Vectors and Matrices Notes.

Jonathan Coulthard
Jonathan.Coulthard@physics.ox.ac.uk

1 Index Notation

Index notation may seem quite intimidating at first, but once you get used to it, it will allow us to prove some very tricky vector and matrix identities with very little effort. As with most things, it will only become clearer with practice, and so it is a good idea to work through the examples for yourself, and try out some of the exercises.

Example: Scalar Product

Let's start off with the simplest possible example: the dot product. For real column vectors **a** and **b**,

$$\mathbf{a} \cdot \mathbf{b} = \mathbf{a}^T \mathbf{b} = \begin{pmatrix} a_1 & a_2 & a_3 & \cdots \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \end{pmatrix} = a_1 b_1 + a_2 b_2 + a_3 b_3 + \cdots \tag{1}$$

or, written in a more compact notation

$$\mathbf{a} \cdot \mathbf{b} = \sum_i a_i b_i, \tag{2}$$

where the σ means that we sum over all values of i.

Example: Matrix-Column Vector Product

Now let's take matrix-column vector multiplication, $\mathbf{Ax} = \mathbf{b}$.

$$\begin{bmatrix} A_{11} & A_{12} & A_{23} & \cdots \\ A_{21} & A_{22} & A_{23} & \cdots \\ A_{31} & A_{32} & A_{33} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \end{bmatrix} \tag{3}$$

You are probably used to multiplying matrices by visualising multiplying the elements highlighted in the red boxes. Written out explicitly, this is

$$b_2 = A_{21} x_1 + A_{22} x_2 + A_{23} x_3 + \cdots \tag{4}$$

If we were to shift the A box and the b box down one place, we would instead get

$$b_3 = A_{31} x_1 + A_{32} x_2 + A_{33} x_3 + \cdots \tag{5}$$

It should be clear then, that in general, for the ith element of b, we can write

$$b_i = A_{i1}x_1 + A_{i2}x_2 + A_{i3}x_3 + \cdots \tag{6}$$

Or, in our more compact notation,

$$b_i = \sum_j A_{ij}x_j. \tag{7}$$

Note that if the matrix $\mathbf{A}$ had only one column, then i would take only one value ($i = 1$). $\mathbf{b}$ would then also only have one element (b_1) making it a scalar. Our matrix-column vector product would therefore reduce exactly to a dot product. In fact, we can interpret the elements b_i as the dot product between the row vector which is the ith row of $\mathbf{A}$, and the column vector $\mathbf{x}$.

Example: Matrix-Matrix multiplication

One last simple example before we start proving some more nontrivial stuff. Consider the matrix product $\mathbf{AB} = \mathbf{C}$.

$$\begin{bmatrix} A_{11} & A_{12} & A_{23} & \cdots \\ A_{21} & A_{22} & A_{23} & \cdots \\ \boxed{A_{31} \quad A_{32} \quad A_{33} \quad \cdots} \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} B_{11} & \boxed{B_{12}} & B_{23} & \cdots \\ B_{21} & B_{22} & B_{23} & \cdots \\ B_{31} & B_{32} & B_{33} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} = \begin{bmatrix} C_{11} & C_{12} & C_{23} & \cdots \\ C_{21} & C_{22} & C_{23} & \cdots \\ C_{31} & \boxed{C_{32}} & C_{33} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \tag{8}$$

I have once again marked with a box the way that you are probably used to seeing these multiplications done. Explicitly,

$$C_{32} = A_{31}B_{12} + A_{32}B_{22} + A_{33}B_{32} + \cdots \tag{9}$$

As with the previous example, you might interpret C_{23} as the dot product between the row vector A_{3i}, and the column vector B_{i2}, i.e.

$$C_{23} = \sum_k A_{2k}B_{k3} \tag{10}$$

it is clear how this generalises to any element C_{ij},

$$C_{ij} = \sum_k A_{ik}B_{kj} \tag{11}$$

The rule for matrix multiplication is: "make the inner index (k, in this case) the same, and sum over it."

Example: Trace of a product of matrices.

The trace of a matrix is defined to be the sum of its diagonal elements

$$\mathrm{Tr}(\mathbf{C}) = \sum_i C_{ii} \tag{12}$$

We would like to prove that

$$\mathrm{Tr}(\mathbf{AB}) = \mathrm{Tr}(\mathbf{BA}) \tag{13}$$

First, let's take the definition of the matrix product (in index form) and plug it into the definiton of the trace, i.e. plug

$$C_{ij} = \sum_k A_{ik} B_{kj} \tag{14}$$

into

$$\mathrm{Tr}(\mathbf{C}) = \sum_i C_{ii}. \tag{15}$$

We obtain

$$\mathrm{Tr}(\mathbf{AB}) = \sum_i \left(\sum_k A_{ik} B_{ki} \right) \tag{16}$$

Now A_{ik}, and B_{ki} are just scalars (i.e. single elements of the matrices $\mathbf{A}$ and $\mathbf{B}$ respectively), and so we can commute them.

$$\mathrm{Tr}(\mathbf{AB}) = \sum_i \left(\sum_k B_{ki} A_{ik} \right) \tag{17}$$

Now, the thing inside the brackets is *almost* a matrix product, but we are summing over the wrong index (the outer index rather than the inner one). The question is, can we swap the order of the two summations? Because addition is commutative ($a + b = b + a$), we can [1].

Exercise: By writing out the sums with a small number of terms, convince yourself that you can indeed commute $\sum_i$ and $\sum_k$.

Finally, by swapping $\sum_i$ and $\sum_k$, we obtain

$$\begin{aligned} \mathrm{Tr}(\mathbf{AB}) &= \sum_k \left(\sum_i B_{ki} A_{ik} \right), \\ &= \sum_k (\mathbf{BA})_{kk}, \\ &= \mathrm{Tr}(\mathbf{BA}). \end{aligned} \tag{18}$$

Exercise: Now try a slightly more complicated example for yourself. Using index notation, prove that $\mathrm{Tr}(\mathbf{AB} \cdots \mathbf{YZ})$ is invariant under cyclic permutations of the matrices.

1.1 Einstein Summation convention

Our notation is much more compact than writing out huge matrices and trying to figure out how the multiplications, etc. work in general. However, writing out Σs can become very cumbersome. Since we have

$$\sum_i \sum_j (\text{stuff}) = \sum_i \sum_j (\text{stuff}) \tag{19}$$

$$C \sum_j (\text{stuff}) = \sum_i \sum_j C(\text{stuff}), \tag{20}$$

$$\tag{21}$$

we can just drop the Σs entirely, and adopt what is called Einstein summation convention. It is simply summarised as follows:

[1] Small health warning: Swapping the order of the summations amounts to reordering the terms in the sum. This is fine if we are just summing over a finite number of terms. However, if the sums are infinite, we can only reorder the terms if the sum converges *absolutely*. If it converges only *conditionally*, then we *cannot* reorder the terms (see wikipedia for the definitions of these terms). You are unlikely to encounter a situation where this is an issue, but be aware that such situations exist.

- If an index appears twice in a term, then it is implied that we sum over it. For example, $A_{ij}B_{jk} \equiv \sum_j A_{ij}B_{jk}$.

This is a very powerful convention that you will use extensively when you learn special and general relativity in your third year. However, in the immediate future, it can also make proofs of various matrix and vector identities very easy! For instance our proof of $\mathrm{Tr}(\mathbf{AB}) = \mathrm{Tr}(\mathbf{BA})$. can be written as

$$\begin{aligned}\mathrm{Tr}(\mathbf{AB}) &= A_{ik}B_{ki} \\ &= B_{ki}A_{ik} \qquad (22) \\ &= \mathrm{Tr}(\mathbf{BA}). \qquad (23)\end{aligned}$$

(Note that in a problem sheet, or in an exam, you should explain each line in this proof—the notation makes it look much more trivial than it really is!)

1.1.1 Tips

As a wise man once said: "With great power comes great responsibility." While Einstein summation convention can indeed make our lives much easier, it can also produce a great deal of nonsense if you are not very careful when keeping track of your indices. Here are some tips for doing so:

- *Free indices* appear only once in an expression and thus are not summed over. *Dummy indices* appear twice, and are implicitly summed over.
- To help avoid confusion, it is a good idea to use roman letters (i, j, k) for free indices, and greek letters (λ, μ, ν) for dummy indicies.
- *Dummy* indices should *never* appear in the "final answer".
- The *free indices* should *always* be the same in every term in an expression.
- An index should *never* appear more than twice in a single term.

1.1.2 The Kronecker Delta

The Kronecker delta is a useful symbol which crops up all the time. It is defined as

$$\delta_{ij} = \begin{cases} 1, & \text{if } i = j \\ 0, & \text{otherwise} \end{cases} \qquad (24)$$

It should be clear that this is basically a representation of the identity matrix. It also has the useful property that if you sum over one of the indices, then it kills the sum, and replaces the dummy index with the other (free) index. For example,

$$\sum_j a_{ij}\delta_{jn} = a_{in}. \qquad (25)$$

(since the only nonzero term in the sum is $j = n$). If we were to use Einstein summation convention, then we would write the above as

$$a_{i\mu}\delta_{\mu n} = a_{in}. \qquad (26)$$

As another example, the scalar product between two vectors $\mathbf{a} \cdot \mathbf{b}$ can be written as:

$$\mathbf{a} \cdot \mathbf{b} = a_\mu b_\nu \delta_{\mu\nu} = a_\mu b_\mu \qquad (27)$$

Where the $\delta_{\mu\nu}$ forces the indices of a and b to be equal.

1.1.3 The Levi-Civita symbol

The Levi-Civita symbol, ε_{ijk} is another handy object. It is defined as

$$\varepsilon_{ijk} = \begin{cases} 1, & \text{if } ijk \text{ is a cyclic permutation of 123} \\ -1, & \text{if } ijk \text{ is an anticyclic permutation of 123} \\ 0, & \text{if any two of } i, j, \text{ or } k \text{ are equal.} \end{cases} \tag{28}$$

Why is such a thing useful? Well, let's consider the object $\varepsilon_{i\mu\nu}a_\mu b_\nu$. It has one free index, i, so it is a vector. What are its components? If we set $i = 1$, from the definition of ε we see that the only nonzero terms in the sum will be $\mu, \nu \neq 0$. This leaves us with $\mu, \nu = 2$ or 3. Writing out these components explicitly, we find

$$\varepsilon_{1\mu\nu}a_\mu b_\nu = \varepsilon_{123}a_2 b_3 + \varepsilon_{132}a_3 b_2 \tag{29}$$
$$= a_2 b_3 - a_3 b_2. \tag{30}$$

Which you probably recognise as the first component of the vector $\mathbf{a} \times \mathbf{b}$. Indeed, if we go through the components, we will indeed find that

$$\varepsilon_{i\mu\nu}a_\mu b_\nu = (\mathbf{a} \times \mathbf{b})_i. \tag{31}$$

An identity you will find useful is

$$\varepsilon_{\mu jk}\varepsilon_{\mu lm} = \delta_{jl}\delta_{km} - \delta_{jm}\delta_{kl}. \tag{32}$$

The proof of this identity is just the evaluation of all of the cases (thankfully most of them are zero!).

Exercise: Prove the identity (32).

Example: The scalar triple-product

In the problem sheet on vectors, we (without proof) made use of the fact that the scalar triple-product, $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})$, is unchanged under cyclic permutations of the vectors **a**, **b**, and **c**. Armed with our new notation, proving this becomes trivial!

$$\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = \varepsilon_{\lambda\mu\nu}a_\lambda b_\mu c_\nu \tag{33}$$
$$= \varepsilon_{\nu\lambda\mu}a_\lambda b_\mu c_\nu \tag{34}$$
$$= \varepsilon_{\nu\lambda\mu}c_\nu a_\lambda b_\mu \tag{35}$$
$$= \mathbf{c} \cdot (\mathbf{a} \times \mathbf{b}) \tag{36}$$

We immediately see that (since everything on the right commutes), cyclicly permuting the vectors corresponds to cyclicly permuting the indices $\lambda\mu\nu$ in the ε. Since $\epsilon_{\lambda\mu\nu}$ is unchanged under cyclic permutations, $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})$ is unchanged under cyclic permutations of the vectors!

Exercise: Which of the following expressions are valid? If possible, write them in matrix-column vector notation. If they *are not* valid, why not?

1. $\nabla_\mu x_\mu$
2. $A_{i\sigma}x_\sigma$
3. $\varepsilon_{\lambda\mu\nu}x_{\lambda\mu}$

4. $a_i b_j + c_k d_j$

5. $\varepsilon_{i\mu\nu}\varepsilon_{\nu\sigma\rho}a_\mu b_\sigma c_\rho$

You now might want to go back and have a look at the problem sheet from week 2 "Vectors and matrices I" (specifically, Q. 24 of the class problems) for some more practice. Try proving them all with index notation and Einstein summation convention!