

Introduction to tensors and indicial notation

Michael Raulli

1 Tensors and tensor multiplication in indicial notation

Indicial notation is a compact way of writing systems of equations. It can be used as a replacement for longhand writing of equations or matrix representation. A matrix is more valuable for representing the storage of values in the system, but for writing equations in a compact form, and especially for higher order tensors, indicial notation is superior.

Unless otherwise stated, all indices go between 1 and 3. Boldface type lowercase letters are matrices and boldface uppercase letters are matrices. Let us start with the simple representation of a vector in indicial notation:

$$\mathbf{a} = a_i \quad i = 1 \dots 3 \quad (1)$$

The three entries in the vector $\mathbf{a}$ are represented by a_1, a_2, a_3 . *The index letter is incremented from its starting value (1) to its ending value (3).* If $i = 1 \dots 1$, then a_i is a scalar, also called a 0th order tensor. A scalar can also be represented simply by a . If $i = 1 \dots k$ and $k > 1$ then a_i is a vector, also called a 1st order tensor.

Let us look at scalar-vector multiplication in indicial notation:

$$c \cdot \mathbf{a} = \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} c \cdot a_1 \\ c \cdot a_2 \\ c \cdot a_3 \end{bmatrix} \quad (2)$$

$$c \cdot a_i = b_i; \quad b_1 = c \cdot a_1; \quad b_2 = c \cdot a_2; \quad b_3 = c \cdot a_3; \quad (3)$$

It is simple to see that the index (i) is incremented between 1 and 3 in equation (3), and the multiplication is carried out, giving the same answer as in equation (2).

Note: The number of indices indicates the order of the tensor. The scalar (c) does not have an index, indicating that it is a 0th order tensor. The vector ($\mathbf{a}$) has one index (i), indicating that it is a 1st order tensor. This is trivial for this case, but becomes useful later.

Let us examine the vector dot product, which has a scalar result. Here we learn a new feature of index notation: *sum over repeated indices*.

$$\mathbf{a} \cdot \mathbf{b} = \begin{bmatrix} a_1 & a_2 & a_3 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = a_1 b_1 + a_2 b_2 + a_3 b_3 = c \quad (4)$$

$$a_i b_i = a_1 b_1 + a_2 b_2 + a_3 b_3 = c \quad (5)$$

It is easily seen that as the index i is incremented, the multiplication takes place and the result is added together. For $i = 1$, we multiply a_1 and b_1 , then i is incremented and we multiply a_2 and b_2 , then i is incremented and we multiply a_3 and b_3 . The repeated index i indicates that the results of the three multiplications should be added.

Note: There are two important things that are shown by the indicial notation in equation (5), but not by the vector notation in equation (4). The first is that the repeated index in a_i and b_i shows that the dimensions of a and b must be the same. The vector notation does not show this, even though one should know that this needs to be the case for this simple example. This becomes more useful in higher order tensors. The second piece of information is the fact that the final result will be a 0^{th} order tensor, or scalar. Anytime there are repeated indices in a tensor multiplication, this degree of the tensor will be canceled out. So in the case where two 1^{st} order tensors are multiplied, the index i cancels out, and this shows that the final result will not have an index, therefore it is a scalar result. Again, these observations may be apparent here, but they are very useful for higher order tensors. The repeated index is called a *dummy index*.

Let us take another step and examine matrices, or 2^{nd} order tensors:

$$\mathbf{A} = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} = A_{ij} \quad i = 1 \dots 3; j = 1 \dots 3 \quad (6)$$

There are now two unique indices (i and j). As noted previously, the number of unique indices indicates the order of the tensor, 2^{nd} order, in this case.

Note: The total number of entries in a tensor can be determined by multiplying the range of all unique indices. Since $i = 1 \dots 3$ and $j = 1 \dots 3$ above, the total number of entries in the tensor is $3 \cdot 3 = 9$, this is confirmed by the 3x3 matrix above.

Let us examine scalar-matrix multiplication:

$$c\mathbf{A} = \begin{bmatrix} cA_{11} & cA_{12} & cA_{13} \\ cA_{21} & cA_{22} & cA_{23} \\ cA_{31} & cA_{32} & cA_{33} \end{bmatrix} = \mathbf{B} = \begin{bmatrix} B_{11} & B_{12} & B_{13} \\ B_{21} & B_{22} & B_{23} \\ B_{31} & B_{32} & B_{33} \end{bmatrix} \quad (7)$$

$$B_{ij} = cA_{ij} \quad (8)$$

This case is simple, since the same scalar multiplies each entry in A_{ij} as i and j are incremented. There are not any repeated indices, therefore no summation. Also, there is no reduction in the order of the tensor, as indicated by the fact that there are no repeated indices. Finally, since the scalar does not have any indices, the bounds on i and j do not matter as far as compatible dimensions are concerned.

For the matrix-vector case, there are several things made apparent by the indicial notation:

$$\mathbf{d} = \mathbf{A} \cdot \mathbf{b} \quad (9)$$

$$d_i = A_{ij}b_j \quad (10)$$

The repeated index (j) indicates that you should sum over the index j (this is illustrated in detail below). It also indicates that the the second dimension of the matrix must match that of the vector. Finally, the repeated index will be canceled, leaving only the index i . One index indicates a 1st order tensor, or a vector, which is, of course, what happens when a matrix and vector are multiplied. Let us carry out the multiplication in equation (10) in detail:

$i = 1; j = 1 \dots 3;$ Keep i fixed and sum for all values of j as it is incremented:

$$d_1 = A_{11}b_1 + A_{12}b_2 + A_{13}b_3 \quad (11)$$

$i = 2; j = 1 \dots 3;$ Keep i fixed, reset j and sum for all values of j as it is incremented:

$$d_2 = A_{21}b_1 + A_{22}b_2 + A_{23}b_3 \quad (12)$$

$i = 3; j = 1 \dots 3;$ Keep i fixed, reset j and sum for all values of j as it is incremented:

$$d_3 = A_{31}b_1 + A_{32}b_2 + A_{33}b_3 \quad (13)$$

This is the same procedure as matrix-vector multiplication, except we are following the indicial notation rules, which say *sum over repeated indices*. This procedure also illustrates another important technique: *increment the repeated indices while keeping the unique indices fixed; when the repeated indices have been fully incremented, increment the unique index, reset the repeated index, and the increment the repeated index fully again. Repeat this until all indices have been fully incremented at least once.*

Matrix-matrix multiplication follows in a similar manner, with an added index:

$$\mathbf{D} = \mathbf{A} \cdot \mathbf{B} \quad (14)$$

$$D_{ik} = A_{ij}B_{jk} \quad i = 1 \dots 2; j = 1 \dots 3; k = 1 \dots 2 \quad (15)$$

What do these indices tell us? We must sum over the index j . The second dimension of A (a 2x3 2nd order tensor) and the first dimension of B (a 3x2 2nd order tensor) must match. Finally, the resulting tensor D , will be a 2x2 2nd order, because the index j cancels out and we are left with two unique indices (i and k), each with a dimension of 2. The indices i and k can be of arbitrary size. Let us examine the mechanics of the multiplication in detail:

$i = 1; k = 1; j = 1 \dots 3;$ Keep i and k fixed and sum for all values of j as it is incremented:

$$D_{11} = A_{11}B_{11} + A_{12}B_{21} + A_{13}B_{31} \quad (16)$$

$i = 1; k = 2; j = 1 \dots 3;$ Increment k , reset j , then keep i and k fixed and sum over j as it is incremented:

$$D_{12} = A_{11}B_{12} + A_{12}B_{22} + A_{13}B_{32} \quad (17)$$

$i = 2; k = 1; j = 1 \dots 3;$ Increment i , reset j and k , then keep i and k fixed and sum over j as it is incremented:

$$D_{21} = A_{21}B_{11} + A_{22}B_{21} + A_{23}B_{31} \quad (18)$$

$i = 2; k = 2; j = 1 \dots 3;$ Increment k , reset j , then keep i and k fixed and sum over j as it is incremented:

$$D_{22} = A_{21}B_{12} + A_{22}B_{22} + A_{23}B_{32} \quad (19)$$

It can be readily verified that the indicial procedure is the same as a standard matrix-vector multiplication. Again, the repeated index is incremented while the unique indices are fixed. The unique indices are also incremented. Even though k as incremented fully twice, and i only once, the results would be the same if the procedure was reversed.

2 Derivatives in indicial notation

The indication of derivatives of tensors is simply illustrated in indicial notation by a comma.

2.1 Gradients of scalar functions

The definition of the gradient of a scalar function is used as illustration. The Cartesian coordinates x, y, z are replaced by x_1, x_2, x_3 in order to facilitate the use of indicial notation.

$$\mathbf{g} = \nabla f = \left[\frac{\partial f}{\partial x_1} \quad \frac{\partial f}{\partial x_2} \quad \frac{\partial f}{\partial x_3} \right] \quad (20)$$

$$g_i = f_{,i} = \frac{\partial f}{\partial x_i} \quad i = 1 \dots 3 \quad (21)$$

The two definitions are equivalent. The comma in the indicial notation indicates to take the derivative of f with respect to each coordinate x_i , which is the definition of a gradient. Since there are no repeated indices, there is no summation. Again, the number of unique indices still indicates what the order of the resulting tensor will be. In this case, there is one unique index (i), therefore a 1st order tensor results, as is the case for gradients of scalar functions.

2.2 Divergence of vector

The divergence of a vector is defined below:

$$h = \nabla \cdot \mathbf{g} = \left[\frac{\partial}{\partial x_1} \quad \frac{\partial}{\partial x_2} \quad \frac{\partial}{\partial x_3} \right] \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \frac{\partial f}{\partial x_3} \end{bmatrix} = \frac{\partial^2 f}{\partial x_1^2} + \frac{\partial^2 f}{\partial x_2^2} + \frac{\partial^2 f}{\partial x_3^2} = \nabla^2 f \quad (22)$$

$$h = g_{i,i} = \frac{\partial g_i}{\partial x_i} = \frac{\partial^2 f}{\partial x_1^2} + \frac{\partial^2 f}{\partial x_2^2} + \frac{\partial^2 f}{\partial x_3^2} = f_{,ii} \quad (23)$$

Again, the definitions are equivalent. The notation $g_{i,i}$ shows that we have started with a vector (g_i) and that for each value of i , a derivative with respect to x_i should be taken. Since the index i is repeated, we sum over it. Also, since the index i is repeated, they cancel and the result is a scalar. The notation $f_{,ii}$ indicates that the second derivative with respect to x_i should be taken and summed, giving the same results as $g_{i,i}$. This is where the real confusion and utility of index notation can start.

2.3 Gradients of vector functions

The gradient of a vector function is defined as follows:

$$\nabla \mathbf{f} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \frac{\partial f_1}{\partial x_3} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \frac{\partial f_2}{\partial x_3} \\ \frac{\partial f_3}{\partial x_1} & \frac{\partial f_3}{\partial x_2} & \frac{\partial f_3}{\partial x_3} \end{bmatrix} \quad (24)$$

$$f_{i,j} = \frac{\partial f_i}{\partial x_j} \quad (25)$$

Both definitions are equivalent. The two unique indices tell you that the result of $f_{i,j}$ will be a 2^{nd} order tensor.

3 Special operators

3.1 Kronecker delta function

The Kronecker delta is a convenient way of expressing the identity in indicial notation:

$$\begin{cases} \delta_{ij} = 1 & i = j \\ \delta_{ij} = 0 & i \neq j \end{cases} \quad (26)$$

The Kronecker delta follows all of the same rules of index notation.

$$A_{ik} = \delta_{ij} A_{jk} \quad (27)$$

Where i and j have the same dimensions, therefore $A_{ik} = A_{jk}$. You should check, using indicial multiplication rules, that the Kronecker delta really does function as an identity. Consider: why does $\delta_{ii} = 3$, not 1?

3.2 Permutation operator

The permutation symbol is also known as the Levi-Civita operator. It is defined as follows:

$$\epsilon_{ijk} = \begin{cases} +1 & \text{for a forward permutation of } ijk \\ -1 & \text{for a backward permutation of } ijk \\ 0 & \text{if } i, j \text{ or } k \text{ are equal} \end{cases} \quad (28)$$

Forward permutations are defined as three unique numbers which read 1,2,3 from left to right (allowing for wrapping), e.g. ϵ_{123} , ϵ_{312} , or ϵ_{231} . Backward permutations read 3,2,1 from left to right (allowing for wrapping), e.g. ϵ_{321} , ϵ_{132} , or ϵ_{213} . A useful relationship for the permutation operator is that:

$$\epsilon_{ijk} = -\epsilon_{ikj} \quad (29)$$

This does not indicate, however, which one is positive or negative (or zero), just that they are of opposite sign.

This more complex operator is useful for defining cross products:

$$\mathbf{d} = \begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix} = \mathbf{a} \times \mathbf{b} = \begin{bmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{bmatrix} \quad (30)$$

$$d_i = \epsilon_{ijk} a_j b_k \quad (31)$$

The two above definitions are identical, please check to verify this.

3.3 Kronecker-permutation relationship

A useful identity is the following:

$$\epsilon_{ijk}\epsilon_{ilm} = \delta_{jl}\delta_{km} - \delta_{jm}\delta_{kl} \quad (32)$$

4 Example

The use of indicial notation can be illustrated through the equations of linear elasticity. First, we define the displacements as a 1st order tensor, with components in the principal Cartesian directions, considering only plane stress, therefore two dimensional.

$$\text{displacements : } u_i \quad i = 1 \dots 2 \quad (33)$$

The well known linearized strains can be computed from the displacements:

$$\epsilon_{ij} = \frac{1}{2} (u_{i,j} + u_{j,i}) \quad j = 1 \dots 2 \quad (34)$$

There are no repeated indices (the same indices on the opposite sides of the plus sign do not count), therefore there is no summation (except for the plus sign, of course). Since there are two unique indices, the result is a 2nd order strain tensor. Written in matrix form, the strain tensor as computed from equation (34) looks as follows:

$$\epsilon_{ij} = \begin{bmatrix} \frac{\partial u_1}{\partial x_1} & \frac{1}{2} \left(\frac{\partial u_1}{\partial x_2} + \frac{\partial u_2}{\partial x_1} \right) \\ \frac{1}{2} \left(\frac{\partial u_2}{\partial x_1} + \frac{\partial u_1}{\partial x_2} \right) & \frac{\partial u_2}{\partial x_2} \end{bmatrix} \quad (35)$$

The constitutive relationship illustrates the power of the indicial notation, since the constitutive matrix is a 4th order tensor:

$$\sigma_{ij} = C_{ijkl} \epsilon_{kl} \quad (36)$$

Since the multiplication of a 4^{th} order tensor with a 2^{nd} order tensor is extremely difficult to visualize, the indicial notation is critical. There are two repeated indices on the righthand side (k and l), therefore they will be summed over. We already know that k and l go from $1 \dots 2$, and we also know that the stress (σ_{ij}) tensor is of the same dimensions. Therefore, all indices go from $1 \dots 2$. This means that the constitutive tensor is $2 \times 2 \times 2 \times 2$ and has 16 entries, for the simplified 2-d problem (81 entries in 3-d). Let us multiply out equation (36), for greater detail.

$$\sigma_{11} = C_{1111}\varepsilon_{11} + C_{1112}\varepsilon_{12} + C_{1121}\varepsilon_{21} + C_{1122}\varepsilon_{22} \quad (37)$$

$$\sigma_{12} = C_{1211}\varepsilon_{11} + C_{1212}\varepsilon_{12} + C_{1221}\varepsilon_{21} + C_{1222}\varepsilon_{22} \quad (38)$$

$$\sigma_{21} = C_{2111}\varepsilon_{11} + C_{2112}\varepsilon_{12} + C_{2121}\varepsilon_{21} + C_{2122}\varepsilon_{22} \quad (39)$$

$$\sigma_{22} = C_{2211}\varepsilon_{11} + C_{2212}\varepsilon_{12} + C_{2221}\varepsilon_{21} + C_{2222}\varepsilon_{22} \quad (40)$$

Notice how, i and j are held fixed, while k and l are fully incremented and summed over. A new equation is started when the values of i or j change. For isotropic, linear elastic materials, Hooke's law gives us the following relationships between stress and strain for plane stress:

$$\sigma_{xx} = \sigma_{11} = \frac{E}{1-\nu^2} (\varepsilon_{11} + \nu\varepsilon_{22}) \quad (41)$$

$$\tau_{xy} = \sigma_{12} = \frac{E}{1-\nu^2} (1-\nu) \varepsilon_{12} = 2G\varepsilon_{12} = G\gamma_{12} \quad (42)$$

$$\tau_{yx} = \sigma_{21} = \frac{E}{1-\nu^2} (1-\nu) \varepsilon_{21} = 2G\varepsilon_{21} = G\gamma_{21} \quad (43)$$

$$\sigma_{yy} = \sigma_{22} = \frac{E}{1-\nu^2} (\varepsilon_{22} + \nu\varepsilon_{11}) \quad (44)$$

Based on Hooke's law, we can determine the entries in the 4^{th} order tensor C_{ijkl} :

$$C_{ijkl} = \frac{E}{1-\nu^2} \begin{bmatrix} \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & \nu \end{bmatrix}}_{C_{11}} & \underbrace{\begin{bmatrix} 0 & 1-\nu \\ 0 & 0 \end{bmatrix}}_{C_{12}} \\ \underbrace{\begin{bmatrix} 0 & 0 \\ 1-\nu & 0 \end{bmatrix}}_{C_{21}} & \underbrace{\begin{bmatrix} \nu & 0 \\ 0 & 1 \end{bmatrix}}_{C_{22}} \end{bmatrix} \quad (45)$$

$$\sigma_{ij} = \frac{E}{1-\nu^2} \begin{bmatrix} \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & \nu \end{bmatrix}}_{C_{11}} & \underbrace{\begin{bmatrix} 0 & 1-\nu \\ 0 & 0 \end{bmatrix}}_{C_{12}} \\ \underbrace{\begin{bmatrix} 0 & 0 \\ 1-\nu & 0 \end{bmatrix}}_{C_{21}} & \underbrace{\begin{bmatrix} \nu & 0 \\ 0 & 1 \end{bmatrix}}_{C_{22}} \end{bmatrix} \begin{bmatrix} \begin{bmatrix} \varepsilon_{11} & \varepsilon_{12} \\ \varepsilon_{21} & \varepsilon_{22} \end{bmatrix} & \begin{bmatrix} \varepsilon_{11} & \varepsilon_{12} \\ \varepsilon_{21} & \varepsilon_{22} \end{bmatrix} \\ \begin{bmatrix} \varepsilon_{11} & \varepsilon_{12} \\ \varepsilon_{21} & \varepsilon_{22} \end{bmatrix} & \begin{bmatrix} \varepsilon_{11} & \varepsilon_{12} \\ \varepsilon_{21} & \varepsilon_{22} \end{bmatrix} \end{bmatrix} \quad (46)$$

$$\sigma_{ij} = C_{11}\varepsilon_{kl} + C_{12}\varepsilon_{kl} + C_{21}\varepsilon_{kl} + C_{22}\varepsilon_{kl} \quad (47)$$

Equation (47) is equivalent to equations (37)–(40).

The final step is to apply the equilibrium conditions. The repeated index j indicates that there will be a summation over j and the result of σ_{ij} , will be a vector, since the index j cancels out.

$$\sigma_{ij,j} + b_i = 0 \quad (48)$$

$$\frac{\partial \sigma_{ij}}{\partial x_j} + b_i = 0 \quad (49)$$

where b_i are the body forces acting in the Cartesian directions. The two above definitions are equivalent. Let us look at the two equations that result by expanding these equations:

$$\frac{\partial \sigma_{11}}{\partial x_1} + \frac{\partial \sigma_{12}}{\partial x_2} + b_1 = 0 \quad (50)$$

$$\frac{\partial \sigma_{21}}{\partial x_1} + \frac{\partial \sigma_{22}}{\partial x_2} + b_2 = 0 \quad (51)$$

These strain and stress derivatives illustrate how taking derivatives can either increase or decrease the order of the tensor result. Indicical notation makes it easy to keep track of the appropriate derivatives and also makes it clear what the final order of the tensor will be.

Note: For numerical solvers, such as the finite element method, the dimensionality of the strain, constitutive and stress tensors are often reduced to 1,2 and 1, respectively, in order to ease the software implementation. The results obtained are equivalent to the full method above.

5 Summary

Common name	Tensor Order	Example
Scalar	0	–
Vector	1	u_i
Matrix	2	$\varepsilon_{ij}, \sigma_{ij}$
–	4	C_{ijkl}

Rules of indicial notation:

1. Sum over all repeated indices
2. Increment all unique indices fully at least once, covering all combinations.
3. Increment repeated indices first
4. A comma indicates differentiation, with respect to coordinate x_i
5. The number of non-repeated indices indicates the order of the tensor result of a tensor multiplication or derivative