Index Notation and Rotation Operators

Ross Dempsey

September 2014

1 Index Notation

Index notation, also known as Einstein notation, is a compact form of representing complicated linear algebraic expressions in applications to physics. The most practically important feature is that repeated indices represent summation. For example, we can transform a simple matrix multiplication into Einstein notation like so:

$$\begin{aligned} A\boldsymbol{x} &= \boldsymbol{b} \\ \sum_j A^i_j x^j &= b_i \\ A^i_j x^j &= b_i \end{aligned}$$

Other than the removal of the cumbersome summation sign, this notation is unfamiliar in that it represents some indices as superscripts. The primary virtue of the notation, other than its brevity, is that it codifies the difference between contravariant and covariant tensors by raising or lowering indices. For completeness' sake, we'll briefly discuss this difference.

1.1 Contravariance and Covariance

The way in which a vector's components change in a change of basis determines whether it is contra- or covariant. If we start in a basis $\mathcal{B} = \{\boldsymbol{b_1}, \boldsymbol{b_2}, \ldots, \boldsymbol{b_n}\}$, then a contravariant vector $\boldsymbol{v}$ can be represented in terms of its components as

$$\boldsymbol{v} = v^i b_i$$

Note that we write the index on top for a contravariant vector; this is the standard convention of Einstein notation. The majority of vectors encounted in basic physics are in fact contravariant. For example, the displacement vector $\boldsymbol{x}$ is a contravariant vector. We can explore what this means by performing a change of basis. Given a displacement vector $\boldsymbol{x}$ with components x^i, we seek to find the new components in a second basis, $\mathcal{B}' = \{b'_1, b'_2, \ldots, b'_n\}$. Let the matrix describing the change of basis be A, so that

$$b'_j = b_i A^i_j$$

Then we seek to find the new set of coefficients $x^{j'}$ such that

$$x^i b_i = \boldsymbol{x} = x^{j'} b'_j \tag{1}$$

We can rewrite the right hand side using the relationship between the bases. Thus,

$$x^i b_i = x^{j'} \left(b_i A^i_j\right) = \left(A^i_j x^{j'}\right) b_i \tag{2}$$

This shows that the components of $\boldsymbol{x}$ transform in the opposite direction as the basis vectors; i.e., converting from the original components to the primed components would require multiplication by A^{-1}.

On the contrary, covariant vectors have components that transform with the coordinate system. They can be represented as linear functionals or one-forms; that is, mappings from the vector space to its underlying scalar field. Since they are linear, they can be characterized by their actions on the basis vectors; in fact, for a functional $\boldsymbol{\alpha}$, the components can be considered to be

$$\alpha_i = \boldsymbol{\alpha}\left(b_i\right)$$

Note that we use lower subscripts to denote the components of a covariant vector, in accordance with Einstein notation. Now, we perform the change of basis and invoke the multilinearity of α:

$$\begin{aligned} \alpha'_i &= \boldsymbol{\alpha}\left(b'_i\right) \\ &= \boldsymbol{\alpha}\left(b_j A^j_i\right) \\ &= A^j_i \boldsymbol{\alpha}\left(b_j\right) \\ &= A^j_i \alpha_j \end{aligned}$$

Thus, the components of $\boldsymbol{\alpha}$ transform with the basis vectors.

In simple situations, it is easy to determine whether a vector is contravariant or covariant. Consider a scaling of the coordinate system; i.e., for the case of a displacement vector, change the units from meters to millimeters. This obviously leads to a scaling of the basis vectors by 0.001 and a scaling of the vector components by 1000; since these transformations are inverse, the vector is contravariant. Alternatively, however, define a temperature field T and then consider the vector ∇T. If the units of the basis are changed from meters to millimeters, then both the basis vectors and ∇T will be scaled by 0.001 (because each component of ∇T represents a change in T *per unit length*). Thus, ∇T (and gradient vectors in general) are covariant.

In nearly all cases, an object known as the metric tensor $g_{\mu\nu}$ exists which effectively permits the raising and lowering of indices at will. Thus, considerations of contravariance and covariance are unimportant in the following sections.

2 Deriving the Rotation Matrix

We start with a vector expression that describes rotations of a vector about an arbitrary axis $\hat{\boldsymbol{n}}$ with arbitrary angle θ:

$$\boldsymbol{r}'(\hat{\boldsymbol{n}}, \theta) = (\boldsymbol{r} \cdot \hat{\boldsymbol{n}})\hat{\boldsymbol{n}} + \cos\theta\,(\boldsymbol{r} - (\boldsymbol{r} \cdot \hat{\boldsymbol{n}})\hat{\boldsymbol{n}}) + \sin\theta\,(\hat{\boldsymbol{n}} \times \boldsymbol{r})$$

Now we proceed to convert this expression into an indexed representation of a matrix formula. We start by defining the Kronecker delta and the Levi-Civita symbol.

2.1 Kronecker Delta

Ordinarily, the Kronecker delta δ_{ij} is a simple function of i and j:

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

In this context, we interpret it as a tensor. Its meaning remains the same, except of course that if one or more of its indices is repeated it is summed over. When both of its indices are contracted, it becomes a representation for the dot product:

$$\boldsymbol{u} \cdot \boldsymbol{v} = \delta_{ij} u^i v^j$$

2.2 Levi-Civita Symbol

The Levi-Civita symbol is a tensor codifying the cross product. Consider computing the cross product of $\boldsymbol{u}$ and $\boldsymbol{v}$ using a determinant:

$$\boldsymbol{u} \times \boldsymbol{v} = \begin{vmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ u^1 & u^2 & u^3 \\ v^1 & v^2 & v^3 \end{vmatrix} = |A|$$

From linear algebra, we know that the determinant can be expressed as a sum of permutation products. That is, for all permutations (a, b, c) of the indices $(1, 2, 3)$, we form the product $A_{1a} A_{2b} A_{3c}$ and negate it if (a, b, c) is an odd permutation. Summing over all permutations yields the determinant. The Levi-Civita symbol encodes this algorithm. The value of ϵ_{ijk} is defined to be 0 if any of the indices are equal (i.e., if (i, j, k) is not a permutation), 1 if (i, j, k) is an even permutation of $(1, 2, 3)$, and -1 if it is an odd permutation. Thus, contracting ϵ_{ijk} with two vectors gives a component of their cross product:

$$(\boldsymbol{u} \times \boldsymbol{v})_i = \epsilon^i_{jk} u^j v^k$$

2.3 Manipulating the Vector Formula

Equipped with these two tensors, we can convert the vector formula above to matrix form. We start by utilizing the Levi-Civita symbol to rewrite the third term:

$$\sin\theta \, (\hat{\boldsymbol{n}} \times \boldsymbol{r})_i = \sin\theta \left(\epsilon^i_{jk} \hat{n}^j r^k \right)$$

We then rewrite each dot product in the first and second terms using a δ symbol, and obtain a matrix representation of the formula:

$$\begin{aligned} r'^i(\hat{\boldsymbol{n}}, \theta) &= \left(\delta_{jk} r^j \hat{n}^k \right) \hat{n}^i + \cos\theta \left(r^i - \left(\delta_{jk} r^j \hat{n}^k \right) \hat{n}^i \right) + \sin\theta \left(\epsilon^i_{jk} \hat{n}^j r^k \right) \\ &= r^i + (\cos\theta - 1) \left(r^i - \left(\delta_{jk} r^j \hat{n}^k \right) \hat{n}^i \right) + \sin\theta \left(\epsilon^i_{jk} \hat{n}^j r^k \right) \end{aligned}$$

We can at this point define 3 operators to simplify the notation of the formula. They are $I = \delta^\mu_i$, the identity operator; $P_\perp = I - \delta_{ik} \hat{n}^k \hat{n}^\mu$, the perpendicular projection operator for $\hat{n}$; and $C = \epsilon^\mu_{ji} \hat{n}^j$, the cross product operator for $\hat{n}$. We can combine these operators into a single composite rotation operator,

$$R(\hat{\boldsymbol{n}}, \theta) = I + (\cos\theta - 1) \, P_\perp + (\sin\theta) C$$

Then the original vector equation can be written compactly as

$$r'^\mu(\hat{\boldsymbol{n}}, \theta) = R(\hat{\boldsymbol{n}}, \theta) r^i$$

3 Combining

Starting from the three operators I, $P_\perp$, and C, we seek to determine the products of all pairs. This amounts to forming a multiplication table for the operators. We start by stating without proof (the proof is very simple) that I serves as the identity of this group, so that $IX = XI = X$ for all X. We can then proceed to fill in the nontrivial sections of the multiplication table, starting with the square of the perpendicular projection operator. Note that because it is a projection operator, we expect to find $P_\perp^2 = P_\perp$.

$$
\begin{aligned}
P_\perp^2 &= \left(\delta_i^\mu - \delta_{ik}\hat{n}^k\hat{n}^\mu\right)\left(\delta_i^\nu - \delta_{il}\hat{n}^l\hat{n}^\nu\right) \\
&= \delta^{\mu\nu} - \delta_{\mu l}\hat{n}^l\hat{n}^\nu - \delta_{\nu k}\hat{n}^k\hat{n}^\mu + \delta_{kl}\hat{n}^k\hat{n}^l\hat{n}^\mu\hat{n}^\nu
\end{aligned}
$$

At this point, we notice the $\delta_{kl}\hat{n}^k\hat{n}^l$ factor in the last term is simply $|\hat{\boldsymbol{n}}|$, which we have assumed to be 1. In addition, we contract the second term into $-\hat{n}^\mu\hat{n}^\nu$, so that it cancels the last term. This leaves us with

$$
P_\perp^2 = \delta_\nu^\mu - \delta_{\nu k}\hat{n}^k\hat{n}^\mu
$$

This is clearly of the same form as $P_\perp$; to formally show the equality, we could multiply by δ_i^ν. Next, we find the square of C.

$$
\begin{aligned}
C^2 &= \epsilon_{ji}^\mu\hat{n}^j\epsilon_{ki}^\mu\hat{n}^k = \left(\epsilon_{j\nu}^\mu\epsilon_{ki}^\nu\right)\hat{n}^j\hat{n}^k \\
&= \left(\delta_{\mu k}\delta_{ij} - \delta_{\mu i}\delta_{jk}\right)\hat{n}^j\hat{n}^k \\
&= \delta_{ij}\hat{n}^j\hat{n}^\mu - \delta_i^\mu
\end{aligned}
$$

We see that this is equal to $-P_\perp$. The product of C with $P_\perp$ is easily obtained.

$$
\begin{aligned}
CP_\perp &= \epsilon_{j\nu}^\mu\hat{n}^j\left(\delta_i^\nu - \delta_{ik}\hat{n}^k\hat{n}^\nu\right) \\
&= \epsilon_{ji}^\mu\hat{n}^j - \epsilon_{j\nu}^\mu\hat{n}^j\hat{n}^\nu\delta_{ik}\hat{n}^k
\end{aligned}
$$

The second term vanishes because it contains a component of $\hat{\boldsymbol{n}} \times \hat{\boldsymbol{n}}$. Thus, $CP_\perp = C$. The same result can be obtained in a similar fashion for $P_\perp C$.

4 Simple Application

We can utilize the multiplication table we have derived to prove that $R^2(\hat{n},\theta) = R(\hat{n},2\theta)$. This follows by direct multplication.

$$
\begin{aligned}
R^2(\hat{n},\theta) &= R(\hat{n},\theta) + (\cos\theta - 1)\left(P_\perp + (\cos\theta - 1)P_\perp + \sin\theta C\right) \\
&\quad + \sin\theta\left(C + (\cos\theta - 1)C - \sin\theta P_\perp\right) \\
&= R(\hat{n},\theta) + (\cos\theta - 1)\left(\cos\theta P_\perp + \sin\theta C\right) + \sin\theta\left(\cos\theta C - \sin\theta P_\perp\right) \\
&= R(\hat{n},\theta) + (\cos^2\theta - \sin^2\theta)P_\perp + 2\sin\theta\cos\theta C - \cos\theta P_\perp - \sin\theta C \\
&= I + (\cos 2\theta - 1)P_\perp + \sin 2\theta C \\
&= R(\hat{n},2\theta)
\end{aligned}
$$