Einstein Summation Convention

January 4, 2024

1 Motivation

Einstein invented this convention as alternative way to represent tensors that go through mathematical operations. When tensor become high dimensional or have many components it is easer to use representation that will remove clutter.

2 Introduction

Given the following example, we will illustrate how not so complex tensor operations can be notionally expensive to perform in traditional notation.

Example 1 *Consider these two vectors* $\mathbf{a}$, $\mathbf{b} \in \mathbb{R}^n$. *Write out the dot product of these two vectors* $\mathbf{a} \cdot \mathbf{b}$.

$$\mathbf{a} \cdot \mathbf{b} = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} \cdot \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} = a_1 b_1 + \ldots + a_n b_n \equiv \sum_{i=1}^{n} a_i b_i$$

Question, what about $\mathbf{b} \times \{(\mathbf{a} \cdot \mathbf{b})\mathbf{b}\}$? ∎

One can see from the above example that performing tensor operations can get really messy very fast due to the notational representation of tensors. Hence the need to more simplistic notation is vital, which is what presented in Einstein convention.

3 Einstein Summation Convention

In many fields, it is more convenient to represent vectors, matrices and tensors with identical notation, where lower indices "*subscripts*" and upper indices "*superscripts*" are appended to the variable representing the quantities of interest[1]. Even though, in mathematics *subscripts* and *superscripts* are interpreted

[1] this notation was introduced in Ricci calculus by Gregorio Ricci-Curbastro

differently, Einstein in his summation convention do not distinguish between them[2].

$$x_i \equiv x^i$$
$$x_{ij} \equiv x^i_j \equiv x^{ij}$$

Einstein introduced his summation convention to provide a simplified notations for sums (which is a vital operation in tensor calculus). He defined his convention based on the following rules.

Rules

1. If an index appears twice in a single term[3]. Then it implies summation over the two variables that carry this index. This index is referred as "*dummy index*", where index that appears once is called "*free/live index*".

$$\ldots + a_i b_{jk} z_i + \ldots \equiv \ldots + \sum_{i=1}^{n} a_i b_{jk} z_i + \ldots$$

   Hence, in this example i is a dummy index where j, k are live indices.

2. Number of free/live indices determine the order of the tensor (no live indices means scalar, one means vector, two means second order tensor and so on).

3. An index should not appear more than twice in a single term.

$$\ldots + a_i b_{ik} z_i + \ldots \rightarrow \textbf{not valid}$$

   One of possible exceptions is $a_i(b_i + c_i)$ since we add $b_i,\ c_i$ first.

4 Examples

Example 2 *Let $n = 3$ "i.e. cardinality". Write out what is represented by $a_i b_i$*

We have a dummy index i hence this imply summation

$$a_i b_i = \sum_{i=1}^{3} a_i b_i = a_1 b_1 + a_2 b_2 + a_3 b_3 \equiv \mathbf{a} \cdot \mathbf{b}$$

Hence based on Einstein summation, $a_i b_i$ represents a dot product $\mathbf{a} \cdot \mathbf{b}$. This align with that $a_i b_i$ is zero order tensor[4] "constant". ■

[2]that is, x^2 should be understood as the second component of x not the square of x

[3]monom $\ldots + \overbrace{z^n y^n x^n}^{\text{monom}} + \ldots$

[4]since no live indices, refer to rule two

Example 3 *Let $n = 3$. Write out what is represented by this $a_{ij}b_{kj} = c_{ik}$*

We can see in $a_{ij}b_{kj}$ we have a *dummy* index j, which imply that we have a summation, hence the expression reads

$$c_{ik} = \sum_{j=1}^{3} a_{ij}b_{kj}$$

Computing this expression

$$\begin{aligned}
c_{11} &= \sum_{j=1}^{3} a_{1j}b_{1j} = a_{11}b_{11} + a_{12}b_{12} + a_{13}b_{13} \\
&\vdots \\
c_{13} &= \sum_{j=1}^{3} a_{1j}b_{3j} = a_{11}b_{31} + a_{12}b_{32} + a_{13}b_{33} \\
c_{21} &= \sum_{j=1}^{3} a_{2j}b_{1j} = a_{21}b_{11} + a_{22}b_{12} + a_{23}b_{13} \\
&\vdots \\
c_{33} &= \sum_{j=1}^{3} a_{3j}b_{3j} = a_{31}b_{31} + a_{32}b_{32} + a_{33}b_{33}
\end{aligned}$$

We can see that the results of this expression is a second order tensor c_{ik} since we have two live indices "it could be a matrix but not necessarily". ■

Example 4 *Let $n = 2$. Write out what is represented by $d_i = a_{ij}b_{jk}c_k$*

We have two dummy indices j, k hence this imply two summations

$$d_i = \sum_{j=1}^{2}\sum_{k=1}^{2} a_{ij}b_{jk}c_k$$

Computing this expression

$$\begin{aligned}
d_1 = \sum_{j=1}^{2}\sum_{k=1}^{2} a_{1j}b_{jk}c_k &= \sum_{j=1}^{2} a_{1j}b_{j1}c_1 + a_{1j}b_{j2}c_2 \\
&= (a_{11}b_{11}c_1 + a_{11}b_{12}c_2) + (a_{12}b_{21}c_1 + a_{12}b_{22}c_2) \\
d_2 = \sum_{j=1}^{2}\sum_{k=1}^{2} a_{2j}b_{jk}c_k &= \sum_{j=1}^{2} a_{2j}b_{j1}c_1 + a_{2j}b_{j2}c_2 \\
&= (a_{21}b_{11}c_1 + a_{21}b_{12}c_2) + (a_{22}b_{21}c_1 + a_{22}b_{22}c_2)
\end{aligned}$$

The results show that d_i is a first order tensor, since we have one live index. ■

Example 5 *Suppose* $Q = b_{ij}y_ix_j$ *and* $y_i = a_{ij}x_j$*. Substitute* y_i *into* Q *and simplify*

To avoid overlap of the dummy indices in y_i and Q, we will rename the dummy indices in y_i (i.e. $j \to k$)[5]

$$y_i = a_{ik}x_k$$

Substituting

$$Q = b_{ij}a_{ik}x_jx_k$$

Note that we have no live indices, which align with that Q is a zero order tensor. We have 3 dummy indices which imply three sums. ■

References

[5]you can do it the other way around (i.e. in Q rename $j \to k$)