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 Einsum

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

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

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

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

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

Definition 1. Einsum expressions generally evaluate to some computation over tensors, so let $T^{(1)}, \ldots, T^{(n)}$ 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 $\boldsymbol{s_i} \in S^{n_i}$, composed of symbols $s_{ij} \in S$ for $j \in [n_i]$. We refer to the index string that is right of the arrow ($\to$) as the output string $\boldsymbol{s_t}$.

In our example this could be $S = \{i, j, k\}$ with respective index strings $\boldsymbol{s_1} = ijk$ for $T^{(1)} = A$, $\boldsymbol{s_2} = iik$ for $T^{(2)} = B$, $\boldsymbol{s_3} = j$ for $T^{(3)} = v$, and $\boldsymbol{s_t} = ij$. 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.

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 $\boldsymbol{s} = (i, i)$ 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 $\boldsymbol{s}$. E.g. $\sigma(\boldsymbol{s_2}) = \sigma(iik) = \{ik\}$.

All symbols to the right of the arrow ($\rightarrow$) are used as an index for the result of the computation. These symbols are called *bound* symbols $B = \sigma(\boldsymbol{s_t})$. All other symbols used in the expression are called *free* symbols $F = \bigcup_{i\in[n]} \sigma(\boldsymbol{s_i}) \setminus \sigma(\boldsymbol{s_t})$. In einsum, we sum over all axes that belong to free symbols. It follows that the multi-index space we iterate over is $\mathcal{I} = \prod_{s\in B}[d_s]$ and the multi-index space we sum over is $\mathcal{J} = \prod_{s\in F}[d_s]$. In our example, the bound symbols are $B = \{ij\}$ and the free symbols are $F = \{k\}$. 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 $\mathcal{I}$, it follows that d_s has to be defined for all symbols $s \in B$. 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 (s) to index the tensor. To formally express this, we need a projection from a multi-index $IJ \in \mathcal{I} \times \mathcal{J}$ to another multi-index, which includes only the symbols used in $\boldsymbol{s}$, in the same order as present in $\boldsymbol{s}$. We denote this as $IJ : \boldsymbol{s}$. Notice how this still allows duplication of indices given in IJ. This is needed, as can be seen in our example for B_{iik}, where a multi-index, e.g. $(1, 4, 2) \in \mathcal{I} \times \mathcal{J}$, is projected on the index string iik, which results in the multi-index $(1, 4, 2) : iik = (1, 1, 2)$.

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 write down what computation a general einsum expression

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

results in. It means that T is a $|\boldsymbol{s_t}|$-th order tensor with

$$\forall I \in \mathcal{I} : T_{I:\boldsymbol{s_t}} = \bigoplus_{J\in\mathcal{J}} \bigodot_{i=1}^{n} T^{(i)}_{IJ:\boldsymbol{s_t}}$$

Because we also project the indices I on the output string $\boldsymbol{s_t}$, we allow to iterate over duplicate indices, e.g. $\text{diag}(v) = (j \to jj, v)$. This leaves some entries of the result undefined. We define these entries to be the additive neutral element $\mathbb{0}$ in the given semiring R. This may sound arbitrary at first, but will be useful for later theorems.

There are still some special case which need to be considered. If there are no free symbols in the expression, then the sum will be empty. But we still want the result of the computation of the product. Therefore, if $F = \emptyset$, then

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

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

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

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

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

results in the computation of a scalar T with

$$T = \bigoplus_{J \in \mathcal{J}} \bigodot_{i=1}^{n} T^{(i)}_{J:\boldsymbol{s_t}}$$

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

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

- 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

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

- 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

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

1.1.2 Nested Einsum Expressions

Theorem 1: For $i \in [m+n]$, let $T^{(i)}$ be a 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 free 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 B, B', F, F' be the bound and free symbols of the second and first 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

$$\begin{aligned}
& V = (\boldsymbol{s_1}, \ldots, \boldsymbol{s_m}, \boldsymbol{s_u} \rightarrow \boldsymbol{s_v}, T^{(1)}, \ldots, T^{(m)}, U) \\
\iff \forall I \in \mathcal{I} : V_{I:\boldsymbol{s_v}} &= \bigoplus_{J \in \mathcal{J}} \bigodot_{i=1}^{m} T^{(i)}_{IJ:\boldsymbol{s_i}} \odot U_{IJ:\boldsymbol{s_u}} \\
&= \bigoplus_{J \in \mathcal{J}} \bigodot_{i=1}^{m} T^{(i)}_{IJ:\boldsymbol{s_i}} \odot \bigoplus_{J' \in \mathcal{J}'} \bigodot_{i'=m+1}^{m+n} T^{(i')}_{IJJ':\boldsymbol{s_{i'}}} \\
&= \bigoplus_{J \in \mathcal{J}} \bigoplus_{J' \in \mathcal{J}'} \bigodot_{i=1}^{m} T^{(i)}_{IJ:\boldsymbol{s_i}} \odot \bigodot_{i=m+1}^{m+n} T^{(i)}_{IJJ':\boldsymbol{s_{i'}}} \\
&= \bigoplus_{J \in \mathcal{J} \times \mathcal{J}'} \bigodot_{i=1}^{m+n} T^{(i)}_{IJ:\boldsymbol{s_i}} \\
\iff & V = (\boldsymbol{s_1}, \ldots, \boldsymbol{s_{m+n}} \rightarrow \boldsymbol{s_v}, T^{(1)}, \ldots, T^{(m+n)})
\end{aligned}$$

where the third equality follows from

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

$B' \subseteq B \cup F$, and $(B \cup F) \cap F' = \emptyset$. The last two facts are required so that $IJJ' : \boldsymbol{s_{i'}}$ is well-defined and projects on the same indices as $I'J' : \boldsymbol{s_{i'}}$. 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

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

and

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

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

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

Proof. Let $\mathcal{I}' = \prod_{s \in \sigma(\hat{s}_u)} [d_s]$. Let $\mathcal{M} : \boldsymbol{s} := \{M : \boldsymbol{s} \mid M \in \mathcal{M}\}$ for an index string s and a multi-index space $\mathcal{M}$. Then $\mathcal{I}' : \boldsymbol{\hat{s}_u} \subseteq \mathcal{I} : \boldsymbol{s_u}$, because $d_{s_{uj}} = d_{\hat{s}_{uj}}$ per the definition of einsum, and because the amount of axes contributing to $\mathcal{I}'$ $(|\sigma(\boldsymbol{\hat{s}_u})|)$ has to be smaller or equal to the amount of axes contributing to $\mathcal{I}$ $(|\sigma(\boldsymbol{s_u})|)$. This last fact follows from the constraint $s_{uj} = s_{uj'} \implies \hat{s}_{uj} = \hat{s}_{uj'}$.

Then

$$\forall I \in \mathcal{I} : U_{I:\boldsymbol{s_u}} = \bigodot_{i=m+1}^{m+n} T^{(i)}_{IJ:\boldsymbol{s_i}}$$

and therefore

$$\forall I \in \mathcal{I}' : U_{I':\boldsymbol{\hat{s}_u}} = \bigodot_{i=m+1}^{m+n} T^{(i)}_{I'J:\boldsymbol{\hat{s}_i}}$$

because of the previous observation, and because the free symbols of the expression, which are used in J, are not changed by the symbol map ν.

Therefore

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

where ... □

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

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

- 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

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

1.2 Some Other Section

Lorem ipsum dolor sit amet, 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 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.

2.1 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.

2.2 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.

3 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.

3.1 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.

3.2 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.

4 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.

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 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

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 ($n \to m \to l$ 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:

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

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.

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

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.

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 $f : M \to M, g : M \to M$

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:

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

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, 22.05.2023