



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, consectetur 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, consectetur 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

Abstract	iii
1 Introduction	1
1.1 Einsum	1
1.2 Some Other Section	7
2 Methods	9
2.1 Some Section	9
2.2 Some Other Section	10
3 Results	11
3.1 Some Section	12
3.2 Some Other Section	13
4 Discussion	15
4.1 Some Section	15
4.2 Some Other Section	16
5 Conclusion	17
Proofs	23
1 General Einsum Stuff	24
2 Expressing Stuff as Einsum	24
Selbständigkeitserklärung	27

1 Introduction

1.1 Einsum

Given two third-order tensor $A \in \mathbb{R}^{3 \times 4 \times 5}$, $B \in \mathbb{R}^{3 \times 3 \times 5}$ and a vector $v \in \mathbb{R}^4$. Consider the following computation for a matrix $C \in \mathbb{R}^{3 \times 3}$:

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

This can be written in a much shorter form with einsum:

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

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

Definition 1. Einsum expressions generally evaluate to some computation over tensors, so let $T^{(1)}, \dots, T^{(n)}$ be our input tensors, where $T^{(i)}$ is a 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 $\mathbf{s}_i \in S^{n_i}$, composed of symbols $s_{ij} \in S$ for $j \in [n_i]$. We refer to the index string that is right of the arrow (\rightarrow) as \mathbf{s}_t .

In our example this could be $S = \{i, j, k\}$ with respective index strings $\mathbf{s}_1 = ijk$ for $T^{(1)} = A$, $\mathbf{s}_2 = iik$ for $T^{(2)} = B$, $\mathbf{s}_3 = j$ for $T^{(3)} = v$, and $\mathbf{s}_t = ij$. The individual symbols are $s_{11} = i$, $s_{12} = j$, $s_{13} = k$, $s_{21} = i$, $s_{22} = i$, $s_{23} = k$, $s_{31} = j$, $s_{t1} = i$, $s_{t2} = j$.

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.

The next step in the definition is to speak about the axis sizes. If we want to iterate over shared indices, it would be nice if 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 4 (four?). Let us express this formally.

Let d_{ij} 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 assigned the same symbol. E.g. a square matrix could have index strings $\mathbf{s} = (i, i)$ or $\mathbf{s} = (i, j)$.

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 $\sigma(\mathbf{s})$ denote the set with all symbols used in an index string \mathbf{s} . E.g. $\sigma(\mathbf{s}_2) = \sigma(iik) = \{ik\}$.

All symbols that are present in ...

It follows from this, that d_s has to be defined for all symbols $s \in \mathbf{s}_t$. This means we have to add the constraint $\sigma(\mathbf{s}_t) \subseteq \bigcup_{i \in [n]} \sigma(\mathbf{s}_i)$.

Let S be a collection of symbols that correspond to non-empty sets of tensor axes. For $i \in [n]$, let $T^{(i)}$ be a n_i -th order tensor with axis symbols $s_{ij} \in S$ for their unique axes $a_{ij} \in \mathbb{N}$. The actual value of a_{ij} does not matter. These variables are just used as a unique identifier for the axes so that the definition is more rigorous. From the axis symbols we build the index string $\mathbf{s}_i \in S^{n_i}$ which consists of all the axis symbols in order.

The size of the tensor $T^{(i)}$ on the axis a_{ij} is denoted as d_{ij} for $j \in [n_i]$. For this definition it must hold that $s_{ij} = s_{i'j'} \implies d_{ij} = d_{i'j'}$ for all $i, i' \in [n], j \in [n_i], j' \in [n_{i'}]$. This means that the set of axes that a symbol $s \in S$ corresponds to, must all have the same size.

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 assigned the same symbol. E.g. with $S = \{i, j\}$, a square matrix could have index strings $\mathbf{s} = (i, i)$ or $\mathbf{s} = (i, j)$.

Let $\sigma(\mathbf{s})$ denote the set with all symbols used in the index string \mathbf{s} . E.g. $\sigma(iijk) = \{ijk\}$.

Let $T^{(0)}$ be a $d^{(0)}$ -th order tensor with index string $\mathbf{s}^{(0)} \in S^{d^{(0)}}$. This tensor will be the result of the computation. Its index strings decide which axes are kept and which axes will be summed over. The collection of symbols S is partitioned into bound symbols $B = \{s_j^{(0)} \mid j \in [d^{(0)}]\}$ and free symbols $F = S \setminus B$. We will keep all axes corresponding to symbols in B and sum over all axes corresponding to symbols in F .

Let $\mathcal{I} := \prod_{s \in B} [d_s]$ and $\mathcal{J} := \prod_{s \in F} [d_s]$. These will be the space of multi-indices which we iterate over. For $I \in \mathcal{I}, J \in \mathcal{J}$, the concatenation of the indices is denoted as IJ . The projection of the multi-index over all tensors IJ on the multi-index for tensor $T^{(i)}$ is denoted as $IJ : s^{(i)}$. $I : s^{(0)}$ is defined analogously.

Given a semiring $R = (M, \oplus, \odot)$. Then the einsum-notation denotes the following:
 Given $\mathbf{s}_t \in S^{n_t}$ and

$$T := (\mathbf{s}_1, \dots, \mathbf{s}_n \rightarrow \mathbf{s}_t, T^{(1)}, \dots, T^{(n)})_R$$

Then T is a n_t -th order tensor with

$$\forall I \in \mathcal{I} : T_{I:\mathbf{s}_t} = \bigoplus_{J \in \mathcal{J}} \bigodot_{i=1}^n T_{IJ:\mathbf{s}_i}^{(i)}$$

$$\begin{aligned} T^{(0)} &= (s^{(1)}, \dots, s^{(n)} \rightarrow s^{(0)}, T^{(1)}, \dots, T^{(n)})_R \\ : \iff \forall I \in \mathcal{I} : T_{I:s^{(0)}}^{(0)} &= \bigoplus_{J \in \mathcal{J}} \bigodot_{i=1}^n T_{IJ:s^{(i)}}^{(i)} \end{aligned}$$

if B, F are non-empty. However, if there are no free symbols, the summation is removed:

$$\begin{aligned} T^{(0)} &= (s^{(1)}, \dots, s^{(n)} \rightarrow s^{(0)}, T^{(1)}, \dots, T^{(n)})_R \\ : \iff \forall I \in \mathcal{I} : T_{I:s^{(0)}}^{(0)} &= \bigodot_{i=1}^n T_{I:s^{(i)}}^{(i)} \end{aligned}$$

If there are no bound symbols, then $s^{(0)}$ is empty and $T^{(0)}$ is a scalar:

$$\begin{aligned} T^{(0)} &= (s^{(1)}, \dots, s^{(n)} \rightarrow s^{(0)}, T^{(1)}, \dots, T^{(n)})_R \\ : \iff T^{(0)} &= \bigoplus_{J \in \mathcal{J}} \bigodot_{i=1}^n T_{J:s^{(i)}}^{(i)} \end{aligned}$$

Note how we would not need the projection $I : s^{(0)}$ if $s^{(0)}$ had no duplicate symbols, because ...

Note that duplicate entries in $s^{(i)}, i \in [n]$ will result in the expression iterating over the input tensors in a sort of *diagonal* way. Duplicate entries in $s^{(0)}$ will result in some entries of $T^{(0)}$ not being defined. These are set to the additive neutral element of the semiring.

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

1.1.1 Examples

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

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

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

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

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

$$\text{trace}(A) = (ii \rightarrow, A)$$

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

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

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

$$\text{diag}(v) = (i \rightarrow ii, v)$$

1.1.2 Nested Einsum Expressions

Theorem 1: For $i \in [m+n+1]$, let $T^{(i)}$ be a $d^{(i)}$ -th order tensor with index strings $s^{(i)} \in S^{d^{(i)}}$, $o := m+n+1$. Let

$$T^{(0)} := (s^{(1)}, \dots, s^{(m)}, s^{(x)} \rightarrow s^{(0)}, T^{(1)}, \dots, T^{(m)}, T^{(o)})$$

and

$$T^{(o)} = (s^{(m+1)}, \dots, s^{(m+n)} \rightarrow s^{(o)}, T^{(m+1)}, \dots, T^{(m+n)})$$

where the free symbols of the second einsum expression share no symbols with the first einsum expression.

Then

$$T^{(0)} = (s^{(1)}, \dots, s^{(m+n)} \rightarrow s^{(0)}, T^{(1)}, \dots, T^{(m+n)})$$

Proof. Let B, B', F, F' be the bound and free symbols of the first and second einsum expression respectively. W.l.o.g. they are all non-empty. From them we can derive

$\mathcal{I}, \mathcal{I}', \mathcal{J}, \mathcal{J}'$ as above. Then

$$\begin{aligned}
T^{(0)} &= (s^{(1)}, \dots, s^{(m)}, s^{(o)} \rightarrow s^{(0)}, T^{(1)}, \dots, T^{(m)}, T^{(o)}) \\
\iff \forall I \in \mathcal{I} : T_{I:s^{(0)}}^{(0)} &= \bigoplus_{J \in \mathcal{J}} \bigodot_{i=1}^m T_{IJ:s^{(i)}}^{(i)} \odot T_{IJ:s^{(o)}}^{(o)} \\
&= \bigoplus_{J \in \mathcal{J}} \bigodot_{i=1}^m T_{IJ:s^{(i)}}^{(i)} \odot \bigoplus_{J' \in \mathcal{J}'} \bigodot_{i'=m+1}^{m+n} T_{IJJ':s^{(i')}}^{(i')} \\
&= \bigoplus_{J \in \mathcal{J}} \bigoplus_{J' \in \mathcal{J}'} \bigodot_{i=1}^m T_{IJ:s^{(i)}}^{(i)} \odot \bigodot_{i=m+1}^{m+n} T_{IJJ':s^{(i)}}^{(i)} \\
&= \bigoplus_{J \in \mathcal{J} \times \mathcal{J}'} \bigodot_{i=1}^{m+n} T_{IJ:s^{(i)}}^{(i)} \\
\iff T^{(0)} &= (s^{(1)}, \dots, s^{(m+n)} \rightarrow s^{(0)}, T^{(1)}, \dots, T^{(m+n)})
\end{aligned}$$

where the third equality follows from

$$\forall I' \in \mathcal{I}' : T_{I':s^{(o)}}^{(o)} = \bigoplus_{J' \in \mathcal{J}'} \bigodot_{i'=m+1}^{m+n} T_{I'J':s^{(i')}}^{(i')},$$

$B' \subseteq B \cup F$, and $(B \cup F) \cap F' = \emptyset$. The last two facts are required so that $IJJ' : s^{(i')}$ is well-defined and projects on the same indices as $I'J' : s^{(i')}$. The fourth equality follows from the distributivity in a semiring. \square

1.1.3 A More General Result

Theorem 2: For $i \in [m+n+1]$, let $T^{(i)}$ be a $d^{(i)}$ -th order tensor with index strings $s^{(i)} \in S^{d^{(i)}}$, $o := m+n+1$. Also let $\hat{s}^{(o)}$ be alternative index strings for $T^{(o)}$ with $s_j^{(o)} = s_{j'}^{(o)} \implies \hat{s}_j^{(o)} = \hat{s}_{j'}^{(o)}$ for all $j, j' \in [d^{(o)}]$. Let

$$T^{(0)} := (s^{(1)}, \dots, s^{(m)}, \hat{s}^{(o)} \rightarrow s^{(0)}, T^{(1)}, \dots, T^{(m)}, T^{(o)})$$

and

$$T^{(o)} = (s^{(m+1)}, \dots, s^{(m+n)} \rightarrow s^{(o)}, T^{(m+1)}, \dots, T^{(m+n)})$$

where the free symbols of the second einsum expression share no symbols with the first einsum expression. Let $\nu : S \rightarrow S$ such that

$$\nu(s) = \begin{cases} \hat{s}_j^{(o)} & \text{if } \exists j \in [d^{(o)}] : s_j^{(o)} = s \\ s & \text{else} \end{cases}$$

which maps symbols in $s^{(o)}$ to the symbol at the same index in $\hat{s}^{(o)}$ and all other symbols to themselves. ν can be extended to map from axis symbol tuples by setting $\nu(s^{(i)}) \in S^{d^{(i)}}$, $\nu(s^{(i)})_j := \nu(s_j^{(i)})$.

Let $\hat{s}^{(i)} := \nu(s^{(i)})$ Then

$$T^{(0)} = (s^{(1)}, \dots, s^{(m)}, \hat{s}^{(m+1)}, \dots, \hat{s}^{(m+n)} \rightarrow s^{(0)}, T^{(1)}, \dots, T^{(m+n)})$$

Proof. Let ...

$$\begin{aligned} T^{(0)} &= (s^{(1)}, \dots, s^{(m)}, s^{(o)} \rightarrow \hat{s}^{(0)}, T^{(1)}, \dots, T^{(m)}, T^{(o)}) \\ \iff \forall I \in \mathcal{I} : T_{I:s^{(0)}}^{(0)} &= \bigoplus_{J \in \mathcal{J}} \bigodot_{i=1}^m T_{IJ:s^{(i)}}^{(i)} \odot T_{IJ:\hat{s}^{(o)}}^{(o)} \\ &= \bigoplus_{J \in \mathcal{J}} \bigodot_{i=1}^m T_{IJ:s^{(i)}}^{(i)} \odot \mu \left(\bigoplus_{J' \in \mathcal{J}'} \bigodot_{i'=m+1}^{m+n} T_{IJJ':s^{(i')}}^{(i')} \right) \\ &= \bigoplus_{J \in \mathcal{J}} \bigodot_{i=1}^m T_{IJ:s^{(i)}}^{(i)} \odot \bigoplus_{J' \in \mathcal{J}'} \bigodot_{i'=m+1}^{m+n} \mu \left(T_{IJJ':s^{(i')}}^{(i')} \right) \\ &= \bigoplus_{J \in \mathcal{J}} \bigodot_{i=1}^m T_{IJ:s^{(i)}}^{(i)} \odot \bigoplus_{J' \in \mathcal{J}'} \bigodot_{i'=m+1}^{m+n} T_{IJJ':\hat{s}^{(i')}}^{(i')} \\ &= \bigoplus_{J \in \mathcal{J} \times \mathcal{J}'} \bigodot_{i=1}^m T_{IJ:s^{(i)}}^{(i)} \odot \bigodot_{i=m+1}^{m+n} T_{IJ:\hat{s}^{(i)}}^{(i)} \\ \iff T^{(0)} &= (s^{(1)}, \dots, s^{(m)}, \hat{s}^{(m+1)}, \dots, \hat{s}^{(m+n)} \rightarrow s^{(0)}, T^{(1)}, \dots, T^{(m+n)}) \end{aligned}$$

□

1.1.4 More Examples

With these theorems, we can write some more complex expressions as einsum.

- 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

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

- The theorem for this still has to be shown ...: matrix multiplication with a diagonal matrix: Let $A \in \mathbb{R}^{m \times n}, v \in \mathbb{R}^n$. Then

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

1.2 Some Other Section

Lorem ipsum dolor sit amet, consectetur 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, consectetur 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, consectetur 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, consectetur 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 Methods

Lorem ipsum dolor sit amet, consectetur 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, consectetur 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, consectetur 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, consectetur 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.

2.1 Some Section

Lorem ipsum dolor sit amet, consectetur 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, consectetur 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, consectetur 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, consectetur 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.

2.2 Some Other Section

Lorem ipsum dolor sit amet, consectetur 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, consectetur 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, consectetur 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, consectetur 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.

3 Results

Lorem ipsum dolor sit amet, consectetur 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, consectetur 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, consectetur 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, consectetur 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, consectetur 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, consectetur 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, consectetur 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, consectetur 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.

3.1 Some Section

Lorem ipsum dolor sit amet, consectetur 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, consectetur 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, consectetur 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, consectetur 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, consectetur 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, consectetur 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, consectetur 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, consectetur 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.

3.2 Some Other Section

Lorem ipsum dolor sit amet, consectetur 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, consectetur 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, consectetur 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, consectetur 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, consectetur 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, consectetur 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, consectetur 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, consectetur 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 Discussion

Lorem ipsum dolor sit amet, consectetur 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, consectetur 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, consectetur 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, consectetur 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, consectetur 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, consectetur 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, consectetur 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, consectetur 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, consectetur 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, consectetur 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, consectetur 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, consectetur 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 Conclusion

Lorem ipsum dolor sit amet, consectetur 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, consectetur 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

Proof that expanding the matrix multiplication in a neural network results in an exponentially big term in terms of the width of the skipped layer of the neural network:

Given: fully connected neural network with two layers ($n \rightarrow m \rightarrow l$ neurons), ReLU Activations, maps inputs $x \in \mathbb{R}^n$ to outputs $y \in \mathbb{R}^l$, 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 tropical 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:

$$\max(Ax + b, 0) = \max(A_+x + b, A_-x) - A_-x$$

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.

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

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.

1 General Einsum Stuff

Nested einsum expressions in the same semiring can be combined into one smaller einsum expression.

(proof: basically distributive law.) Let $R = (M, \oplus, \odot)$ be a semiring, and $f : M \rightarrow M, g : M \rightarrow M$

2 Expressing Stuff as Einsum

Fully connected Feed-Forward Neural Net with ReLU activations (1 layer) $\nu : \mathbb{R}^n \rightarrow \mathbb{R}^m$ 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

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

Attention with einsum:

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

Lorem ipsum dolor sit amet, consectetur 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, consectetur 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.

Selbstständigkeitserklärung

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.

A handwritten signature in black ink, appearing to read 'Wenig', with a long horizontal stroke extending to the right.

Jena, 17.05.2023