) - \frac{\beta_2}{2}\mathbf{1}\|_2 \leqslant \gamma$ where $0 \leqslant \gamma < \alpha \leqslant \beta_1 \leqslant \beta_2 \leqslant 1$.*

Then, if we let p fixed as p^*, and $0 \leqslant \theta \leqslant 1$ and for any x with $\|x - x^*\|_\infty \leqslant c$ we can find a neighborhood $B_\infty(x, r_X)$ where $r_X \geqslant \theta c + \frac{1}{2}$ such that there exists an element $y \in B_\infty(x, r_X)$ satisfying

$$f_{p^*}(y) - f_{p^*}(x^*) \leqslant (1-\theta)(f_{p^*}(x) - f_{p^*}(x^*)) + \frac{7}{8}K$$

Proof. Since we are letting p to be fixed as p^*, the general problem $\min_{x \in \mathbb{Z}_+^K, p \in \mathbb{P}} f_p(x)$ can be simplified and equivalently rewritten as $\min_{x \in \mathbb{Z}_+^K} f(x)$ Where $f : \mathbb{Z}_+^K \to \mathbb{R}_+$ represents the objective function. By Lemma 4.19 we have

$$f(\hat{q}) - f(x^*) \leqslant (1-\theta)(f(x) - f(x^*)) + \left\langle \frac{\alpha}{2}\mathbf{1} - \Delta f(\hat{q}), \Delta \right\rangle - \frac{\alpha}{2}\|\Delta\|^2$$

Now we will see that there exists an element y inside a neighborhood $B(x, r_x)$ with $r_x \in \mathbb{R}_+$ such that y also belongs to the sub-level cube tangent at $\hat{q}$ from Lemma 4.18 so that the inequality $f(y) \leqslant f(\hat{q})$ holds. We will prove the existance of this point y by showing that the intersection between the sub-level cube tangent at $\hat{q}$ and $B(x, r_x)$ is not empty. For that we will find the distance between $\hat{q}$ and x satisfies that

$$\|x - \hat{q}\|_\infty = \|x - q + \Lambda\|_\infty \leqslant \|x - q\|_\infty + \|\Lambda\|_\infty = \theta\|x - x^*\|_\infty + \|\Lambda\|_\infty \leqslant \theta c + \frac{1}{2}$$

The last inequality follows from $\|\Lambda\| \leqslant \frac{1}{2}$ which holds because we can always find $\hat{q} \in \mathbb{Z}_+^K$ by rounding the entries of q to its closest integers. Therefore we have proven that if $r_x \geqslant \theta c + \frac{1}{2}$ the intersection of the sub-level cube tangent at $\hat{q}$ and $B(x, r_x)$ is not empty, proving the existance of the point y.

Lastly, if we take an element y from the intersection we get that

$$\begin{aligned}
f(y) - f(x^*) \leqslant f(\hat{q}) - f(x^*) &\leqslant (1-\theta)(f(x) - f(x^*)) + \left\langle \frac{\alpha}{2}\mathbf{1} - \Delta f(\hat{q}), \Lambda \right\rangle - \frac{\alpha}{2}\|\Lambda\|^2 \\
&\leqslant (1-\theta)(f(x) - f(x^*)) + \left|\left\langle \frac{\alpha}{2}\mathbf{1}, \Lambda \right\rangle\right| + |\langle \Delta f(\hat{q}), \Lambda \rangle| + \frac{\alpha}{2}\|\Lambda\|^2 \\
&\leqslant (1-\theta)(f(x) - f(x^*)) + \frac{\alpha}{4}K + \|\Delta f(\hat{q})\|_\infty \|\Lambda\|_1 + \frac{\alpha}{2}\|\Lambda\|^2 \\
&\leqslant (1-\theta)(f(x) - f(x^*)) + \frac{\alpha}{4}K + \frac{\beta_1}{2}K + \frac{\alpha}{8}K \\
&\leqslant (1-\theta)(f(x) - f(x^*)) + \frac{3\alpha + 4\beta_1}{8}K \\
&\leqslant (1-\theta)(f(x) - f(x^*)) + \frac{7}{8}K
\end{aligned}$$

The inequality of the fourth line follows from $\|\Delta f(x)\|_\infty \leqslant \beta_1$ and from $\|\Lambda\|_\infty \leqslant \frac{1}{2}$ and the last line comes from that $\alpha \leqslant \beta_1 \leqslant 1$ □

The assumption of the satisfaction of the inequality $\|\Delta f_{p^*}(x^*) - \frac{\beta_2}{2}\mathbf{1}\|_2 \leqslant \gamma$ implies that $\Delta f_{p^*}(x^*)$ should be sufficiently smaller than the constant vector $\frac{\beta_2}{2}\mathbf{1}$, which can be understood as a discrete version of the zero-gradient stationary points of the continuous domain.

The constant K that appears on the proof is due to the fact that $\|\Lambda\|_1 \leqslant K\|\Lambda\|_\infty \leqslant K/2$ and that $\|\Lambda\|_2 \leqslant K\|\Lambda\|_\infty \leqslant \sqrt{K}/2$. This gives us that the norms $\|\Lambda\|_1$ and $\|\Lambda\|_2$ can become larger as the dimension K grows, which it is inevitable except if $\|\Lambda\|_\infty = 0$, which implies as we defined Λ the conventional convex optimizitation over the continuous domain.

Keeping all of the assumption of f, we can prove that local sampling methods, in concrete, TnALE converges:

Corollary 4.21. *Let $(x_n)_{n=0}^\infty$ be a succession of $\mathbb{Z}_+^K$ with x_0 randomly chosen. Suppose that p^* is known and that x_n is equal to the y of Theorem 4.20 for all $n > 0$. Then, if $\Omega(1/K) \leqslant \theta \leqslant 1$, i.e θ scales at least as $1/K$, we have that*

$$\lim_{n\to\infty}(f_{p^*}(x_n) - f_{p^*}(x^*)) = O(1)$$

Proof. Let $C_K = (7/8)K$. By Theorem 4.20,

$$\begin{aligned} f_{p^*}(x_n) - f_{p^*}(x^*) &\leqslant (1-\theta)(f_{p^*}(x_{n-1}) - f_{p^*}(x^*)) + C_K \\ &\leqslant (1-\theta)^2(f_{p^*}(x_{n-2}) - f_{p^*}(x^*)) + C_K + C_K(1-\theta) \\ &\leqslant (1-\theta)^3(f_{p^*}(x_{n-3}) - f_{p^*}(x^*)) + C_K + C_K(1-\theta) + C_K(1-\theta)^2 \\ &\leqslant \dots \\ &\leqslant (1-\theta)^n(f_{p^*}(x_0) - f_{p^*}(x^*)) + C_K\sum_{m=1}^{n}(1-\theta)^{m-1} \end{aligned}$$

Using that $\Omega(1/K) \leqslant \theta \leqslant 1$ we get that

$$\lim_{n\to\infty}(f_{p^*}(x_n) - f_{p^*}(x^*)) \leqslant \lim_{n\to\infty}(1-\theta)^n(f_{p^*}(x_0) - f_{p^*}(x^*)) + C_K\sum_{m=1}^{n}(1-\theta)^{m-1}$$

$$= 0 + C_K\frac{1}{\theta} = K\frac{7}{8\theta} \leqslant O(1)$$

Since in the last equality the term θ scales at least as $1/K$, therefore everything does not scale depending in K or θ, making the limit converge to some constant. □

And with this corollary we have proven that the TnALE algorithm converges. On the last corollary we have taken in account that K could also grow during the limit. In our case K would be the fixed value of our dimension that depends on the number of nodes from the graph space $\mathbb{G}$ that we are considering (See Example 4.7).

In this chapter we have seen how to search optimal structures of tensor networks that guarantee that when we search the optimal cores these minimize a loss function π_D. For achieving this we have seen first that if we fix the tensor network structure to some graph and ranks (G, R) we can apply the TN-ALS algorithm or we can use backpropagation for obtaining the cores $\mathcal{G}_1, \dots, \mathcal{G}_n$ that minimize the loss function fixed (G, R).

Then, we have presented the TnALE algorithm, which objective is to give a general solution to the TN-RS, TN-PS and TN-TS problems, giving us an optimal tensor network structure (G, R) and saving some evaluations when testing for tensor network structure candidates.

Finally, we have proved that under certain assumptions about the loss function, the TnALE algorithm does converge to a certain structure. We have presented gradientless optimization theory to prove this fact.

On the following chapter we will see some applications of all of these algorithms combined. We will use them to compress some images and we will also use them for compressing the weights of fully connected neural network layers.

Chapter 5

Applications of TN low rank approximations

5.1 Image compression

An RGB image is a set of $n \times m$ pixels. Each pixel contains three channels (red, green and blue) in which a color is represented. Therefore, we can encode an image as three tensors $T_R, T_G, R_B \in \mathbb{R}^{n \times m}$ with each of its indexes being on the interval $[0, 1] \subset \mathbb{R}$.

First, we could make the order of the tensor T higher by reshaping accordingly the divisors of n or m. For example, if we have a 256×256 image, we could reshape T_R, T_G and T_B as tensors of $\mathbb{R}^{4 \times 8 \times 8 \times 4 \times 8 \times 8}$

Then after doing this reshaping, for each channel we could pick a graph $G^{(0)}$ with the number of nodes being equal as the order of the tensors. Following our example, G should have 6 nodes. We can also pick some initial ranks $R^{(0)} = (R_1, \ldots, R_c)$. Then we can apply the ALS algorithm for finding some cores $\mathcal{G}_1, \ldots, \mathcal{G}_6$ that once contracted following $G^{(0)}$ will yield to a good approximation of each tensor $\mathcal{T}_R, \mathcal{T}_G, \mathcal{T}_B$

We could also before applying the ALS algorithm, apply the TnALE algorithm for finding if there is a more optimal structure than $(G^{(0)}, R^{(0)})$ and then do again the ALS algorithm.

Following our example, for recovering the image format we reshape back the tensors $\mathcal{T}_R, \mathcal{T}_G, \mathcal{T}_B \in \mathbb{R}^{4 \times 8 \times 8 \times 4 \times 8 \times 8}$ into matrices of $\mathbb{R}^{256 \times 256}$.

5.2 Neural networks

A fully connected neural network is composed on 3 different parts: an input layer, a central part with some layers called hidden layers and an output layer. The idea of neural networks is that they recieve an encoded input through the input layer and this input will be successively processed by the hidden layers and then it will return a certain result through the output layer. (See fig. 1.1)

Let $n_1, \ldots, n_L$ be the number of neurons on each layer. Each layer first does a pre-activation stage when it takes into account the activation of the previous layer, it adds some weights to each connection and it adds some bias to some of them.

The preactivation of the entire layer is computed by:

$$z^{(l)} = W^{(l)}a^{(l-1)} + b^{(l)}$$

Where $z^{(l)} \in \mathbb{R}^{n_l}$ is the pre-activation of the layer, $W^{(l)} \in \mathbb{R}^{n_l \times n_{l-1}}$ is the weight matrix, $a^{(l-1)} \in \mathbb{R}^{l-1}$ is the activation of the previous layer and and $b^{(l)} \in \mathbb{R}^{n_l}$ is the bias of the layer l. Note that each element of the vector $z^{(l)}$ denotes the preactivation of one neuron of the layer l. Then, once the preactivation is computed, the activation of a neuron is defined by composing the preactivation with some non-linear activation function $\sigma : \mathbb{R} \to \mathbb{R}$:

$$a^{(l)} = \sigma(z^{(l)})$$

Where in $\sigma(z^{(l)})$ is the vector resulting of applying σ to each element of $z^{(l)}$. Some examples of common activation functions are the Sigmoid $S(x)$, the ReLU $R(x)$ and $T(x)$.

$$S(x) = \frac{1}{1+e^{-x}} \qquad R(x) = \max(x, 0) \qquad T(x) = \frac{1-e^{-2x}}{1+e^{-2x}}$$

So, a neural network can be described as a composite function

$$N(x;\Theta) = \sigma^{(L)}(W^{(L)} \cdot \sigma^{(L-1)}(W^{(L-1)} \dots \sigma^{(1)}(W^{(1)}x + b^{(1)}) + b^{(2)} \dots) + b^{(L)})$$

Where $\Theta = \{W^{(l)}, b^{(l)}\}_{l=1}^{L}$ denotes the set of all adjustable parameters of the network. By varying Θ we can change the outputs that $N(x;\Theta)$. The parameters start at some given values and then they are adjusted usually using the backpropagation algorithm, the idea behind the backpropagation algorithm is very similar as we have described in the Subsection 4.1.2: we will have a dataset D with a series of pairs of inputs and outputs (x_i, y_i) and we will define some loss function $\pi_D : \mathbb{R}^{n_l} \to \mathbb{R}_+$ that measures the error of the output of the neural network $N(x)$ respect of the desired output y_i.

Then, exploiting the chain rule we can compute the contribution of each parameter to the total loss function $\frac{\delta\pi}{\delta W_{i,j}^{(l)}}$ and $\frac{\delta\pi}{\delta b_i^{(l)}}$. While computing these derivatives, we start by the layer $l-1$ and then compute the layer $l-2$ until we get to the input layer of neurons, while we can reuse the partial derivatives of the previous computed parameters.

Once we have all the partial derivatives computed, we can get the gradient of all the parameters and then do gradient descent onto the direction of the gradients, i.e, we update the parameters as

$$W^{(l)} \leftarrow W^{(l)} - \eta\frac{\partial\pi_D}{\partial W^{(l)}}$$

$$b^{(l)} \leftarrow b^{(l)} - \eta\frac{\partial\pi_D}{\partial b^{(l)}}$$

Where η is the learning rate. The compression that we will do is that we will reshape the matrices $W^{(l)}$ as tensors of higher order (say for example of order d_l) and then represent these reshaped matrices as the element of a tensor network state $\mathrm{TNS}(G;R)$. Since obtaining the representation consists of evaluating the core tensors through the contraction mapping $\mathcal{C}(\mathcal{G}_1, \dots, \mathcal{G}_n)$, we can apply the chain rule and compute the partial derivative of the loss function respect to each parameter of each core. In this way we can find the cores of the tensor network while training the neural network.

Another approach on compressing the neural network is to compress the weight matrices $W^{(l)}$ once the neural network is already trained using the TN-ALS algorithm. On (TODO: Citar apendíce) we see the results of these two different methods.

Chapter 6

Conclusions

In this bachelor's thesis we have formally constructed the tensor product space, we reviewed the basics of tensor algebra, we have presented tensor networks and tensor network states and we have seen that it is posible for some tensor networks to express a tensor with less rank than its usual tensor rank. Then, we have given some algorithms for computing the cores of the tensor networks and we have also presented algorithms that search over the space of possible tensor networks for a optimal structure. Finally, we have applied these algorithms for compressing images and neural networks.

Further research on this topic would include implementing other algorithms that improve the tensor network structure search, such as finding the tensor structures using program synthesis which Zhang et. al. in [20] claim that is superior to TnALE. Finding how the structure found by this algorithm improves compression on neural networks would be a good experiment.

Another interesting topic that could be also researched is about representing big tensors by splitting its parameters into multiple low-dimension tensors and then applying the tensor network low-rank approximation with these, since we have found that computing the core tensors for big tensors is very computationally expensive.

Finally, I would like to add some words about how hard it has been as a total non initiated to write this thesis since all of it is based on recent research papers, and they are quite often not very explanatory and are very summarized. However, I liked researching on the topic of tensor networks, and I would like to do some more research on the topic of this thesis.

Bibliography

[1] Xiangbin Meng et al. "The Application of Large Language Models in Medicine: A Scoping Review". In: *iScience* 27.5 (May 17, 2024), p. 109713. ISSN: 2589-0042. DOI: 10.1016/j.isci.2024.109713. URL: https://www.sciencedirect.com/science/article/pii/S2589004224009350 (visited on 06/04/2025).

[2] Tong Nie, Jian Sun, and Wei Ma. *Exploring the Roles of Large Language Models in Reshaping Transportation Systems: A Survey, Framework, and Roadmap*. Mar. 27, 2025. DOI: 10.48550/arXiv.2503.21411. arXiv: 2503.21411 [cs]. URL: http://arxiv.org/abs/2503.21411 (visited on 06/04/2025). Pre-published.

[3] Alaa Abd-alrazaq et al. "Large Language Models in Medical Education: Opportunities, Challenges, and Future Directions". In: *JMIR Medical Education* 9.1 (June 1, 2023), e48291. DOI: 10.2196/48291. URL: https://mededu.jmir.org/2023/1/e48291 (visited on 06/04/2025).

[4] Román Orús. "Tensor Networks for Complex Quantum Systems". In: *Nature Reviews Physics* 1.9 (Sept. 2019), pp. 538–550. ISSN: 2522-5820. DOI: 10.1038/s42254-019-0086-7. URL: https://www.nature.com/articles/s42254-019-0086-7 (visited on 05/14/2025).

[5] Ke Ye and Lek-Heng Lim. *Tensor Network Ranks*. Feb. 9, 2019. DOI: 10.48550/arXiv.1801.02662. arXiv: 1801.02662 [math]. URL: http://arxiv.org/abs/1801.02662 (visited on 03/24/2025). Pre-published.

[6] Christopher Hillar and Lek-Heng Lim. *Most Tensor Problems Are NP-hard*. July 1, 2013. DOI: 10.48550/arXiv.0911.1393. arXiv: 0911.1393 [cs]. URL: http://arxiv.org/abs/0911.1393 (visited on 03/31/2025). Pre-published.

[7] Matthias Christandl, Asger Kjærulff Jensen, and Jeroen Zuiddam. "Tensor Rank Is Not Multiplicative under the Tensor Product". In: *Linear Algebra and its Applications* 543 (Apr. 2018), pp. 125–139. ISSN: 00243795. DOI: 10.1016/j.laa.2017.12.020. arXiv: 1705.09379 [math]. URL: http://arxiv.org/abs/1705.09379 (visited on 05/28/2025).

[8] Roger Penrose. "Applications of Negative Dimensional Tensors". In: *Combinatorial Mathematics and its Applications* (1971), pp. 221–244. URL: https://homepages.math.uic.edu/~kauffman/Penrose.pdf (visited on 03/04/2025).

[9] Steven R. White. "Density Matrix Formulation for Quantum Renormalization Groups". In: *Physical Review Letters* 69.19 (Nov. 9, 1992), pp. 2863–2866. DOI: 10.1103/PhysRevLett.69.2863. URL: https://link.aps.org/doi/10.1103/PhysRevLett.69.2863 (visited on 03/27/2025).

[10] Melven Röhrig-Zöllner, Jonas Thies, and Achim Basermann. "Performance of the Low-Rank Tensor-Train SVD (TT-SVD) for Large Dense Tensors on Modern Multi-Core CPUs". In: *SIAM Journal on Scientific Computing* 44.4 (Aug. 2022), pp. C287–C309. ISSN: 1064-8275, 1095-7197. DOI: 10.1137/21M1395545. arXiv: 2102.00104 [math]. URL: http://arxiv.org/abs/2102.00104 (visited on 05/28/2025).

[11] Qibin Zhao et al. *Tensor Ring Decomposition*. June 17, 2016. DOI: 10.48550/arXiv.1606.05535. arXiv: 1606.05535 [cs]. URL: http://arxiv.org/abs/1606.05535 (visited on 03/04/2025). Pre-published.

[12] Osman Asif Malik, Vivek Bharadwaj, and Riley Murray. *Sampling-Based Decomposition Algorithms for Arbitrary Tensor Networks*. Oct. 7, 2022. DOI: 10.48550/arXiv.2210.03828. arXiv: 2210.03828 [math]. URL: http://arxiv.org/abs/2210.03828 (visited on 04/30/2025). Pre-published.

[13] *PyTorch Documentation — PyTorch 2.7 Documentation*. URL: https://docs.pytorch.org/docs/stable/index.html (visited on 06/04/2025).

[14] R. Penrose. "A Generalized Inverse for Matrices". In: *Mathematical Proceedings of the Cambridge Philosophical Society* 51.3 (July 1955), pp. 406–413. ISSN: 1469-8064, 0305-0041. DOI: 10.1017/S0305004100030401. URL: https://www.cambridge.org/core/journals/mathematical-proceedings-of-the-cambridge-philosophical-society/article/generalized-inverse-for-matrices/5F4516D6B9989BB6563A4B267CC7D615 (visited on 06/03/2025).

[15] *Automatic Differentiation with Torch.Autograd — PyTorch Tutorials 2.7.0+cu126 Documentation*. URL: https://docs.pytorch.org/tutorials/beginner/basics/autogradqs_tutorial.html (visited on 06/03/2025).

[16] Chao Li et al. *Alternating Local Enumeration (TnALE): Solving Tensor Network Structure Search with Fewer Evaluations*. May 29, 2023. DOI: 10.48550/arXiv.2304.12875. arXiv: 2304.12875 [cs]. URL: http://arxiv.org/abs/2304.12875 (visited on 05/15/2025). Pre-published.

[17] Chao Li et al. *Permutation Search of Tensor Network Structures via Local Sampling*. June 14, 2022. DOI: 10.48550/arXiv.2206.06597. arXiv: 2206.06597 [cs]. URL: http://arxiv.org/abs/2206.06597 (visited on 03/24/2025). Pre-published.

[18] Chao Li et al. *Alternating Local Enumeration (TnALE): Solving Tensor Network Structure Search with Fewer Evaluations*. May 29, 2023. DOI: 10.48550/arXiv.2304.12875. arXiv: 2304.12875 [cs]. URL: http://arxiv.org/abs/2304.12875 (visited on 03/24/2025). Pre-published.

[19] Daniel Golovin et al. "Gradientless Descent: High-Dimensional Zeroth-Order Optimization". In: International Conference on Learning Representations. Sept. 25, 2019. URL: https://openreview.net/forum?id=Skep6TVYDB (visited on 06/03/2025).

[20] Zheng Guo et al. *Tensor Network Structure Search Using Program Synthesis*. Feb. 22, 2025. DOI: 10.48550/arXiv.2502.02711. arXiv: 2502.02711 [cs]. URL: http://arxiv.org/abs/2502.02711 (visited on 05/12/2025). Pre-published.

Appendix A

Experimental results