Tensor Notation (Advanced)

Introduction

This page addresses advanced aspects of tensor notation. A key strength of tensor notation is its ability to represent systems of equations with a single tensor equation. This makes it possible to recognize relationships among tensor terms, and manipulate them, that would otherwise be nearly impossible to do using matrix notation.

Kronecker Delta and Derivatives of Axis Variables

The Kronecker Delta is related to the derivatives of coordinate axis variables with respect to themselves. Start with an x, y, z coordinate system and ask, "What are $\frac{\partial x}{\partial x}$, $\frac{\partial x}{\partial y}$, and $\frac{\partial x}{\partial z}$?" The answers are simple: 1, 0, and 0. Also,

$$\frac{\partial y}{\partial x} = 0 \qquad \frac{\partial y}{\partial y} = 1 \qquad \frac{\partial y}{\partial z} = 0$$

And similarly, the derivatives of z with respect to the three variables are

$$\frac{\partial z}{\partial x} = 0 \qquad \frac{\partial z}{\partial y} = 0 \qquad \frac{\partial z}{\partial z} = 1$$

At this point, the pattern should be obvious. It can all be summarized in tensor notation as

$$\frac{\partial x_i}{\partial x_j} = x_{i,j} = \delta_{ij}$$

Example

Loading [MathJax]/extensions/MathMenu.js ronecker delta and derivatives of coordinates can be seen in the following example. First, recall that a derivative such as dy/dx

represents $y_{new} - y_{old}$ divided by $x_{new} - x_{old}$. Next start at some reference point, say (2, 5, 9). This is the "old" point. Now change the x coordinate by 1 from 2 to 3, the "new" value. So $x_{new} - x_{old} = 1$.

But nothing has changed the y value, so it is still equal to 5. (Note that the coordinates are independent, so y is not required to change just because x did.) So $y_{new} - y_{old} = 5 - 5 = 0$ and $dy/dx = 0$.

In contrast, $dx/dx = (3 - 2)/(3 - 2) = 1$. Clearly $\frac{dx_i}{dx_j} = x_{i,j} = \delta_{ij}$.

Kronecker Delta Multiplication

The Kronecker Delta is nicknamed the substitution operator because of the following simple property of multiplication, best explained by example. Multiplying v_i by δ_{ij} gives

$$v_i \delta_{ij} = v_1 \delta_{1j} + v_2 \delta_{2j} + v_3 \delta_{3j}$$

where the i subscript has been automatically summed from 1 to 3 because it appeared twice in the term, once in v_i and once in δ_{ij}.

Now arbitrarily select a value for j, say $j = 3$. This gives

$$v_i \delta_{i3} = v_1 \delta_{13} + v_2 \delta_{23} + v_3 \delta_{33} = v_3$$

The result is simply v_3 (the value chosen for j) because $\delta_{33} = 1$ while $\delta_{13} = \delta_{23} = 0$. In general, $v_i \delta_{ij}$ will always equal v_j for whatever value is selected for j. Therefore, the general rule is

$$v_i \delta_{ij} = v_j$$

hence the nickname of substitution operator.

The rule applies regardless of the complexity of the term. Recall the vector cross-product term $\epsilon_{ijk} \omega_j r_k$. Multiply it by δ_{im} for example.

$$\delta_{im} \epsilon_{ijk} \omega_j r_k \quad = \quad \epsilon_{mjk} \omega_j r_k$$

It has the sole effect of substituting m for i in the alternating tensor.

Scalar Equations

It is very important to be able to recognize the rank of any tensor term. Quantities such as $v_{i,i}$, $a_i b_i$, $A_{ij} B_{ij}$, and $\sigma_{ij} \epsilon_{ij}$ are all scalar terms in fact, which also means they would be part of a single scalar equation. For example, the strain energy density, W, of a

linear elastic material is a scalar quantity given by

$$W = \frac{1}{2}\sigma_{ij}\epsilon_{ij}$$

Although there are several subscripts in the equation, they all expand out (all sum from 1 to 3) because each letter occurs exactly twice. This is amplified by the fact that on the left hand side (LHS) of the equation, W is undeniably a scalar because it has no subscript(s) at all.

Vector Equations

A simple vector equation is the cross product of an angular rotation vector, ω_j, with a position vector, r_k, to obtain the velocity vector, v_i. (Yes, this is just $\mathbf{v} = \boldsymbol{\omega} \times \mathbf{r}$.)

$$v_i = \epsilon_{ijk}\omega_j r_k$$

This very small tensor equation represents so much. For starters, it represents three equations, not one, because the v_i term shows clearly that there is an equation for each $i = 1$, 2, and 3. Second, each of the three equations contains nine terms because j and k both sum to three. However, we've already seen here that ϵ_{ijk} is zero in most terms, leaving

$$v_1 = \omega_2 r_3 - \omega_3 r_2$$
$$v_2 = \omega_3 r_1 - \omega_1 r_3$$
$$v_3 = \omega_1 r_2 - \omega_2 r_1$$

That is a lot packed into one small tensor equation. Of course, $\mathbf{v} = \boldsymbol{\omega} \times \mathbf{r}$ actually represents the same equations as well, except that the subscripts in the tensor equation explicitly lead one to the three component equations.

It is important to recognize that tensor notation often provides the freedom to write any tensor term in many forms. In this case, alternative examples include

$$v_i = \omega_j \epsilon_{ijk} r_k \qquad v_i = r_k \epsilon_{ijk} \omega_j \qquad v_i = r_k \omega_j \epsilon_{ijk} \qquad v_k = \epsilon_{ijk}\omega_i r_j$$

All are equivalent, and technically correct, because the multiplication details are dictated by the subscripts, not the order of the factors. However, it is customary for readability considerations to write ϵ_{ijk} first in a term, and write the remaining factors in approximately the same order they would appear in vector or matrix notation. This leads to the original form of $v_i = \epsilon_{ijk}\omega_j r_k$ as being preferred.

It is also important to recognize forms that are incorrect, and why. One incorrect form is $v_m = \epsilon_{ijk}\omega_j r_k$. This is wrong because the free indices on the two sides of the equation are different: m on the LHS and i on the RHS. While one might guess the author's intent in such a case, such a mismatch of free indices should be avoided.

A similar bad form is the addition of $a_i + b_j$. This is simply meaningless.

A second incorrect form is $v_i = \epsilon_{ijj}\omega_j r_j$. This is wrong because j appears more than twice in the RHS term.

2nd Order Tensor Equations

A very common 2nd order tensor equation is Hooke's Law, relating stress to strain in linear elastic materials. It is written in matrix notation as

$$\boldsymbol{\epsilon} = \frac{1}{E}\left[(1+\nu)\boldsymbol{\sigma} - \nu\,\mathbf{I}\,\mathrm{tr}(\boldsymbol{\sigma})\right]$$

where: $\boldsymbol{\sigma}$ is the stress tensor,

$\boldsymbol{\epsilon}$ is the strain tensor,

E is the elastic modulus,

ν is Poisson's ratio,

$\mathbf{I}$ is the identity matrix,

$\mathrm{tr}(...)$ is the trace of a tensor, in this case, $\sigma_{11} + \sigma_{22} + \sigma_{33}$

Hooke's Law is written in tensor notation as

$$\epsilon_{ij} = \frac{1}{E}\left[(1+\nu)\sigma_{ij} - \nu\,\delta_{ij}\,\sigma_{kk}\right]$$

This shows clearly that each term is 2nd rank because of the i and j indices. Note that σ_{kk} is just a scalar quantity because k is summed from 1 to 3 since it appears twice. So σ_{kk} is the trace of $\boldsymbol{\sigma}$. Another key insight is that the entire $\nu\,\delta_{ij}\,\sigma_{kk}$ term is nonzero only when $i = j$. This corresponds to the normal stresses and strains and is the classic Poisson's Effect. On the other hand, the entire term is zero when $i \neq j$ because of the presence of δ_{ij}.

Writing all the matrices out gives

$$\begin{bmatrix} \epsilon_{11} & \epsilon_{12} & \epsilon_{13} \\ \epsilon_{21} & \epsilon_{22} & \epsilon_{23} \\ \epsilon_{31} & \epsilon_{32} & \epsilon_{33} \end{bmatrix} = \frac{1}{E}\left\{ (1+\nu) \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{bmatrix} - \nu\,(\sigma_{11}+\sigma_{22}+\sigma_{33}) \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right\}$$

So what are the individual component equations represented by the tensor notation equation, and how many are there? Since there are two free indices, there would normally be $3 * 3 = 9$ equations. But it turns out that all terms are symmetric in Hooke's Law, so there are only six independent equations. They are initially

$$\epsilon_{11} = \frac{1}{E}\left[(1+\nu)\sigma_{11} - \nu\,(\sigma_{11}+\sigma_{22}+\sigma_{33})\right] \qquad \epsilon_{12} = \frac{1}{E}\left[(1+\nu)\sigma_{12}\right]$$

$$\epsilon_{22} = \frac{1}{E} \left[(1+\nu)\sigma_{22} - \nu \left(\sigma_{11} + \sigma_{22} + \sigma_{33} \right) \right] \qquad \epsilon_{13} = \frac{1}{E} \left[(1+\nu)\sigma_{13} \right]$$

$$\epsilon_{33} = \frac{1}{E} \left[(1+\nu)\sigma_{33} - \nu \left(\sigma_{11} + \sigma_{22} + \sigma_{33} \right) \right] \qquad \epsilon_{23} = \frac{1}{E} \left[(1+\nu)\sigma_{23} \right]$$

but then clean up into more concise forms

$$\epsilon_{11} = \frac{1}{E} \left[\sigma_{11} - \nu \left(\sigma_{22} + \sigma_{33} \right) \right] \qquad \epsilon_{12} = \frac{(1+\nu)}{E}\sigma_{12}$$

$$\epsilon_{22} = \frac{1}{E} \left[\sigma_{22} - \nu \left(\sigma_{11} + \sigma_{33} \right) \right] \qquad \epsilon_{13} = \frac{(1+\nu)}{E}\sigma_{13}$$

$$\epsilon_{33} = \frac{1}{E} \left[\sigma_{33} - \nu \left(\sigma_{11} + \sigma_{22} \right) \right] \qquad \epsilon_{23} = \frac{(1+\nu)}{E}\sigma_{23}$$

Again, this is an amazing amount of information packed into a single tensor equation.

⚠ For anyone diving directly into this page here, note that ϵ_{12}, ϵ_{13}, and ϵ_{23} are one-half of the normal shear quantities, i.e., $\epsilon_{12} = \gamma_{12}/2$.

☞ Tensor Notation Algebra - Inverting Hooke's Law

So far, tensor notation has not actually provided any capabilities beyond matrix notation. After all, the matrix form of Hooke's Law does contain all the same information that is available in the tensor equation. However, the great power of tensor notation over matrix notation becomes evident when one starts to manipulate tensor equations. The example here will demonstrate how to invert Hooke's Law from strain-as-a-function-of-stress to stress-as-a-function-of-strain. Such an inversion is all but impossible using matrix notation.

Returning to the fundamental equation...

$$\epsilon_{ij} = \frac{1}{E} \left[(1+\nu)\sigma_{ij} - \nu \, \delta_{ij} \, \sigma_{kk} \right]$$

It is easy to solve for σ_{ij} to get

$$\sigma_{ij} = \frac{1}{(1+\nu)} \left[E \, \epsilon_{ij} + \nu \, \delta_{ij} \, \sigma_{kk} \right]$$

But there is a problem, a major problem. There is still a stress term, σ_{kk}, on the RHS that must be resolved. It cannot be combined with σ_{ij} on the LHS because they are different animals. σ_{ij} is an individual component of the stress tensor, σ_{kk} is the trace of the stress tensor, $(\sigma_{11} + \sigma_{22} + \sigma_{33})$.

This hurdle is overcome in a couple steps by first multiplying the equation through (both sides of course) by δ_{ij} as follows.

$$\delta_{ij}\sigma_{ij} = \frac{1}{(1+\nu)}\left[E\,\delta_{ij}\,\epsilon_{ij} + \nu\,\delta_{ij}\,\delta_{ij}\,\sigma_{kk}\right]$$

At this point, several advanced properties of tensor notation kick in. First, note that the entire equation has been transformed from a 2nd rank tensor equation to a scalar equation because there are no longer any free indices. Each term now has i and j occurring twice in it, so both are automatically summed from 1 to 3.

Second, the terms containing $\delta_{ij}\sigma_{ij}$ and $\delta_{ij}\epsilon_{ij}$ can be simplified by recalling the substitution property of the Kronecker Delta. Since $\delta_{ij} = 1$ only when $i = j$, then

$$\delta_{ij}\sigma_{ij} = \sigma_{ii} = \sigma_{jj} = \sigma_{11} + \sigma_{22} + \sigma_{33}$$

and

$$\delta_{ij}\epsilon_{ij} = \epsilon_{ii} = \epsilon_{jj} = \epsilon_{11} + \epsilon_{22} + \epsilon_{33}$$

So choose, arbitrarily, to write each term with j subscripts as σ_{jj} and ϵ_{jj}.

The final interesting term contains $\delta_{ij}\delta_{ij}$. This equals 3 because

$$\delta_{ij}\delta_{ij} = \delta_{ii} = \delta_{11} + \delta_{22} + \delta_{33} = 1 + 1 + 1 = 3$$

So the now-scalar equation reduces to

$$\sigma_{jj} = \frac{1}{(1+\nu)}\left[E\,\epsilon_{jj} + 3\,\nu\,\sigma_{kk}\right]$$

It's now time for another round of insight into tensor notation. Note that σ_{jj} and σ_{kk} are in fact equal because both expand to $\sigma_{11} + \sigma_{22} + \sigma_{33}$. So they can be combined. The σ_{jj} is simply rewritten as σ_{kk}, producing

$$\sigma_{kk} = \frac{1}{(1+\nu)}\left[E\,\epsilon_{jj} + 3\,\nu\,\sigma_{kk}\right]$$

Now the σ_{kk} terms on the LHS and RHS can be combined to obtain

$$\sigma_{kk} = \frac{E\,\epsilon_{jj}}{(1 - 2\nu)}$$

which is the simple clean equation we need to insert back into

$$\sigma_{ij} = \frac{1}{(1+\nu)} \left[E \, \epsilon_{ij} + \nu \, \delta_{ij} \, \sigma_{kk} \right]$$

to obtain

$$\sigma_{ij} = \frac{E}{(1+\nu)} \left[\epsilon_{ij} + \frac{\nu}{(1-2\nu)} \delta_{ij} \epsilon_{jj} \right]$$

Except this is incorrect because j now appears three times in the last term. But the solution is simple. Simply rewrite ϵ_{jj} as ϵ_{kk}. This is completely acceptable because both forms expand out to give the same trace of the strain tensor.

$$\sigma_{ij} = \frac{E}{(1+\nu)} \left[\epsilon_{ij} + \frac{\nu}{(1-2\nu)} \delta_{ij} \epsilon_{kk} \right]$$

And we are finally done. It wasn't easy, but by comparison it seems absolutely impossible (at least to me) to pull this off using matrix notation.

So what happened? The example is typical of equations that contain tensors occurring in two different forms. In this case, the two different forms were σ_{ij} and σ_{kk}. The first is a component of the stress tensor while the second is the trace of the same stress tensor. (Both involve the same stress tensor, but they are in fact different quantities.) When this occurs, the usual procedure is to multiply the equation through by δ_{ij}.

The second key aspect of the example was the demonstrated freedom to change indices when they occur twice. This is because any indices occurring twice are automatically expanded out so it does not matter what letter they are. This was demonstrated by changing σ_{jj} to σ_{kk} and ϵ_{kk} to ϵ_{jj} during the example.

Epsilon-Delta Identity

The alternating tensor and the Kronecker delta are related to each other through the following identity

$$\epsilon_{ijk}\epsilon_{lmn} = \begin{vmatrix} \delta_{il} & \delta_{im} & \delta_{in} \\ \delta_{jl} & \delta_{jm} & \delta_{jn} \\ \delta_{kl} & \delta_{km} & \delta_{kn} \end{vmatrix}$$

$$= \delta_{il}(\delta_{jm}\delta_{kn} - \delta_{jn}\delta_{km}) + \delta_{im}(\delta_{jn}\delta_{kl} - \delta_{jl}\delta_{kn}) + \delta_{in}(\delta_{jl}\delta_{km} - \delta_{jm}\delta_{kl})$$

which is too complex to be of much use to anyone! However, multiplying through by δ_{il} reduces the above equation to something much more manageable and useful. (Proving this yourself is an excellent homework problem.)

$$\epsilon_{ijk}\epsilon_{imn} = \delta_{jm}\delta_{kn} - \delta_{jn}\delta_{km}$$

This is a 4th rank tensor equation because there are four free indices, j, k, m, and n. The i index is repeated twice on the LHS, so it is summed from 1 to 3. So the equation could be expanded to give

$$\epsilon_{ijk}\epsilon_{imn} = \epsilon_{1jk}\epsilon_{1mn} + \epsilon_{2jk}\epsilon_{2mn} + \epsilon_{3jk}\epsilon_{3mn} = \delta_{jm}\delta_{kn} - \delta_{jn}\delta_{km}$$

The terms can be evaluated once values are chosen for each of the four free indices.

The identity is used when two alternating tensors are present in a term, which usually arises when the term involves cross products. The benefit of employing it is that once the epsilons are transformed into the deltas, then the substitution property of the Kronecker Deltas can be used to simplify the equation. The following example demonstrates the usefulness of this identity.

☞ **Epsilon-Delta Identity Example**

Recall from this discussion of cross products that the area of a triangle is given by

$$Area = \frac{1}{2}|\mathbf{a} \times \mathbf{b}|$$

which can be written in tensor notation as

$$Area = \frac{1}{2}\sqrt{\epsilon_{ijk}a_j b_k \epsilon_{imn}a_m b_n}$$

At this point, the epsilon-delta identity can be employed to obtain

$$Area = \frac{1}{2}\sqrt{(\delta_{jm}\delta_{kn} - \delta_{jn}\delta_{km})a_j b_k a_m b_n}$$

Multiplying through gives

$$Area = \frac{1}{2}\sqrt{\delta_{jm}\delta_{kn}a_j b_k a_m b_n - \delta_{jn}\delta_{km}a_j b_k a_m b_n}$$

And the substitution property of the delta operator can now be exploited to obtain

$$Area = \frac{1}{2}\sqrt{a_m a_m b_n b_n - b_m a_m a_n b_n}$$

But $a_m a_m$ is simply the dot product of the **a** vector with itself, as is $b_n b_n$. And $a_m b_m$ and $a_n b_n$ are both the same dot product of **a** with **b**.

$$Area = \frac{1}{2}\sqrt{(\mathbf{a}\cdot\mathbf{a})(\mathbf{b}\cdot\mathbf{b}) - (\mathbf{a}\cdot\mathbf{b})^2}$$

Which offers an alternative calculation to the cross product approach. For what it's worth, the equation can also be written as a determinant.

$$Area = \frac{1}{2}\sqrt{\begin{vmatrix} (\mathbf{a}\cdot\mathbf{a}) & (\mathbf{a}\cdot\mathbf{b}) \\ (\mathbf{a}\cdot\mathbf{b}) & (\mathbf{b}\cdot\mathbf{b}) \end{vmatrix}}$$

And here is a second example...

Second Epsilon-Delta Identity Example

Prove the following identity, where **v** is a vector.

$$\nabla \times (\nabla \times \mathbf{v}) = \nabla(\nabla \cdot \mathbf{v}) - \nabla^2 \mathbf{v}$$

$(\nabla \times \mathbf{v})$ is $\epsilon_{ijk} v_{k,j}$ and $\nabla \times (\nabla \times \mathbf{v})$ is $\epsilon_{mni}\epsilon_{ijk} v_{k,jn}$. Apply the identity to transform the equation to

$$\epsilon_{mni}\epsilon_{ijk} v_{k,jn} = (\delta_{mj}\delta_{nk} - \delta_{mk}\delta_{jn}) v_{k,jn}$$

Multiply through to get

$$\epsilon_{mni}\,\epsilon_{ijk} v_{k,jn} = \delta_{mj}\delta_{nk} v_{k,jn} - \delta_{mk}\delta_{jn} v_{k,jn}$$

As before, apply the substitution properties to obtain

$$\epsilon_{mni}\epsilon_{ijk} v_{k,jn} = v_{k,mk} - v_{m,nn}$$

Recognize that $v_{k,mk}$ is the gradient of the divergence of **v**. This can be made clearer if the term is written more explicitly as $(v_{k,k})_{,j}$. So $v_{k,jk} = \nabla(\nabla \cdot \mathbf{v})$.

The second term is just the Laplacian of **v**, i.e., $\nabla^2 \mathbf{v}$. This proves

$$\nabla \times (\nabla \times \mathbf{v}) = \nabla(\nabla \cdot \mathbf{v}) - \nabla^2 \mathbf{v}$$