FRIEDRICH-SCHILLER-
UNIVERSITÄT
JENA

Exploring Einsum as a Universal Inference Language

Masterarbeit

zur Erlangung des akademischen Grades

Master of Science (M.Sc.)

im Studiengang Informatik

Friedrich-Schiller-Universität Jena

Fakultät für Mathematik und Informatik

von Maurice Wenig **geboren am** 11.08.1999 **in** Jena

Betreuer: Dr. Julien Klaus, Prof. Dr. Joachim Giesen

Abstract

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

Contents

1 Introduction

1.1 Some Section

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

1.2 Some Other Section

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

2 Einsum

Given two third-order tensors $A \in \mathbb{R}^{3\times4\times5}$ and $B \in \mathbb{R}^{3\times3\times5}$, and a vector $v \in \mathbb{R}^4$. Consider the following computation resulting in a matrix $C \in \mathbb{R}^{3\times3}$:

$$\forall i \in [3] : \forall j \in [4] : C_{ij} = \sum_{k=1}^{5} A_{ijk} B_{iik} v_j$$

The original Einstein-notation for summation removes redundant formalism ("boilerplate") from this expression:

$$C_{ij} = A_{ijk} B_{iik} v_j$$

where it is assumed that C is defined for all possible i, j. We sum over all indices that are not used to index the output. In this example, we therefore have to sum over all possible values of k, because it is not used to index C_{ij}. Note how it is clear what the shape of C is, because i and j were used to index the tensors A, B, and v, for which we defined the dimensions on every axis.

This notation is essentially the inspiration for Einsum, which might be apparent given the name Einsum. Einsum is just an adaptation of this style, which makes it easier to use in programming. With it, we can write the above expression like this:

$$C := (ijk, iik, j \to ij, A, B, v)$$

Through the following definition, we hope to clear up why this Einsum expression results in the computation above, and what computation is the result of a general Einsum expression.

Definition 1. Einsum expressions specify how several input tensors are combined into a single output tensor. Let $T^{(1)}, \ldots, T^{(n)}$ be our input tensors, where $T^{(i)}$ is an n_i-th order tensor for $i \in [n]$. The core of the Einsum expression are index strings. For this, we first need a collection of symbols S. The respective index string for a tensor $T^{(i)}$ is then just a tuple $\boldsymbol{s_i} \in S^{n_i}$, composed of symbols $s_{ij} \in S$ for $j \in [n_i]$. The index string that is right of the arrow ($\to$) belongs to the output tensor T and is referred to as output string $\boldsymbol{s_t}$.

In our example this could be $S = \{i, j, k\}$. The tensor $T^{(1)} = A$ has the index string $\boldsymbol{s_1} = ijk$, $T^{(2)} = B$ has $\boldsymbol{s_2} = iik$, $T^{(3)} = v$ has $\boldsymbol{s_3} = j$, and the output string is $\boldsymbol{s_t} = ij$. The individual symbols are $s_{11} = i$, $s_{12} = j$, $s_{13} = k$, $s_{12} = i$, $s_{22} = i$, $s_{23} = k$, $s_{31} = j$, $s_{t1} = i$, $s_{t2} = j$.

In order to refer to individual tensor axes, let us numerate them with $a_{ij} \in \mathbb{N}$, where a_{ij} denotes the j-th axis of the tensor i-th tensor $T^{(i)}$. The actual value of a_{ij} does not matter. These variables are just used as unique identifiers for the axes. Note that we can combine the axes to describe the set of axes of a tensor $T^{(i)}$ with $\boldsymbol{a_i} = \{a_{ij} \mid j \in [n_i]\}$. In contrast to $\boldsymbol{s_i}$, this is a set and not a tuple. This is because tensors can reuse symbols as indices, but cannot reuse axes.

The next step in the definition is to speak about axis sizes. If we want to iterate over shared indices, it is necessary that the axes, that these indices are used for, share the same size. In our example, A_{ijk} and v_j share the symbol $s_{12} = s_{31} = j$. This means that the respective axes a_{12} and a_{31} have to have the same size, which happens to be four. Let us express this formally.

Let $d_{ij} \in \mathbb{N}$ denote the size of the axis a_{ij} for $i \in [n], j \in [n_i]$. Then it must hold that $s_{ij} = s_{i'j'} \implies d_{ij} = d_{i'j'}$ for all $i, i' \in [n], j \in [n_i], j' \in [n_{i'}]$.

Therefore we can also denote the size of all axes, that a symbol $s \in S$ corresponds to, as $d_s := d_{ij}$ for all $i \in [n], j \in [n_i]$ with $s = s_{ij}$. Note that not all same size axes have to be assigned the same symbol. For instance a square matrix could have index strings $\boldsymbol{s} = (i, i)$ or $\boldsymbol{s} = (i, j)$.

The next step of the definition is figuring out which symbols are used for summation and which symbols are used for saving the result of the computation. In order to do this, it is useful to know which symbols are in an index string, because symbols can occur more than once in just one index string (as seen in B_{iik} in our example). Therefore, let $\sigma(\boldsymbol{s})$ denote the set with all symbols used in an index string $\boldsymbol{s}$. That is, in out example $\sigma(\boldsymbol{s_2}) = \sigma(iik) = \{i, k\}$.

All symbols to the right of the arrow ($\rightarrow$) are used as an index for the result of the computation. These symbols are called *free* symbols $F = \sigma(\boldsymbol{s_t})$. All other symbols used in the expression are called *bound* symbols $B = \bigcup_{i\in[n]} \sigma(\boldsymbol{s_i}) \setminus \sigma(\boldsymbol{s_t})$. The reasoning behind this name is, that these symbols are bound by the summation symbol in the original computation. In Einsum, we sum over all axes that belong to bound symbols. It follows that the multi-index space that we iterate over is $\mathcal{F} = \prod_{s\in F}[d_s]$ and the multi-index space we sum over is $\mathcal{B} = \prod_{s\in B}[d_s]$. In our example, the free symbols are $F = \{ij\}$ and the bound symbols are $B = \{k\}$. The multi-index space we iterate over is $d_i \times d_j = [3] \times [4]$. The multi-index space we sum over is $d_k = [5]$.

From the definition of $\mathcal{F}$, it follows that d_s has to be defined for all symbols $s \in F$. This means we have to add the constraint $\sigma(\boldsymbol{s_t}) \subseteq \bigcup_{i\in[n]} \sigma(\boldsymbol{s_i})$.

However, we do not use every symbol in the multi-index spaces to index every input tensor. Instead, we use the index strings $\boldsymbol{s}$ to index the tensor. To formally express this, we need a projection from a multi-index $(\boldsymbol{f}, \boldsymbol{b}) \in \mathcal{F}\times\mathcal{B}$ [1] to another multi-index, which includes only the indices, that are represented by the symbols used in $\boldsymbol{s}$, in the same order as present in $\boldsymbol{s}$. We denote this as $(\boldsymbol{f}, \boldsymbol{b}) : \boldsymbol{s}$. Notice how this still allows

[1] Here, we use $(\boldsymbol{f}, \boldsymbol{b})$ as the notation for concatenating the tuples $\boldsymbol{f}$ and $\boldsymbol{b}$. This means, $(\boldsymbol{f}, \boldsymbol{b})$ is not a tuple of multi-indices, but another multi-index.

duplication of indices given in $(\boldsymbol{f}, \boldsymbol{b})$. This is needed, as can be seen in our example for B_{iik}, where a multi-index, e.g. $(i = 1, j = 4, k = 2) \in \mathcal{F} \times \mathcal{B}$, is projected onto a different multi-index, by the index string iik. With this index string, the index that is represented by the symbol i is projected onto the first and second position, and the index that is represented by the symbol k is projected onto the third position. Therefore, the resulting multi-index is $(i = 1, j = 4, k = 2) : iik = (1, 1, 2)$.

In our example, we used the standard sum and multiplication as operators for computing our result. But with Einsum, we allow the more general use of any semiring $R = (M, \oplus, \odot)$. With this, we can finally define a general Einsum expression

$$T := (\boldsymbol{s_1}, \ldots, \boldsymbol{s_n} \to \boldsymbol{s_t}, T^{(1)}, \ldots, T^{(n)})_R$$

in terms of semiring operations. Namely, T is the $|\boldsymbol{s_t}|$-th order tensor

$$\forall \boldsymbol{f} \in \mathcal{F} : T_{\boldsymbol{f}:\boldsymbol{s_t}} = \bigoplus_{\boldsymbol{b} \in \mathcal{B}} \bigodot_{i=1}^{n} T^{(i)}_{(\boldsymbol{f},\boldsymbol{b}):\boldsymbol{s_i}}.$$

Because we also project the indices $\boldsymbol{f}$ with the output string $\boldsymbol{s_t}$, we allow to iterate over duplicate indices, e.g. $\mathrm{diag}(v) = (j \to jj, v)$. This leaves some entries of the result undefined. We define these entries to be the additive neutral element in the given semiring R. This may sound arbitrary at first, but will be useful later.

There are still some special cases which need to be considered. If there are no bound symbols in the expression, then the sum in the original definition would be empty. But the definition is still meaningful. It boils down to computing the product of the tensor entries, without summing over them. Therefore, if $F = \emptyset$, then

$$T := (\boldsymbol{s_1}, \ldots, \boldsymbol{s_n} \to \boldsymbol{s_t}, T^{(1)}, \ldots, T^{(n)})_R$$

results in the computation of a $|\boldsymbol{s_t}|$-th order tensor T with

$$\forall \boldsymbol{f} \in \mathcal{F} : T_{\boldsymbol{f}:\boldsymbol{s_t}} = \bigodot_{i=1}^{n} T^{(i)}_{\boldsymbol{f}:\boldsymbol{s_i}}.$$

If there are no free symbols, we will sum over all axes given by the symbols in the expression. Therefore, if $B = \emptyset$, then

$$T := (\boldsymbol{s_1}, \ldots, \boldsymbol{s_n} \to, T^{(1)}, \ldots, T^{(n)})_R$$

results in the computation of a scalar T with

$$T = \bigoplus_{\boldsymbol{b} \in \mathcal{B}} \bigodot_{i=1}^{n} T^{(i)}_{\boldsymbol{b}:\boldsymbol{s_i}}.$$

In case the semiring can be derived from the context, or if it is irrelevant, it can be left out from the expression.

All following examples use the standard semiring $R = (\mathbb{R}, +, \cdot)$.

- matrix-vector multiplication: Let $A \in \mathbb{R}^{m \times n}, v \in \mathbb{R}^n$. Then

$$A \cdot v = (ij, j \to i, A, v)$$

- matrix-matrix multiplication: Let $A \in \mathbb{R}^{m \times r}, B \in \mathbb{R}^{r \times n}$. Then

$$A \cdot B = (ik, kj \to ij, A, B)$$

- trace: Let $A \in \mathbb{R}^{n \times n}$. Then

$$\mathrm{trace}(A) = (ii \to, A)$$

- squared Frobenius norm: Let $A \in \mathbb{R}^{n \times n}$. Then

$$|A|_2^2 = (ij, ij \to, A, A)$$

- diagonal matrix: Let $v \in \mathbb{R}^n$. Then

$$\mathrm{diag}(v) = (i \to ii, v)$$

3 Nested Expressions

In practice, concatenations of operations come naturally, e.g. computing the squared norm of a matrix-vector product $|A \cdot v|_2^2$ for $A \in \mathbb{R}^{m \times n}, v \in \mathbb{R}^n$. This would lead to a nested Einsum expression $|A \cdot v|_2^2 = (i, i \rightarrow, (ij, j \rightarrow i, A, v), (ij, j \rightarrow i, A, v))$. This expression dictates the order of evaluating the expression. In the example of the norm, the expression $(ij, j \rightarrow i, A, v)$ has to be evaluated before squaring and summing over the results of this computation.

This is limiting, because the order of evaluation might not yield optimal runtime. This can be seen with a simple matrix-matrix-vector multiplication, which can be written as follows:

$$(A \cdot B) \cdot v = (ij, j \rightarrow i, (ik, kj \rightarrow ij, A, B), v)$$

which is clearly worse than the optimal contraction order

$$A \cdot (B \cdot v) = (ij, j \rightarrow i, A, (ij, j \rightarrow i, B, v))$$

for $A \in \mathbb{R}^{m \times r}, B \in \mathbb{R}^{r \times n}, v \in \mathbb{R}^n$.

But fortunately, all nested Einsum expressions can be compressed into a single Einsum expression, if they are computed over the same semiring. This leaves the path of contraction up to the implementation. In the following theorems, we assume that the computations are all over the same semiring $R = (M, \oplus, \odot)$.

3.1 Simple Nested Expressions

Theorem 1: For $i \in [m+n]$, let $T^{(i)}$ be an n_i-th order tensor with index strings $\boldsymbol{s_i} \in S^{n_i}$. Let $\boldsymbol{s_u}, \boldsymbol{s_v}$ be index strings. Let

$$U := (\boldsymbol{s_{m+1}}, \ldots, \boldsymbol{s_{m+n}} \rightarrow \boldsymbol{s_u}, T^{(m+1)}, \ldots, T^{(m+n)})$$

and

$$V := (\boldsymbol{s_1}, \ldots, \boldsymbol{s_m}, \boldsymbol{s_u} \rightarrow \boldsymbol{s_v}, T^{(1)}, \ldots, T^{(m)}, U)$$

where the bound symbols of the second Einsum expression share no symbols with the first Einsum expression. Then

$$V = (\boldsymbol{s_1}, \ldots, \boldsymbol{s_{m+n}} \rightarrow \boldsymbol{s_v}, T^{(1)}, \ldots, T^{(m+n)})$$

Proof. Let F, F', B, B' be the free and bound symbols of the second (outer) and first (inner) einsum expression respectively. W.l.o.g. they are all non-empty. From them we can derive $\mathcal{F}, \mathcal{F}', \mathcal{B}, \mathcal{B}'$ as in the definition. Then

$$
\begin{aligned}
& V = (\boldsymbol{s_1}, \dots, \boldsymbol{s_m}, \boldsymbol{s_u} \to \boldsymbol{s_v}, T^{(1)}, \dots, T^{(m)}, U) \\
\iff \forall \boldsymbol{f} \in \mathcal{F} : V_{\boldsymbol{f}:\boldsymbol{s_v}} &= \bigoplus_{\boldsymbol{b} \in \mathcal{B}} \bigodot_{i=1}^{m} T^{(i)}_{(\boldsymbol{f},\boldsymbol{b}):\boldsymbol{s_k}} \odot U_{(\boldsymbol{f},\boldsymbol{b}):\boldsymbol{s_u}} \\
&= \bigoplus_{\boldsymbol{b} \in \mathcal{B}} \bigodot_{i=1}^{m} T^{(i)}_{(\boldsymbol{f},\boldsymbol{b}):\boldsymbol{s_i}} \odot \left[\bigoplus_{\boldsymbol{b'} \in \mathcal{B}'} \bigodot_{i'=m+1}^{m+n} T^{(i')}_{(\boldsymbol{f},\boldsymbol{b},\boldsymbol{b'}):\boldsymbol{s_{i'}}} \right] \\
&= \bigoplus_{\boldsymbol{b} \in \mathcal{B}} \bigoplus_{\boldsymbol{b'} \in \mathcal{B}'} \bigodot_{i=1}^{m} T^{(i)}_{(\boldsymbol{f},\boldsymbol{b}):\boldsymbol{s_i}} \odot \bigodot_{i=m+1}^{m+n} T^{(i)}_{(\boldsymbol{f},\boldsymbol{b},\boldsymbol{b'}):\boldsymbol{s_i}} \\
&= \bigoplus_{\boldsymbol{b} \in \mathcal{B} \times \mathcal{B}'} \bigodot_{i=1}^{m+n} T^{(i)}_{(\boldsymbol{f},\boldsymbol{b}):\boldsymbol{s_i}} \\
\iff & V = (\boldsymbol{s_1}, \dots, \boldsymbol{s_{m+n}} \to \boldsymbol{s_v}, T^{(1)}, \dots, T^{(m+n)})
\end{aligned}
$$

where the third equality follows from the definition of U:

$$\forall \boldsymbol{f'} \in \mathcal{F}' : U_{\boldsymbol{f'}:\boldsymbol{s_u}} = \bigoplus_{\boldsymbol{b'} \in \mathcal{B}'} \bigodot_{i'=m+1}^{m+n} T^{(i')}_{(\boldsymbol{f'},\boldsymbol{b'}):\boldsymbol{s_{i'}}}$$

and from the fact, that the symbols in $\boldsymbol{s_u}$ are used in the outer expression as an input string, and in the inner expression as the output string, and therefore $F' \subseteq B \cup F$. Additionally, because of the stated requirement $(B \cup F) \cap B' = \emptyset$, the symbols representing $\boldsymbol{b'}$ do not clash with the symbols representing $(\boldsymbol{f}, \boldsymbol{b})$, and therefore $(\boldsymbol{f}, \boldsymbol{b}, \boldsymbol{b'}) : \boldsymbol{s_{i'}}$ is well-defined and projects on the same indices as $(\boldsymbol{f'}, \boldsymbol{b'}) : \boldsymbol{s_{i'}}$. The fourth equality follows from the distributivity in a semiring. □

This means that we can compress all nested Einsum expressions, where the output string of the inner expression, which is used to compute U, is exactly the same as the respective input string in the outer expression, when U is used as an input tensor. This is already helpful for naturally occuring expressions in linear algebra, e.g.

$$
\begin{aligned}
|A \cdot v|_2^2 &= (i, i \to, (ij, j \to i, A, v), (ij, j \to i, A, v)) \\
&= (ij, j, ij, j \to, A, v, A, v)
\end{aligned}
$$

for $A \in \mathbb{R}^{m \times n}, v \in \mathbb{R}^n$, or

$$
\begin{aligned}
A \cdot B \cdot v &= (ij, j \to i, (ik, kj \to ij, A, B), v) \\
&= (ik, kj, j \to i, A, B, v)
\end{aligned}
$$

for $A \in \mathbb{R}^{m \times r}, B \in \mathbb{R}^{r \times n}, v \in \mathbb{R}^n$. But sometimes, we need to access a different multi-index set than the one we computed, e.g.

$$\text{trace}(A \cdot B) = (ii \to, (ik, kj \to ij, A, B))$$

or

$$A \cdot \text{diag}(v) = (ik, kj \to ij, A, (i \to ii, v))$$

for $A \in \mathbb{R}^{m \times n}, B \in \mathbb{R}^{n \times m}, v \in \mathbb{R}^n$. For this, we need more general ways of compressing nested Einsum expressions.

3.2 Introducing Duplications

The following example is an expression, which we cannot compress with the previous theorem:

$$(ij, jjj \to i, A, (kl, lo \to kko, B, C))$$

for $A \in \mathbb{R}^{a \times b}, B \in \mathbb{R}^{b \times c}, C \in \mathbb{R}^{c \times b}$. In the following theorem, we will explore how to compress expressions such as this one. Note that, for the theorem, we use disjoined sets of symbols for the inner and outer expression. This helps in the proof, and is not a real constraint in practice, because we can just rename the symbols in different scopes. For example, we could also write the above expression as

$$(ij, jjj \to i, A, (ik, kj \to iij, B, C)).$$

Theorem 2: For $i \in [m+n]$, let $T^{(i)}$ be an n_i-th order tensor with index strings $\boldsymbol{s_i} \in S^{n_i}$. Let $\boldsymbol{s_u}$ be an index string for the n_u-th order tensor U, which is defined as follows:

$$U := (\boldsymbol{s_{m+1}}, \dots, \boldsymbol{s_{m+n}} \to \boldsymbol{s_u}, T^{(m+1)}, \dots, T^{(m+n)})$$

Also let $\boldsymbol{\hat{s}_u}$ be alternative index strings for U with $s_{uj} = s_{uj'} \implies \hat{s}_{uj} = \hat{s}_{uj'}$ for all $j, j' \in [n_u]$, which means that $\boldsymbol{\hat{s}_u}$ can only introduce new symbol duplications, and cannot remove any.

In our example, $\boldsymbol{s_u} = kko$ and $\boldsymbol{\hat{s}_u} = jjj$. This does not break the symbol duplication of the first and second index, and introduces a new duplication on the third index.

Let s_v be an index string and

$$V := (\boldsymbol{s_1}, \dots, \boldsymbol{s_m}, \boldsymbol{\hat{s}_u} \to \boldsymbol{s_v}, T^{(1)}, \dots, T^{(m)}, U)$$

such that the first and second Einsum expression share no symbols. Then these nested Einsum expressions can also be compressed into a single Einsum expression.

In contrast to Theorem 1, we cannot just replace the input index string $\boldsymbol{\hat{s}_u}$ by all the input index strings in the inner Einsum expression $\boldsymbol{s_{m+1}}, \dots, \boldsymbol{s_{m+n}}$. Instead, we first need to apply a symbol map. Let $\nu : S \to S$ such that

$$\nu(s) := \begin{cases} \hat{s}_{uj} & \text{if } \exists j \in [n_u] : s_{uj} = s \\ s & \text{else} \end{cases}$$

which maps symbols in $\boldsymbol{s_u}$ to the symbol at the same index in $\boldsymbol{\hat{s}_u}$ and all other symbols to themselves.

This symbol map holds information about which symbols will be iterated over at the same time in the outer expression. In our example, these are the important mappings:

$$k \to j$$
$$o \to j$$

This means that k and o will be iterated over at the same time.

ν can be extended, such that it maps entire index strings instead of just symbols, by setting $\nu(\boldsymbol{s_i}) \in S^{n_i}, \nu(\boldsymbol{s_i})_j := \nu(s_{ij})$. Then we can write the substituted index strings by setting $\boldsymbol{\hat{s}_i} := \nu(\boldsymbol{s_i})$ for $i \in [m+1, m+n]$.

Then the compressed Einsum expression is the following:

$$V = (\boldsymbol{s_1}, \dots, \boldsymbol{s_m}, \boldsymbol{\hat{s}_{m+1}}, \dots, \boldsymbol{\hat{s}_{m+n}} \to \boldsymbol{s_v}, T^{(1)}, \dots, T^{(m+n)})$$

which helps us to compress the example:

$$(ij, jjj \to i, A, (kl, lo \to kko, B, C)) = (ij, jl, lj \to i, A, B, C)$$

Proof. The fundamental idea behind this theorem is, that by using the index string $\boldsymbol{\hat{s}_u}$, we only iterate over a sub-space of the indices that we defined for the computation of U. To formulate this, we need some idea of which multi-indices we iterate over. Therefore, let $\mathcal{M} : \boldsymbol{s} := \{M : \boldsymbol{s} \mid M \in \mathcal{M}\}$ for an index string $\boldsymbol{s}$ and a multi-index space $\mathcal{M}$.

Let F, F', B, B' be the free and bound symbols of the second (outer) and first (inner) einsum expression respectively. W.l.o.g. they are all non-empty. From them we can derive $\mathcal{F}, \mathcal{F}', \mathcal{B}, \mathcal{B}'$ as in the definition. Then $(\mathcal{F} \times \mathcal{B}) : \boldsymbol{\hat{s}_u} \subseteq \mathcal{F}' : \boldsymbol{s_u}$. This follows from $d_{s_{uj}} = d_{\hat{s}_{uj}}$ for $j \in [n_u]$, and the amount of symbols in the projection of $(\mathcal{F} \times \mathcal{B}) : \boldsymbol{\hat{s}_u}$ being smaller or equal to the amount of symbols in the projection of $\mathcal{F}' : \boldsymbol{s_u}$. The first fact is true per the definition of einsum. The second fact can be rewritten as $|\sigma(\boldsymbol{\hat{s}_u})| \leq |\sigma(\boldsymbol{s_u})|$ and follows directly from the constraint $s_{uj} = s_{uj'} \implies \hat{s}_{uj} = \hat{s}_{uj'}$ for all $j, j' \in [n_u]$.

Then

$$\forall \boldsymbol{f'} \in \mathcal{F}' : U_{\boldsymbol{f'}:\boldsymbol{s_u}} = \bigoplus_{\boldsymbol{b'} \in \mathcal{B}'} \bigodot_{i=m+1}^{m+n} T^{(i)}_{(\boldsymbol{f'}, \boldsymbol{b'}):\boldsymbol{s_i}}$$

and therefore

$$\forall (\boldsymbol{f}, \boldsymbol{b}) \in \mathcal{F} \times \mathcal{B} : U_{(\boldsymbol{f}, \boldsymbol{b}):\boldsymbol{\hat{s}_u}} = \bigoplus_{\boldsymbol{b'} \in \mathcal{B}'} \bigodot_{i=m+1}^{m+n} T^{(i)}_{(\boldsymbol{f}, \boldsymbol{b}, \boldsymbol{b'}):\boldsymbol{\hat{s}_i}}$$

because of the previous observation, and because the bound symbols of the expression, which are used in $\boldsymbol{b'}$, do not occur in $\boldsymbol{s_u}$, and are therefore not changed by the symbol map ν.

Therefore

$$\begin{aligned}
& V = (\boldsymbol{s_1}, \ldots, \boldsymbol{s_m}, \boldsymbol{\hat{s}_u} \rightarrow \boldsymbol{s_v}, T^{(1)}, \ldots, T^{(m)}, U) \\
\iff \forall \boldsymbol{f} \in \mathcal{F} : V_{\boldsymbol{f}:\boldsymbol{s_v}} &= \bigoplus_{\boldsymbol{b} \in \mathcal{B}} \bigodot_{i=1}^{m} T^{(i)}_{(\boldsymbol{f},\boldsymbol{b}):\boldsymbol{s_i}} \odot U_{(\boldsymbol{f},\boldsymbol{b}):\boldsymbol{\hat{s}_u}} \\
&= \bigoplus_{\boldsymbol{b} \in \mathcal{B}} \bigodot_{i=1}^{m} T^{(i)}_{(\boldsymbol{f},\boldsymbol{b}):\boldsymbol{s_i}} \odot \bigoplus_{\boldsymbol{b'} \in \mathcal{B}'} \bigodot_{i'=m+1}^{m+n} T^{(i')}_{(\boldsymbol{f},\boldsymbol{b},\boldsymbol{b'}):\boldsymbol{\hat{s}_{i'}}} \\
&= \bigoplus_{\boldsymbol{b} \in \mathcal{B} \times \mathcal{B}'} \bigodot_{i=1}^{m} T^{(i)}_{(\boldsymbol{f},\boldsymbol{b}):\boldsymbol{s_i}} \odot \bigodot_{i=m+1}^{m+n} T^{(i)}_{(\boldsymbol{f},\boldsymbol{b}):\boldsymbol{\hat{s}_i}} \\
\iff & V = (\boldsymbol{s_1}, \ldots, \boldsymbol{s_m}, \boldsymbol{\hat{s}_{m+1}}, \ldots, \boldsymbol{\hat{s}_{m+n}} \rightarrow \boldsymbol{s_v}, T^{(1)}, \ldots, T^{(m+n)})
\end{aligned}$$

where the third equality holds because we only iterate over a sub-space of the indices that we defined for the computation of U, and because the first and second einsum expression share no symbols. The rest of the steps are the same as in Theorem 1. □

With this theorem, we can prove a property of the trace in a relatively simple manner, namely that for $A \in \mathbb{R}^{m \times n}, B \in \mathbb{R}^{n \times m}$, it holds that

$$\operatorname{trace}(A \cdot B) = \operatorname{trace}(B \cdot A).$$

Proof.

$$\begin{aligned}
\operatorname{trace}(A \cdot B) &= (ii \rightarrow, (ik, kj \rightarrow ij, A, B)) \\
&= (ik, ki \rightarrow, A, B) \\
&= (ki, ik \rightarrow, B, A) \\
&= \operatorname{trace}(B \cdot A)
\end{aligned}$$

where the second equality holds because of Theorem 2, and the third equality holds because of the commutativity of multiplication in the standard semiring. □

This is already a useful tool for compressing nested expressions, but there are still some naturally occuring expressions we cannot compress with this, e.g.:

$$A \cdot \operatorname{diag}(v) = (ik, kj \rightarrow ij, A, (i \rightarrow ii, v))$$

for $A \in \mathbb{R}^{m \times n}, v \in \mathbb{R}^n$. This is because the symbol duplication ii is broken by the index string kj, and therefore we access more entries than the ones we computed.

3.3 Removing Duplications

The following example is an expression, which we cannot compress with the previous theorems:

$$(ij, kl, mn, ijklmn \rightarrow ijk, A, B, C, (abc \rightarrow aabbcc, D))$$

for $A \in \mathbb{R}^{m\times b\times b\times c}, B \in \mathbb{R}^{b\times c}$. In the following theorem, we will explore how to compress expressions such as this one. Again, we use disjoined sets of symbols for the inner and outer expression to help us in the formulation and the proof.

Theorem 3: For $i \in [m+n]$, let $T^{(i)}$ be an n_i-th order tensor with index strings $\boldsymbol{s_i} \in S^{n_i}$. Let $\boldsymbol{s_u}$ be an index string for the n_u-th order tensor U, which is defined as follows:

$$U := (\boldsymbol{s_{m+1}}, \dots, \boldsymbol{s_{m+n}} \rightarrow \boldsymbol{s_u}, T^{(m+1)}, \dots, T^{(m+n)})$$

Also let $\boldsymbol{\hat{s}_u}$ be alternative index strings for U with $s_{uj} \neq s_{uj'} \implies \hat{s}_{uj} \neq \hat{s}_{uj'}$ for all $j, j' \in [n_u]$, which means that $\boldsymbol{\hat{s}_u}$ can only remove symbol duplications, and cannot introduce any. Note that this is the converse of the constraint in Theorem 2.

In our example, $\boldsymbol{s_u} = oopp$ and $\boldsymbol{\hat{s}_u} = jklm$. This removes the symbol duplication of the first and second index, as well as the symbol duplication of the third and fourth index.

Let s_v be an index string and

$$V := (\boldsymbol{s_1}, \dots, \boldsymbol{s_m}, \boldsymbol{\hat{s}_u} \rightarrow \boldsymbol{s_v}, T^{(1)}, \dots, T^{(m)}, U)$$

where the first and second Einsum expression share no symbols. Then these nested Einsum expressions can also be compressed into a single Einsum expression.

As in Theorem 2, we need to apply a symbol map before substituting $\boldsymbol{\hat{s}_u}$. Interestingly, the symbol map is not applied to the index strings in the computation of U $(\boldsymbol{s_{m+1}}, \dots, \boldsymbol{s_{m+n}})$, but to the index strings in the computation of V $(\boldsymbol{s_1}, \dots, \boldsymbol{s_m})$. Similarly, it does not map $\boldsymbol{s_u}$ to $\boldsymbol{\hat{s}_u}$, but $\boldsymbol{\hat{s}_u}$ to $\boldsymbol{s_u}$.

Let $\nu : S \rightarrow S$ such that

$$\nu(s) := \begin{cases} s_{uj} & \text{if } \exists j \in [n_u] : \hat{s}_{uj} = s \\ s & \text{else} \end{cases},$$

which can be extended to map entire index strings as in Theorem 2. In our example, these are the important mappings:

$$\begin{aligned} j &\rightarrow o \\ k &\rightarrow o \\ l &\rightarrow p \\ m &\rightarrow p \end{aligned}$$

This means that j and k will be iterated over at the same time, and l and m will be iterated over at the same time.

Let $\boldsymbol{\hat{s}_i} := \nu(\boldsymbol{s_i})$ for $i \in [m]$, $\boldsymbol{\hat{s}_v} := \nu(\boldsymbol{s_v})$, then the compressed Einsum expression is the following:

$$V = (\boldsymbol{\hat{s}_1}, \dots, \boldsymbol{\hat{s}_m}, \boldsymbol{s_{m+1}}, \dots, \boldsymbol{s_{m+n}} \rightarrow \boldsymbol{\hat{s}_v}, T^{(1)}, \dots, T^{(m+n)})$$

which helps us to compress the example:

$$(ijkl, jklm \to ij, A, (op \to oopp, B)) = (ioop, op \to io, A, B)$$

Note how the index string for the output $\boldsymbol{s_v}$ was changed into $\boldsymbol{\hat{s}_v}$. This will become apparent in the proof.

Proof. The key idea behind this proof, is that the entries of U, which were not defined in the computation, are set to the additive neutral element $\mathbb{0}$. This is useful, because in a semiring over some set M, the additive neutral element *annihilates* M. This means, that for any $a \in M$, $a \cdot \mathbb{0} = \mathbb{0} \cdot a = \mathbb{0}$. Therefore, for any multi-index where U is set to $\mathbb{0}$, V is also set to $\mathbb{0}$. This means, that in the computation of V, only the indices which respect the duplications in $\boldsymbol{s_u}$ are defined.

Let F, F', B, B' be the free and bound symbols of the second (outer) and first (inner) einsum expression respectively. W.l.o.g. they are all non-empty. From them we can derive $\mathcal{F}, \mathcal{F}', \mathcal{B}, \mathcal{B}'$ as in the definition. Then $U_{(\boldsymbol{f},\boldsymbol{b}):\boldsymbol{\hat{s}_u}}$ is only non-zero for multi-indices $(\boldsymbol{f}, \boldsymbol{b}) \in \mathcal{F} \times \mathcal{B}$ with $(\boldsymbol{f}, \boldsymbol{b}) : \hat{s}_{uj} = (\boldsymbol{f}, \boldsymbol{b}) : \hat{s}_{uj'}$, where $j, j' \in [n_u]$ are indices of $\boldsymbol{s_u}$ where the symbols are duplicated, i.e. $s_{uj} = s_{uj'}$. In our example, this means that $(op \to oopp, B)$ is only non-zero for $(j, k, l, m) \in [d_j] \times [d_k] \times [d_l] \times [d_m]$ with $j = k$ and $l = m$, because $s_{u1} = s_{u2} = o$ and $s_{u3} = s_{u4} = p$.

Therefore, when U is multiplied with the other tensors, the resulting entry

$$\bigodot_{i=1}^{m} T^{(i)}_{(\boldsymbol{f},\boldsymbol{b}):\boldsymbol{s_i}} \odot U_{(\boldsymbol{f},\boldsymbol{b}):\boldsymbol{\hat{s}_u}}$$

is only non-zero for multi-indices $(\boldsymbol{f}, \boldsymbol{b}) \in \mathcal{F} \times \mathcal{B}$ that respect the same conditions. In our example, this is equivalent to

$$A_{ij} B_{k,l} C_{mn} \odot U_{ijklmn} = \begin{cases} A_{ij} B_{kl} C_{mn} \odot U_{ijklmn} & \text{if } i = j, k = l, m = n \\ \mathbb{0} & \text{else} \end{cases}.$$

Now, this already looks like not all symbols are needed for this computation. But to see, in which way we can replace symbols, we need to consider the three ways in which duplications can be broken. Either a duplication is broken only by free symbols, only by bound symbols, or by a combination of both. In our example, we have all of these cases. The duplication aa is broken by i and j, which are both free symbols. The duplication cc is broken by m and n, which are both bound symbols. The duplication bb is broken by k and l, where k is a free symbol and l is a bound symbol. Every one of these cases leads to the same result, but in a slightly different way.

First let us consider the case where a duplication is broken only by free symbols. In this case, the free symbols that break the duplication can be replaced by a single symbol, because all entries of V, with a multi-index that does not respect the duplication, is $\mathbb{0}$. In our example, this is equivalent to replacing i and j by a single symbol a:

$$\forall i, j, k : V_{ijk} = \bigoplus_{l,m,n} A_{ij} B_{kl} C_{mn} \odot U_{ijklmn}$$
$$\forall i, j, k : V_{aak} = \bigoplus_{l,m,n} A_{aa} B_{kl} C_{mn} \odot U_{aaklmn}$$

For the next two cases, we need to use that $a \oplus \mathbb{0} = a$ for any $a \in M$. This means, that only those summands that respect the duplications will be summed over, because all summands which do not respect the summation are $\mathbb{0}$. This affects the remaining two cases in different ways. If a duplication is broken only by bound symbols, then we need a single symbol to sum over all the multi-indices that respect the duplication. In our example, this is equivalent to replacing m and n by a single symbol c:

$$\bigoplus_{l,m,n} A_{aa} B_{kl} C_{mn} \odot U_{aaklmn} = \bigoplus_{l,c} A_{aa} B_{kl} C_{cc} \odot U_{aaklcc}$$

Now, if a duplication is broken by free symbols and by bound symbols, then the free symbols can again be replaced by a single symbol, as in the first case. In our example, this is useless, because there is already only one symbol k where this can be applied. Let us do it anyway, because it might clear up what needs to be done in this step.

$$\forall i, j, k : V_{aak} = \bigoplus_{l,c} A_{aa} B_{kl} C_{cc} \odot U_{aaklcc}$$
$$\forall i, j, k : V_{aab} = \bigoplus_{l,c} A_{aa} B_{bl} C_{cc} \odot U_{aablcc}$$

Then, the values held by the breaking bound symbols are already defined by the value held by the now only breaking free symbol. Therefore, the summation over the values of this symbol is useless, because there is only one combination of indices, which respects the duplication. Therefore, the breaking bound symbols need to hold the same exact value as the breaking free symbol, and these symbols can be replaced by the same symbol, that was used to replace the free symbols. Because of this, the occurence of the symbol also needs to be removed from the sum. In our example, this is equivalent to replacing l by b and removing it from the sum:

$$\begin{aligned}\forall i, j, k : V_{aab} &= \bigoplus_{l,c} A_{aa} B_{bl} C_{cc} \odot U_{aablcc} \\ &= \bigoplus_{c} A_{aa} B_{bb} C_{cc} \odot U_{aabbcc}\end{aligned}$$

$$\bigoplus_{l,c} A_{aa} B_{bl} C_{cc} \odot U_{aablcc} = \bigoplus_{l,c} A_{aa} B_{bl} C_{cc} \odot U_{aablcc}$$

Now, because $a \oplus \mathbb{0} = a$ for any $a \in M$, if a duplication is broken by symbols

Therefore, and because $a \oplus \mathbb{0} = a$ for any $a \in M$, the summation is reduced to only those summands, which have multi-indices that respect the duplications. Therefore we can iterate over the summands in such a way, that only those multi-indices are considered, that respect the duplications. For this, we define a different set of bound symbols $\hat{B} =$

$\left(\bigcup_{i\in[m]}\sigma(\hat{\boldsymbol{s}}_{\boldsymbol{i}})\cup\sigma(\boldsymbol{s_u})\right)\setminus\sigma(\hat{\boldsymbol{s}}_{\boldsymbol{v}})$. From these bound symbols, we can derive its corresponding multi-index space $\hat{\mathcal{B}}=\prod_{s\in\hat{B}}[d_s]$. In our example, the new set of bound symbols is $\hat{B}=\{p\}$, and the original set of bound symbols is $B=\{k,l,m\}$.

Now, in order to use these multi-indices in a well-defined manner together with the multi-indices of the free symbols, we can define an incomplete symbol map $\mu : S\to S$. This symbol map is incomplete because it does not contain all mappings, that are contained in the complete symbol map ν. It is restricted on only those symbols, that will be part of the bound symbols after application of the complete map. The symbol map also has to produce index strings that are able to project the multi-indices $(\boldsymbol{f},\hat{\boldsymbol{b}})\in\mathcal{F}\times\hat{\mathcal{B}}$ in a well-defined manner, which is possible in the first place because $F\cap\hat{B}=\emptyset$. Therefore we define

$$\mu(s):=\begin{cases} s_{uj} & \text{if } \exists j:\hat{s}_{uj}=s \text{ and } s_{uj}\in\hat{B} \\ s & \text{else}\end{cases}.$$

To indicate the incompletely mapped index strings, we denote $\mu(\boldsymbol{s_i})$ as $\check{\boldsymbol{s}}_{\boldsymbol{i}}$. In our example, these are the important mappings:

$$\begin{aligned} l&\to p\\ m&\to p\end{aligned}$$

Then

$$\bigoplus_{\boldsymbol{b}\in\mathcal{B}}\bigodot_{i=1}^{m}T^{(i)}_{(\boldsymbol{f},\boldsymbol{b}):\boldsymbol{s_i}}\odot U_{(\boldsymbol{f},\boldsymbol{b}):\hat{\boldsymbol{s}}_{\boldsymbol{u}}}=\begin{cases}\bigoplus_{\hat{\boldsymbol{b}}\in\hat{\mathcal{B}}}\bigodot_{i=1}^{m}T^{(i)}_{(\boldsymbol{f},\hat{\boldsymbol{b}}):\check{\boldsymbol{s}}_{\boldsymbol{i}}}\odot U_{(\boldsymbol{f},\hat{\boldsymbol{b}}):\boldsymbol{s_u}} & \text{if } \boldsymbol{f} \text{ respects the duplications}\\ \mathbb{0} & \text{else}\end{cases}$$

for all $\boldsymbol{f}\in\mathcal{F}$. When applied to our example, this is equivalent to:

$$\bigoplus_{l,m,n}A_{ij}B_{kl}C_{mn}\odot U_{ijklmn}=\begin{cases}\bigoplus_{l,m,n}A_{aa}B_{b,l}C_{mn}\odot U_{aablmn} & \text{if } a=i=j, b=k\\ \mathbb{0} & \text{else}\end{cases}$$

for all (i,j,k).

When applied to our example, this is equivalent to:

$$\bigoplus_{l,m,n}A_{ij}B_{kl}C_{mn}\odot U_{ijklmn}=\begin{cases}\bigoplus_{b,c}A_{ij}B_{bb}C_{cc}\odot U_{ijbbcc} & \text{if } i=j\\ \mathbb{0} & \text{else}\end{cases}$$

for all (i,j,k).

Now, because all $\boldsymbol{f}\in\mathcal{F}$ that do not respect the duplications are $\mathbb{0}$, we can also iterate over the free variables in such a way, that only those multi-indices are considered, that respect the duplications. For this, we define a new multi-index space $\hat{\mathcal{F}}=\sigma(\hat{\boldsymbol{s}}_{\boldsymbol{v}})$, and project these multi-indices with index strings $\hat{\boldsymbol{s}}_{\boldsymbol{i}}$ and $\hat{\boldsymbol{s}}_{\boldsymbol{v}}$ that respect the duplications. Then

$$\begin{aligned}
& V = (\boldsymbol{s_1}, \dots, \boldsymbol{s_m}, \hat{\boldsymbol{s}}_{\boldsymbol{u}} \rightarrow \boldsymbol{s_v}, T^{(1)}, \dots, T^{(m)}, U) \\
\iff & \forall \boldsymbol{f} \in \mathcal{F} : V_{\boldsymbol{f}:\boldsymbol{s_v}} = \bigoplus_{\boldsymbol{b} \in \mathcal{B}} \bigodot_{i=1}^{m} T^{(i)}_{(\boldsymbol{f},\boldsymbol{b}):\boldsymbol{s_i}} \odot U_{(\boldsymbol{f},\boldsymbol{b}):\hat{\boldsymbol{s}}_{\boldsymbol{u}}} \\
\iff & \forall \hat{\boldsymbol{f}} \in \hat{\mathcal{F}} : V_{\hat{\boldsymbol{f}}:\hat{\boldsymbol{s}}_{\boldsymbol{v}}} = \bigoplus_{\hat{\boldsymbol{b}} \in \hat{\mathcal{B}}} \bigodot_{i=1}^{m} T^{(i)}_{(\hat{\boldsymbol{f}},\hat{\boldsymbol{b}}):\hat{\boldsymbol{s}}_{\boldsymbol{i}}} \odot U_{(\hat{\boldsymbol{f}},\hat{\boldsymbol{b}}):\boldsymbol{s_u}} \\
& = \bigoplus_{\hat{\boldsymbol{b}} \in \hat{\mathcal{B}}} \bigodot_{i=1}^{m} T^{(i)}_{(\hat{\boldsymbol{f}},\hat{\boldsymbol{b}}):\hat{\boldsymbol{s}}_{\boldsymbol{i}}} \odot \left[\bigoplus_{\boldsymbol{b'} \in \mathcal{B}'} \bigodot_{i'=m+1}^{m+n} T^{(i')}_{(\hat{\boldsymbol{f}},\hat{\boldsymbol{b}},\boldsymbol{b'}):\boldsymbol{s_{i'}}} \right] \\
& = \bigoplus_{\hat{\boldsymbol{b}} \in \hat{\mathcal{B}} \times \mathcal{B}'} \bigodot_{i=1}^{m} T^{(i)}_{(\hat{\boldsymbol{f}},\hat{\boldsymbol{b}}):\hat{\boldsymbol{s}}_{\boldsymbol{i}}} \odot T^{(i)}_{(\hat{\boldsymbol{f}},\hat{\boldsymbol{b}}):\boldsymbol{s_i}} \\
\iff & V = (\hat{\boldsymbol{s}}_{\boldsymbol{1}}, \dots, \hat{\boldsymbol{s}}_{\boldsymbol{m}}, \boldsymbol{s_{m+1}}, \dots, \boldsymbol{s_{m+n}} \rightarrow \hat{\boldsymbol{s}}_{\boldsymbol{v}}, T^{(1)}, \dots, T^{(m+n)})
\end{aligned}$$

TODO: expand? □

With these theorems, we can write some more complex expressions from linear algebra as a single Einsum expression respectively:

- squared norm of matrix-vector multiplication: Let $A \in \mathbb{R}^{m \times n}, v \in \mathbb{R}^n$. Then

$$\begin{aligned}
|A \cdot v|_2^2 &= (i, i \rightarrow, (ij, j \rightarrow i, A, v), (ij, j \rightarrow i, A, v)) \\
&= (ij, j, ij, j \rightarrow, A, v, A, v)
\end{aligned}$$

- trace of matrix-matrix multiplication: Let $A \in \mathbb{R}^{m \times n}, B \in \mathbb{R}^{n \times m}$. Then

$$\begin{aligned}
\operatorname{trace}(A \cdot B) &= (ii \rightarrow, (ik, kj \rightarrow ij, A, B)) \\
&= (ik, ki \rightarrow, A, B)
\end{aligned}$$

- matrix multiplication with a diagonal matrix: Let $A \in \mathbb{R}^{m \times n}, v \in \mathbb{R}^n$. Then

$$\begin{aligned}
A \cdot \operatorname{diag}(v) &= (ik, kj \rightarrow ij, A, (i \rightarrow ii, v)) \\
&= (ij, j \rightarrow ij, A, v)
\end{aligned}$$

3.4 General Nested Expressions

In the final generalisation of compressing nested Einsum expressions, all duplication breaking is allowed. This has no application for linear algebra, as all the previous theorems had, because this first comes into play with third order tensors. This is because, with two or less axes, there is no possibility of simultaniously breaking a duplication and introducing a new one. Nevertheless, it serves as a useful tool for compressing all kinds of nested Einsum expressions.

In the following theorem, we explore a way of compressing the expression

$$(a, b, c, d, e, abbde \to bc, v^{(1)}, v^{(2)}, v^{(3)}, v^{(4)}, v^{(5)}, (i, j, k, l \to iijkkl, v^{(6)}, v^{(7)}, v^{(8)}, v^{(9)}))$$

for $v^{(i)} \in \mathbb{R}^{d_{vi}}$ with $i \in [9]$, $\boldsymbol{d_v} = (f, f, g, g, h, f, f, g, h) \in \mathbb{N}^9$. Again, we use disjoined sets of symbols for the inner and outer expression to help us in the formulation and the proof.

Theorem 4: For $i \in [m + n]$, let $T^{(i)}$ be an n_i-th order tensor with index strings $\boldsymbol{s_i} \in S^{n_i}$. Let $\boldsymbol{s_u}$ be an index string for the n_u-th order tensor U, which is defined as follows:

$$U := (\boldsymbol{s_{m+1}}, \dots, \boldsymbol{s_{m+n}} \to \boldsymbol{s_u}, T^{(m+1)}, \dots, T^{(m+n)})$$

Also let $\boldsymbol{\hat{s}_u}$ be alternative index strings for U.

Let s_v be an index string and

$$V := (\boldsymbol{s_1}, \dots, \boldsymbol{s_m}, \boldsymbol{\hat{s}_u} \to \boldsymbol{s_v}, T^{(1)}, \dots, T^{(m)}, U)$$

where the first and second Einsum expression share no symbols. Then these nested Einsum expressions can also be compressed into a single Einsum expression.

Once again, a map $\nu : S \to S$ has to be applied to the index strings before substituting index strings. The definition of the map this time is somewhat more complex. As in the previous two theorems, this map holds information about which symbols are essentially used together as one index.

For the definition of the map ν, we first construct an undirected graph $G = (V, E)$ that we call *symbol graph*. In the symbol graph, the nodes consist of all symbols from both expressions. The edges are $E = \{\{s_{uj}, \hat{s}_{uj}\} \mid j \in [n_u]\}$, which connects all symbols from $\boldsymbol{s_u}$ and $\boldsymbol{\hat{s}_u}$ that share an index. The symbol graph for our example is displayed in Figure 3.1.

In the symbol graph, if two symbols are connected, then they need to be iterated over at the same time in the compressed expression, because they are essentially the same index. Therefore, it makes sense assigning a symbol $s_C \in S \setminus V$ to each of the graphs components C. Then we can define ν as follows:

$$\nu(s) := \begin{cases} s_C & \text{if } s \in C \\ s & \text{else} \end{cases}$$

In our example, the components are $\{a, b, i, j\}$, $\{c, d, k\}$, and $\{e, l\}$. Therefore we could use

$$\nu(s) := \begin{cases} x & \text{if } s \in \{a, b, i, j\} \\ y & \text{if } s \in \{c, d, k\} \\ z & \text{if } s \in \{e, l\} \\ s & \text{else} \end{cases},$$

which can be extended to map entire index strings as in Theorem 2.

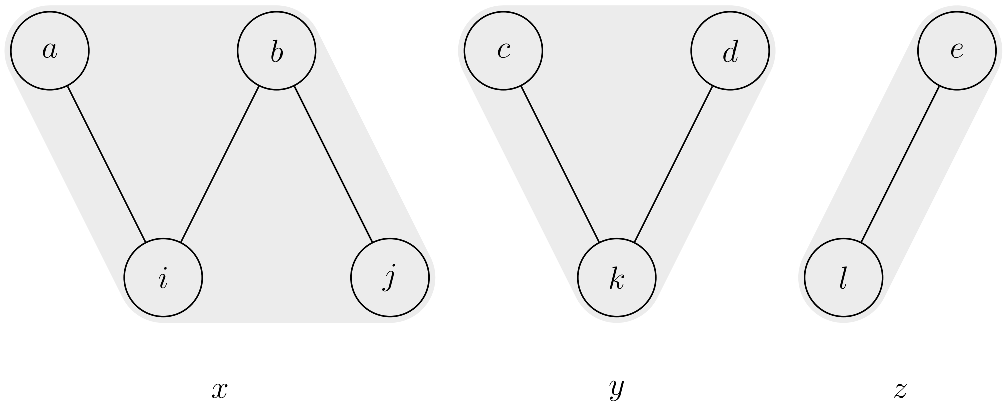


Figure 3.1: Symbol graph for the example

In this general form, this map is applied to all index strings from both expressions before the substitution. Let $\hat{\boldsymbol{s}}_{\boldsymbol{i}} := \nu(\boldsymbol{s}_{\boldsymbol{i}})$ for $i \in [m+n]$, $\hat{\boldsymbol{s}}_{\boldsymbol{v}} := \nu(\boldsymbol{s}_{\boldsymbol{v}})$, then the compressed Einsum expression is the following:

$$V = (\hat{\boldsymbol{s}}_{\boldsymbol{1}}, \ldots, \hat{\boldsymbol{s}}_{\boldsymbol{m+n}} \to \hat{\boldsymbol{s}}_{\boldsymbol{v}}, T^{(1)}, \ldots, T^{(m+n)})$$

which helps us to compress the example:

$$\begin{aligned}&(a, b, c, d, e, abbde \to bc, v^{(1)}, v^{(2)}, v^{(3)}, v^{(4)}, v^{(5)}, (i, j, k, l \to iijkkl, v^{(6)}, v^{(7)}, v^{(8)}, v^{(9)}))\\ &\quad = (x, x, y, y, z, x, x, y, z \to xy, v^{(1)}, v^{(2)}, v^{(3)}, v^{(4)}, v^{(5)}, v^{(6)}, v^{(7)}, v^{(8)}, v^{(9)})\end{aligned}$$

Proof. Wird spannend. $\square$

4 Methods

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

4.1 Some Section

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor.

Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

4.2 Some Other Section

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

5 Results

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a,

turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

5.1 Some Section

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a,

turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

5.2 Some Other Section

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a,

turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

6 Discussion

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

6.1 Some Section

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor.

Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

6.2 Some Other Section

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

7 Conclusion

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

List of Figures

List of Tables

Proofs

1 Exponential Blow-Up

Given: fully connected neural network with two layers ($n \to m \to l$ neurons), ReLU Activations, maps inputs $x \in \mathbb{R}^n$ to outputs $y \in R^l$, with parameters $A^{(0)} \in \mathbb{R}^{m\times n}, A^{(1)} \in \mathbb{R}^{l\times m}, b^{(0)} \in \mathbb{R}^m, b^{(1)} \in \mathbb{R}^l$. Then the computation of the neural network is:

$$y = \max(A^{(1)} \max(A^{(0)}x + b^{(0)}, 0) + b^{(1)}, 0)$$

To reasonably work with matrix multiplication in the tropcial semiring, we can only view matrices with positive integer entries. Making the entries integers does not impact the strength of the neural network, because [...] (quote the paper).

Now to only use positive valued matrices, we can rewrite the expression of computing the next layer from a previous layer:

$$\begin{aligned}\max(Ax + b, 0) &= \max(A_+x - A_-x + b, A_-x - A_-x)\\ &= \max(A_+x + b, A_-x) - A_-x\end{aligned}$$

where $A_+ = \max(A, 0), A_- = \max(-A, 0)$ and therefore $A = A_+ - A_-$.

This turns the network output into a tropical rational function (quote):

$$\begin{aligned}y = \max(&\overbrace{A_+^{(1)} \max(A_+^{(0)}x + b^{(0)}, A_-^{(0)}x)}^{z} + A_-^{(1)}A_+^{(0)}x + b^{(1)},\\ &A_-^{(1)} \max(A_+^{(0)}x + b^{(0)}, A_-^{(0)}x) + A_+^{(1)}A_+^{(0)}x)\\ &- \left[A_-^{(1)} \max(A_+^{(0)}x + b^{(0)}, A_-^{(0)}x) + A_+^{(1)}A_+^{(0)}x\right]\end{aligned}$$

We focus on the subexpression z, which makes the calculation a bit simpler, but keeps the point.

Now if we want to avoid switching semirings, we need to apply the distributive law a bunch of times.

$$z = A_+^{(1)} \max(A_+^{(0)}x + b^{(0)}, A_-^{(0)}x)$$

$$\begin{aligned}
z_i &= \bigodot_{j=1}^{m} \left(b_j^{(0)} \odot \bigodot_{k=1}^{n} x_k^{\odot A_{jk+}^{(0)}} \oplus \bigodot_{k=1}^{n} x_k^{\odot A_{jk-}^{(0)}} \right)^{\odot A_{ij+}^{(1)}} \\
&= \bigodot_{j=1}^{m} \left(\left(b_j^{(0)}\right)^{\odot A_{ij+}^{(1)}} \odot \bigodot_{k=1}^{n} x_k^{\odot\left(A_{ij+}^{(1)} + A_{jk+}^{(0)}\right)} \oplus \bigodot_{k=1}^{n} x_k^{\odot\left(A_{ij+}^{(1)} + A_{jk-}^{(0)}\right)} \right) \\
&= \bigoplus_{J \in 2^{[m]}} \bigodot_{j \in J} \left[\left(b_j^{(0)}\right)^{\odot A_{ij+}^{(1)}} \odot \bigodot_{k=1}^{n} x_k^{\odot\left(A_{ij+}^{(1)} + A_{jk+}^{(0)}\right)} \right] \odot \bigodot_{j \in [n] \setminus J} \left[\bigodot_{k=1}^{n} x_k^{\odot\left(A_{ij+}^{(1)} + A_{jk-}^{(0)}\right)} \right]
\end{aligned}$$

Where the second equality is just the first equality written with the operations of the tropical semiring, the third equality follows from the distributive law with standard operations, and the last equality follows from the distributive law in the tropical semiring.

This expression maximizes over a number of subexpressions that grows exponentially in the width of the inner layer. Which subexpressions can be removed before the evaluation remains an open question. Note that it depends on the non-linearities of the neural network, which might make it hard to find a general answer to this question.

2 Expressing Stuff as Einsum

Fully connected Feed-Forward Neural Net with ReLU activations (1 layer) $\nu : \mathbb{R}^n \to \mathbb{R}^m$ with weights $A \in \mathbb{R}^{m \times n}$ and biases $b \in \mathbb{R}^m$, input $x \in \mathbb{R}^n$

- $(\text{einsum\_expression})_R$ indicates that the einsum expression uses the semiring R
- $R_{(+,\cdot)}$ indicates the standard semiring
- $R_{(\max,+)}$ indicates the tropical semiring
- $R_{(\min,\max)}$ indicates the minimax semiring

$$\begin{aligned}
\nu(x) &= \max(Ax + b, 0) \\
&= (i, i \to i, 0, (i, i \to i, b, (ij, j \to i, A, x)_{R_{(+,\cdot)}})_{R_{(\max,+)}})_{R_{(\min,\max)}}
\end{aligned}$$

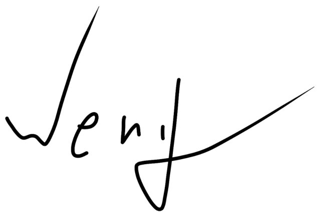
Attention with einsum:

$$\begin{aligned}
(QK^\top)_{ij} &= \sum_k Q_{ik}K_{jk} \\
QK^\top &= (ik, jk \to ij, Q, K) \\
\text{softmax}(X)_{ij} &= \frac{\exp(X_{ij})}{\sum\limits_{j'} \exp(X_{ij'})} \\
&= \exp(X_{ij}) \cdot \sum_{j'} \exp(-X_{ij'}) \\
\text{softmax}(X) &= (ij, i \to ij, \exp(X), (ij \to i, \exp(-X))) \\
(XV)_{ij} &= \sum_k X_{ik}V_{kj} \\
XV &= (ik, kj \to ij, X, V) \\
\text{Attention}(Q, K, V) &= \text{softmax}\left(\frac{QK^\top}{\sqrt{d_K}}\right)V \\
&= (ik, kj \to ij, (ij, i \to ij, \exp(\frac{1}{\sqrt{d_k}} \cdot (ik, jk \to ij, Q, K)), \\
&\quad (ij \to i, \exp(-\frac{1}{\sqrt{d_k}} \cdot (ik, jk \to ij, Q, K)))), V) \\
&= (ik, i, kj \to ij, \exp(\frac{1}{\sqrt{d_k}} \cdot (ik, jk \to ij, Q, K)), \\
&\quad (ij \to i, \exp(-\frac{1}{\sqrt{d_k}} \cdot (ik, jk \to ij, Q, K))), V) \\
&= (ik, ij, kj \to ij, \exp(\frac{1}{\sqrt{d_k}} \cdot (ik, jk \to ij, Q, K)), \\
&\quad \exp(-\frac{1}{\sqrt{d_k}} \cdot (ik, jk \to ij, Q, K)), V)
\end{aligned}$$

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

Selbständigkeitserklärung

Ich erkläre, dass ich die vorliegende Arbeit selbstständig und nur unter Verwendung der angegebenen Quellen und Hilfsmittel angefertigt habe. Seitens des Verfassers bestehen keine Einwände die vorliegende Bachelorarbeit für die öffentliche Benutzung im Universitätsarchiv zur Verfügung zu stellen.



Jena, 16.06.2023