EOS 352 Continuum Dynamics
Subscript notation

April 30, 2014

Overview

These notes cover the following

- Subscript notation as an alternative to standard vector notation
- Writing equations for generic vector components
- Dot products, repeated indices and the summation convention
- Algebra and differentiation rules in subscript notation

Vectors as defined by components

In vector algebra, we have so far used boldface letters to denote vectors and defined a number of vector specific operations such as dot and cross products to write down equations that define relationships between different vectors. The simplest example would be for instance to say that one vector, say $\mathbf{a}$, is the sum of two other vectors, $\mathbf{b}$ and $\mathbf{c}$:

$$\mathbf{a} = \mathbf{b} + \mathbf{c} \tag{1}$$

A more sophisticated example would be the following: if the normal to a surface is $\mathbf{n}$, then we know the normal component of a flux field $\mathbf{q}$ would be $\mathbf{q} \cdot \mathbf{n}$. If we wanted to give this component a direction, we would write it $(\mathbf{q} \cdot \mathbf{n})\mathbf{n}$, as its direction is clearly $\mathbf{n}$. Now, if we wanted to say 'what is the part of the flux field that causes transport along rather than across the surface', we would probably say that it's whatever is left over after subtracting the normal part,

$$\mathbf{q}_{\parallel} = \mathbf{q} - (\mathbf{q} \cdot \mathbf{n})\mathbf{n}. \tag{2}$$

We have used a similarly symbolic notation in vector calculus, writing for instance

$$\mathbf{q} = -k\nabla T \tag{3}$$

to write the relationship between heat flux $\mathbf{q}$ and the temperature field T.

This notation has the advantage of brevity, but in any practical application, what we really need to do is translate this shorthand notation into individual components. We might have $\mathbf{a} = (a_x, a_y, a_z)$ (or equally, $\mathbf{a} = a_x\mathbf{i} + a_y\mathbf{j} + a_z\mathbf{k}$). If similarly $\mathbf{b} = (b_x, b_y, b_z)$, and $\mathbf{c} = (c_x, x_y, x_z)$, then (1) stands for

$$a_x = b_x + c_x \tag{4a}$$
$$a_y = b_y + c_y \tag{4b}$$
$$a_z = b_z + c_z \tag{4c}$$

These are simple but tedious to write out, even for such a trivial example. For (2), things get even more tedious: if $\mathbf{q} = (q_x, q_y, q_z)$ and $\mathbf{n} = (n_x, n_y, n_z)$, then $\mathbf{q} \cdot \mathbf{n} = q_x n_x + q_y n_y + q_z n_z$, and $\mathbf{q}_\parallel = (q_{\parallel,x}, q_{\parallel,y}, q_{\parallel,z})$ is given in component form by

$$q_{\parallel,x} = q_x - (q_x n_x + q_y n_y + q_z n_z)n_x, \tag{5a}$$
$$q_{\parallel,y} = q_y - (q_x n_x + q_y n_y + q_z n_z)n_y, \tag{5b}$$
$$q_{\parallel,z} = q_z - (q_x n_x + q_y n_y + q_z n_z)n_z. \tag{5c}$$

Not only do you need to remember what operations like dot products mean in terms of components, writing this out in components quickly churns out lengthy expressions. We can also do the same with vector calculus-based relations like Fourier's law (3), which becomes

$$q_x = -k\frac{\partial T}{\partial x}, \qquad q_y = -k\frac{\partial T}{\partial y}, \qquad q_z = -k\frac{\partial T}{\partial z} \tag{6}$$

The point is that, in order to fully specify a vector, we have to specify all its components. This is what is really meant by saying that 'a vector has a length and a direction'. A relationship between different vectors and scalars is therefore really a set of equations that links all their components. This is what standard vector notation does, but it hides the individual components in boldface symbols like $\mathbf{q}$ and $\mathbf{n}$ and in vector operations like $\mathbf{q} \cdot \mathbf{n}$ and ∇T.

Subscript notation offers a more explicit alternative to this that becomes extremely useful later — for instance, when we try to write down conservation laws for vector-valued quantities like momentum. Rather than saying that a vector is denoted by a single symbol $\mathbf{q}$, subscript notation tries to identify its individual components. Ordinarily, we would expect this to mean writing the vector out as (q_x, q_y, q_z) and then writing out equations for each component. However, as we've seen above, writing out equations for components is tedious, and we need a shorthand.

The first thing to recognize when trying to write out equations for components is that there is nothing special about the x-, y- and z-directions. We can equally label them the first, second and third axes, and treat them as basically being the same. Instead of of writing the coordinates of a point as (x, y, z), an alternative notation is to write (x_1, x_2, x_3) (so $x = x_1$, $y = x_2$, $z = x_3$). Similarly, instead of writing a vector $\mathbf{a}$ in component form as (a_x, a_y, a_z), we can denote the components by (a_1, a_2, a_3). We can then say that we know the vector $\mathbf{a}$ if we know a_i for each $i = 1, 2, 3$. Likewise, to specify the position of a point, we need to know the corresponding x_i for each $i = 1, 2, 3$

What do we gain by this change in notation? By identifying components by a numerical subscript (i.e. a_1 rather than a_x), we can write down equations for a generic (rather than specific) subscript. For instance, a naive take on (4) would be that it now reads

$$a_1 = b_1 + c_1, \tag{7a}$$
$$a_2 = b_2 + c_2, \tag{7b}$$
$$a_3 = b_3 + c_3. \tag{7c}$$

But it should be clear that this is just a lot of repetition, replacing the '1' in the first equation by '2' and '3' to obtain the second and third equations. A more succinct way of saying the same thing would be that

$$a_i = b_i + c_i \tag{8}$$

for $i = 1, 2, 3$. So we need to write down only a single equation, rather than one separately for each component. This is the basic idea of subscript notation: to write down equations for generic vector components. For instance, instead of $\mathbf{a} = \mathbf{b} + \mathbf{c}$, one would simply write that $a_i = b_i + c_i$, on the understanding that this must hold for $i = 1, 2, 3$.

While it is common to use the subscript i to denote the index of a generic vector component, one could use any other letter also, so long as it is used consistently. For instance, (8) could also be written as

$$a_j = b_j + c_j$$

for $j = 1, 2, 3$.

Dot products and the summation convention

Simple summation of vectors is relatively easy to understand in component notation. The real power of component notation lies in its ability to render much more tedious equations like (5) in succinct form. Recall that we have

$$\mathbf{q} \cdot \mathbf{n} = q_x n_x + q_y n_y + q_z n_z \tag{9}$$

Switching to numerical subscript notation, we have to write $\mathbf{q} = (q_1, q_2, q_3)$ instead of (q_x, q_y, q_z), and similarly $\mathbf{n} = (n_1, n_2, n_3)$ instead of (n_x, n_y, n_z). In that case,

$$\mathbf{q} \cdot \mathbf{n} = q_1 n_1 + q_2 n_2 + q_3 n_3 \tag{10}$$

But this can be abbreviated succinctly as

$$\mathbf{q} \cdot \mathbf{n} = \sum_{i=1}^{3} q_i n_i. \tag{11}$$

Exercise 1 *In subscript notation, Fourier's law (3) is*

$$q_i = -k \frac{\partial T}{\partial x_i}$$

which, with $i = 1$, 2, 3, *reproduces (6) when we recognize that* $q_x = q_1$, $q_y = q_2$, $q_z = q_3$ *and* $x = x_1$, $y = x_2$, $z = x_3$. *Consider now the equation*

$$\nabla \cdot \mathbf{q} = a.$$

How would you write this out as succinctly as possible in subscript notation?

Suppose we would like to write (2) in component form. Using (11), we might be tempted to write

$$q_{\parallel,i} = q_i - (\mathbf{q} \cdot \mathbf{n}) n_i = q_i - \left(\sum_{i=1}^{3} q_i n_i \right) n_i. \tag{12}$$

But now we have a problem: Suppose I want to know the first component $q_{\parallel,1}$, so I'd like to put $i = 1$. But then, on the right-hand side, it looks like I should sum over i going from 1 to 3. So should I put $i = 1$ everywhere or sum over i? Clearly, comparing with (5), the answer should be $q_{\parallel,1} = q_1 - (q_1 n_1 + q_2 n_2 + q_3 n_3) n_1 = q_1 - (\sum_{i=1}^{3} q_i n_i) n_1$, so the sum only applies to the indices i inside the bracket. But this notation is highly ambiguous (and actually plain wrong): we cannot simultaneously have a quantity fixed ($i = 1$) and sum over the same quantity going from 1 to 3.

The correct way to go about this is to realize that, in (11), the index i can be changed to any other index and the resulting sum will always be the same. In other words, we can put

$$\sum_{i=1}^{3} q_i n_i = \sum_{j=1}^{3} q_j n_j = \sum_{k=1}^{3} q_k n_k, \tag{13}$$

because all three expressions stand for the sum $q_1 n_1 + q_2 n_2 + q_3 n_3$. To make (12) unambiguous, a better notation would therefore be

$$q_{\parallel,i} = q_i - \left(\sum_{j=1}^{3} q_j n_j \right) n_i, \tag{14}$$

so i is reserved for labelling the generic component $q_{\|,i}$ that we are defining, while j labels the components in the sum, and there is no ambiguity about what needs to be summed. The index j is then called a *dummy index*.[1] It can be changed to any other index apart from i, and the equation retains its meaning. In other words, we could equally write

$$q_{\|,i} = q_i - \left(\sum_{k=1}^{3} q_k n_k \right) n_i. \tag{15}$$

or even

$$q_{\|,j} = q_j - \left(\sum_{k=1}^{3} q_k n_k \right) n_j.$$

Sums over products involving repeated indices are extremely common when using subscript notation: basically, they occur whenever there would have been a dot product in classical vector notation. To shorten subscript notation, the following convention is therefore adopted almost universally: if a subscript index is repeated in a product or derivative, a sum over that index from 1 to 3 is implied. A subscript index must not be repeated more than once in the same product or derivative.

In other words, if we write $a_j b_j$, what is meant by this *by definition* is $a_j b_j \equiv \sum_{j=1}^{3} a_j b_j$, but the summation sign will be omitted. Similarly, $a_i b_i$ or $a_k b_k$ mean the same thing, and we can legitimately write $a_j b_j = a_k b_k = \sum_{l=1}^{3} a_l b_l$.

This is the first part of the summation convention. The second part is that we cannot have expressions like $a_i b_i c_i$: the same index (here i) can only be repeated once in the same product of terms. This is because dot products only involve two factors, and expressions like $a_i b_i c_i$ are mostly likely to be incorrect attempts to write out a product of a vector with a dot product as in (12), which might have suggested that the last term on the right-hand side should be written as $q_i n_i n_i$ — but as we saw above, this does not answer the question of which index was fixed and which was to be summed over.

Example 1 *With the summation convention, (2) can be written in subscript notation as*

$$q_{\|,i} = q_i - q_j n_j n_i.$$

Compare this for efficiency in notation with the lengthier expressions in (5).

Note 1 *An index must not be repeated more than once when the product is formed — this still allows an index to appear more than twice if there is an operation that must be performed before the product is taken. For instance, it is legitimate to write*

$$a_i(b_i + c_i)$$

[1] This is the same as the integration variable in definite integration being a *dummy variable*. Take an integral like $\int_0^\pi \sin(x)\,\mathrm{d}x$. This evaluates to 2. There is no reason however why we need to have x as the integration variable here. We could equally write $\int_0^\pi \sin(y)\,\mathrm{d}y$, which also evaluates to 2. In fact, quite generally we have $\int_a^b f(x)\,\mathrm{d}x = \int_a^b f(y)\,\mathrm{d}y$ if a and b are fixed.

because the term in brackets must be evaluated first: we really have $a_i(b_i + c_i) = a_i d_i$, where $d_i = b_i + c_i$, and hence the index i is only repeated once in the actual product. This expression in fact stands for $\mathbf{a} \cdot (\mathbf{b} + \mathbf{c})$ in standard vector notation.

Now, the definition of the summation convention allows for derivatives as well as products. This is because divergences lead to sums over repeated indices, just as dot products do. For instance, we have

$$\nabla \cdot \mathbf{q} = \frac{\partial q_x}{\partial x} + \frac{\partial q_y}{\partial y} + \frac{\partial q_z}{\partial z}.$$

In numerical subscript notation, this becomes

$$\nabla \cdot \mathbf{q} = \frac{\partial q_1}{\partial x_1} + \frac{\partial q_2}{\partial x_2} + \frac{\partial q_3}{\partial x_3} = \sum_{i=1}^{3} \frac{\partial q_i}{\partial x_i}.$$

Using the summation convention, this can be written more succinctly as

$$\nabla \cdot \mathbf{q} = \frac{\partial q_i}{\partial x_i},$$

the summation over i being implied.

Exercise 2 *Write the following in subscript notation, using the summation convention throughout where applicable:*

1. $\mathbf{a} = \lambda \mathbf{b} - \mathbf{c}$,
2. $\mathbf{q} \cdot \mathbf{n}$ *if* $\mathbf{q} = -k\nabla T$,
3. $\rho c \frac{\partial T}{\partial t} + \rho c \mathbf{u} \cdot \nabla T - \nabla \cdot (k \nabla T) = a.$
4. $\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0$
5. $\nabla^2 T = a.$
6. $\nabla \cdot (\mathbf{q} + \mathbf{p})$
7. $|\mathbf{a}|^2$

Exercise 3 *Find a_1, a_2 and a_3 if*

1. $a_i = b_i + c_i$ *with* $\mathbf{b} = (1, 2, 3)$, $\mathbf{c} = (4, -5, 6)$
2. $a_i = b_i + (b_j c_j) c_i$ *with* $\mathbf{b} = (2, -1, 4)$, $\mathbf{c} = (1, 3, 4)$
3. $a_i = -\frac{\partial f}{\partial x_i}$ *with* $f = x_1^2 + 2x_1 x_2 + x_2^2 - x_3^2$

4. $a_i = u_j \frac{\partial u_i}{\partial x_j}$ *with* $\mathbf{u} = (x_2, -x_1, x_3^2)$

5. $a_i = u_j \frac{\partial u_i}{\partial x_j} - u_j \frac{\partial u_j}{\partial x_i}$ *with* $\mathbf{u} = (x_1 + x_2, -x_1 + x_2, x_3)$

6. $a_i = u_i \frac{\partial u_j}{\partial x_j}$ *with* $\mathbf{u} = (x_1, x_2, x_3)$

Exercise 4 *Write in standard (non-component) vector notation*

1. $\int_V \frac{\partial q_i}{\partial x_i} \,\mathrm{d}V = \int_S q_i n_i \,\mathrm{d}S$

2. $u_i \frac{\partial u_j}{\partial x_j}$

3. $\frac{\partial (u_j u_j)}{\partial x_i}$

4. $\frac{1}{2} u_j \frac{\partial u_j}{\partial x_i}$

5. $\frac{\partial h}{\partial t} + \frac{\partial (h u_i)}{\partial x_i} + \frac{\partial q_i}{\partial x_i} = a$

Algebra and calculus rules

The usual rules of algebra and calculus continue to hold when subscript notation is used together with the summation convention. For instance, multiplication is distributive:

$$a_i(b_i + c_i) = a_i b_i + a_i c_i.$$

How can we show this? Recall that the repeated index i indicates a summation, so

$$a_i(b_i + c_i) = \sum_{i=1}^{3} a_i(b_i + c_i)$$

But for each term in the sum, the product is distributive. In other words, for $i = 1$, 2 or 3 fixed, $a_i(b_i + c_i) = a_i b_i + a_i c_i$ holds, Hence[2]

$$\sum_{i=1}^{3} a_i(b_i + c_i) = \sum_{i=1}^{3} (a_i b_i + a_i c_i) = \sum_{i=1}^{3} a_i b_i + \sum_{i=1}^{3} a_i c_i.$$

[2] In case the last step is not obvious, write the sums out explicitly:

$$\begin{aligned}
\sum_{i=1}^{3} (a_i b_i + a_i c_i) &= (a_1 b_1 + a_1 c_1) + (a_2 b_2 + a_2 c_2) + (a_3 b_3 + a_3 c_3) \\
&= a_1 b_1 + a_2 b_2 + a_3 b_3 + a_1 c_1 + a_2 c_2 + a_3 c_3 \\
&= \sum_{i=1}^{3} a_i b_i + \sum_{i=1}^{3} a_i c_i.
\end{aligned}$$

But we can then drop the summation signs again if we revert to using the summation condition, or

$$\sum_{i=1}^{3} a_i b_i + \sum_{i=1}^{3} a_i c_i = a_i b_i + a_i c_i.$$

In a similar vein, we can apply the usual rules of calculus. For instance, differentiation is distributive over sums,

$$\frac{\partial (f_i + g_i)}{\partial x_j} = \frac{\partial f_i}{\partial x_j} + \frac{\partial g_i}{\partial x_j},$$

and the product rule for sums over repeated indices can be written as

$$\frac{\partial (f_i g_i)}{\partial x_j} = \frac{\partial f_i}{\partial x_j} g_i + f_i \frac{\partial g_i}{\partial x_j} \tag{16}$$

Exercise 5 *The following are examples of the product rule:*

1. *Show that (16) holds by writing out the terms involved in the sum over i*
2. *Show similarly that*

$$\frac{\partial (f_i g_j)}{\partial x_i} = \frac{\partial f_i}{\partial x_i} g_j + f_i \frac{\partial g_j}{\partial x_i}.$$

Clearly, this denotes a vector field (with each component of the vector field being described by index j, as i is summed over). Would it be easy to write this identity in standard (non-subscript) vector notation?