Mathematical Preliminaries
Vectors, Tensors and all that

Jeevanjyoti Chakraborty
jeevan@mech.iitkgp.ernet.in

1 Motivation for tensors

In fluid mechanics (also in solid mechanics, and continuum mechanics, in general) we often refer to something called tensors. You may have already come across a statement like "stress is a tensor". In what follows, we try to develop some idea about tensors. In later chapters, we will show its connection with stress in fluid mechanics.

Consider functions ...

For a given scalar quantity, s, let us define a function as $f(s) = s^2$ (might as well be s^3 or $\sqrt{2}s^4$ or $\sin(s)$ and so on). The important thing to note here is that the input to the function f is a scalar and its output is also a scalar. We say that the function f maps a scalar to a scalar.

For a given vector, $\vec{r}$, let us define a function $f(\vec{r}) = |\vec{r}| = \sqrt{r_1^2 + r_2^2 + r_3^2}$, where r_1, r_2, and r_3 are the components of $\vec{r}$ referred to some chosen coordinate axes. Here, even though the input to f is a vector, the output is a scalar. So, we say that the function f maps a vector to a scalar.

Again, for a given vector, $\vec{r}$, let us define a function $f(\vec{r}) = \vec{a} \cdot \vec{r}$ where $\vec{a}$ is some vector intrinsic to the definition of f. Again, the function f maps a vector to a scalar.

Now, consider the definition $f(\vec{r}) = \alpha\vec{r}$, where α is some scalar intrinsic to the definition of f. This time, the input to f is a vector and the output is also a vector. So, f maps a vector to a vector. The mapping is through the scalar α.

Similarly, consider the function definition

$$f(\vec{r}) = \begin{bmatrix} \alpha_1 r_1 \\ \alpha_2 r_2 \\ \alpha_3 r_3 \end{bmatrix},$$

which is a vector with the three components $\alpha_1 r_1$, $\alpha_2 r_2$, and $\alpha_3 r_3$. The only difference from the previous definition is that while there was the same scalar α being multiplied with all the components of $\vec{r}$, here there is a different scalar being multiplied with the three components of $\vec{r}$. The spirit is the same. So, again the function f maps a vector to another vector. The mapping is through the three scalars α_1, α_2, and α_3.

Finally, we can have another type of function defintion which maps a vector to another vector in such a way that each of the components of the output vector is made up by using a linear combination of the components of the input vector. Consider the following definition

$$f(\vec{r}) = \begin{bmatrix} \alpha_{11} r_1 + \alpha_{12} r_2 + \alpha_{13} r_3 \\ \alpha_{21} r_1 + \alpha_{22} r_2 + \alpha_{23} r_3 \\ \alpha_{31} r_1 + \alpha_{32} r_2 + \alpha_{33} r_3 \end{bmatrix}.$$

Here, the input is a vector and the output is a vector also just like before. Let us call the output vector $\vec{v}$. The above definition can equivalently be written as

$$\begin{aligned}\vec{v} = f(\vec{r}) &= \begin{bmatrix}\alpha_{11} & \alpha_{12} & \alpha_{13}\\ \alpha_{21} & \alpha_{22} & \alpha_{23}\\ \alpha_{31} & \alpha_{32} & \alpha_{33}\end{bmatrix}\begin{bmatrix}r_1\\ r_2\\ r_3\end{bmatrix}\\ &= [\alpha][\vec{r}],\end{aligned}$$

where $[\alpha]$ and $[\vec{r}]$ are shorthand matrix representations. Instead of this matrix representation, we can simply write the definition for $\vec{v}$ as

$$\vec{v} = f(\vec{r}) = \boldsymbol{\alpha} \cdot \vec{r}.$$

We say that $f(\vec{r})$ maps a vector $\vec{r}$ to another vector $\vec{v}$. The mapping itself is through the entity $\boldsymbol{\alpha}$ whose matrix representation comprises nine scalars that are "packed" in a special way in three groups of three. We say that the special entity $\boldsymbol{\alpha}$ is a general linear mapping from $\vec{r}$ to $\vec{v}$. This general linear mapping from one vector to another vector is what referred to as a 2nd-order tensor.

It is extremely important to distinguish between the tensor itself and its matrix representation. Tensor is the word used to refer to the linear mapping. The matrix representation is just a way of denoting it. Very importantly, the matrix representation of a tensor is not unique. What do we really mean by "not unique"? To answer this question, we look back at vectors.

2 Matrix representation is not unique

A vector is an entity that has both magnitude and direction (and, very strictly speaking, also follows the rules of vector addition). The point to note is that the existence of this entity does not depend on its mathematical representation. It would be there even if we did not have any mathematical machinery to represent it. However, in order to talk about vectors and work with them we do need to represent them mathematically. So we set up a set of coordinate axes and denote the vector through certain coordinates or through functions of certain coordinates. Corresponding to a particular choice of the coordinate axes, the vector representation is unique. However, if we choose another set of coordinate axes, the vector representation changes. Nevertheless, the actual entity, i.e. the vector itself is, of course, the same regardless of the choice of the coordinate axes.

Now again consider the definition $\vec{v} = \boldsymbol{\alpha} \cdot \vec{r}$. Here $\vec{r}$ and $\vec{v}$ are both vectors. Suppose their matrix representations corresponding to one particular set of coordinate axes, say, $Ox_1x_2x_3$ are $[r_1\ r_2\ r_3]^T$ and $[v_1\ v_2\ v_3]^T$, respectively. Then corresponding to these representations, the matrix representation of the tensor $\boldsymbol{\alpha}$ will be through a unique set of nine scalars. Now, if we choose another set of coordinate axes, say, $O'x_1'x_2'x_3'$, then the matrix representation of $\vec{r}$ and $\vec{v}$ will also change to something different, say, $[r_1'\ r_2'\ r_3']^T$ and $[v_1'\ v_2'\ v_3']^T$, respectively, Simultaneously, the matrix representation of $\boldsymbol{\alpha}$ will also change to something different comprising another unique set of nine scalars. If we keep on choosing different coordinate axes, the matrix representations of $\vec{r}$ and $\vec{v}$ will keep changing with simultaneous changes in the matrix representation of $\boldsymbol{\alpha}$. But the relation $\vec{v} = \boldsymbol{\alpha} \cdot \vec{r}$ must necessarily continue to hold regardless of the choice of the coordinate axes.

3 Index notation and its relation to matrix representation

- The symbol v_i denotes the three components v_1, v_2, and v_3 of a vector $\vec{v}$. We have

$$\vec{v} = v_i \equiv \begin{bmatrix}v_1\\ v_2\\ v_3\end{bmatrix}.$$

Note that we have *chosen* a column matrix representation. We could easily have chosen a row matrix representation. But, once, we choose one of these, we have to be consistent throughout because it has implications in other representation and calculations.

- The symbol T_{ij} denotes the nine components T_{11}, T_{12}, T_{13}, ..., T_{33} of a 2nd-order tensor $\mathbf{T}$. We have

$$\mathbf{T} \equiv T_{ij} = \begin{bmatrix} T_{11} & T_{12} & T_{13} \\ T_{21} & T_{22} & T_{23} \\ T_{31} & T_{32} & T_{33} \end{bmatrix}.$$

  Why is the matrix representation as above and not like the following

$$\begin{bmatrix} T_{11} & T_{21} & T_{31} \\ T_{12} & T_{22} & T_{32} \\ T_{13} & T_{23} & T_{33} \end{bmatrix}.$$

  We will answer this question a little later.

- An index that is repeated (appearing twice) means summation. For instance, the dot product between two vectors is

$$\vec{a} \cdot \vec{b} \equiv a_i b_i = a_1 b_1 + a_2 b_2 + a_3 b_3 = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}^T \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

- An index that appears more than two times makes the whole thing meaningless! For instance, $a_i b_i c_i$ does not mean anything; it's just wrong.

- If an index is not repeated in a product, it is referred to as a "free" index. The number of free indices determines which order tensor the resulting entity is (remember scalars are tensors of order zero, vectors are tensors of order one). For instance, in the product $T_{ij} r_j$ the index j is repeated while the index i is free. Since one index is free, the product is a tensor of order one, i.e. a vector.

$$\begin{aligned} T_{ij} r_j &= T_{i1} r_1 + T_{i2} r_2 + T_{i3} r_3 \\ &\equiv \begin{bmatrix} T_{11} r_1 + T_{12} r_2 + T_{13} r_3 \\ T_{21} r_1 + T_{22} r_2 + T_{23} r_3 \\ T_{31} r_1 + T_{32} r_2 + T_{33} r_3 \end{bmatrix}, \end{aligned}$$

  which is indeed a vector.

- We have seen that $a_i b_i$ is a scalar. But what about $a_i b_j$? Here there are two indices, neither of which is repeated. So we have two free indices. Accordingly, the product $a_i b_j$ must represent a tensor of order two that can be denoted as T_{ij}. For the equivalent matrix form we have

$$T_{ij} = a_i b_j \equiv \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}^T = \begin{bmatrix} a_1 b_1 & a_1 b_2 & a_1 b_3 \\ a_2 b_1 & a_2 b_2 & a_2 b_3 \\ a_3 b_1 & a_3 b_2 & a_3 b_3 \end{bmatrix}$$

  Note how the resulting matrix representation has shaped up. It is not

$$\begin{bmatrix} a_1 b_1 & a_2 b_1 & a_3 b_1 \\ a_1 b_2 & a_2 b_2 & a_3 b_2 \\ a_1 b_3 & a_2 b_3 & a_3 b_3 \end{bmatrix}.$$

  This explains why

$$T_{ij} \equiv \begin{bmatrix} T_{11} & T_{12} & T_{13} \\ T_{21} & T_{22} & T_{23} \\ T_{31} & T_{32} & T_{33} \end{bmatrix} \quad \text{and not} \quad \begin{bmatrix} T_{11} & T_{21} & T_{31} \\ T_{12} & T_{22} & T_{32} \\ T_{13} & T_{23} & T_{33} \end{bmatrix}$$

- We have seen that $T_{ij}r_j$ is a vector. What about $T_{ji}r_j$? Again there is only one free index and so $T_{ji}r_j$ must be a tensor of order one, i.e. a vector. Let's expand it and find the equivalent matrix form

$$\begin{aligned} T_{ji}r_j &= T_{1i}r_1 + T_{2i}r_2 + T_{3i}r_3 \\ &\equiv \begin{bmatrix} T_{11}r_1 + T_{21}r_2 + T_{31}r_3 \\ T_{12}r_1 + T_{22}r_2 + T_{32}r_3 \\ T_{13}r_1 + T_{23}r_2 + T_{33}r_3, \end{bmatrix} \end{aligned}$$

which is a vector indeed. But this vector is definitely different from the one that we obtained from $T_{ij}r_j$.

- Note that if $[\mathbf{T}]$ and $[\vec{r}]$ are the matrix representations of the 2nd-order tensor $\mathbf{T}$ and the vector $\vec{r}$ respectively, then

$$T_{ij}r_j = [\mathbf{T}][\vec{r}] \qquad \text{whereas} \qquad T_{ji}r_j = [\mathbf{T}]^T[\vec{r}].$$

- The gradient operator is like a vector. So it can be represented using one free index. Referred to a particular coordinate system, say, $Ox_1x_2x_3$, we have

$$\nabla \equiv \frac{\partial}{\partial x_i} \equiv \begin{bmatrix} \frac{\partial}{\partial x_1} & \frac{\partial}{\partial x_2} & \frac{\partial}{\partial x_3} \end{bmatrix}^T. \quad \text{(Note the transpose making it a column matrix)}$$

- The divergence of a vector is like the dot product of two vectors. The result must be a scalar. We have

$$\nabla \cdot \vec{v} \equiv \frac{\partial v_i}{\partial x_i} \equiv \begin{bmatrix} \partial/\partial x_1 \\ \partial/\partial x_2 \\ \partial/\partial x_3 \end{bmatrix}^T \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \frac{\partial v_1}{\partial x_1} + \frac{\partial v_2}{\partial x_2} + \frac{\partial v_3}{\partial x_3}.$$

- A shorthand notation for representing partial differentiation is through the use of commas. Thus, for instance

$$\frac{\partial v_1}{\partial x_2} \equiv v_{1,2} \qquad \text{and} \qquad \frac{\partial v_i}{\partial x_i} \equiv v_{i,i}$$

- What about gradient of a vector, $\nabla\vec{v}$? This in indical notation is $\partial v_i/\partial x_j$ and is like the product a_ib_j. There being two free indices, it is a tensor of order two so that it can be represented in the form of a 3×3 matrix.

$$\nabla\vec{v} \equiv \frac{\partial v_i}{\partial x_j} \equiv v_{i,j} \equiv \begin{bmatrix} v_{1,1} & v_{1,2} & v_{1,3} \\ v_{2,1} & v_{2,2} & v_{2,3} \\ v_{3,1} & v_{3,2} & v_{3,3} \end{bmatrix}$$

- The Kronecker delta δ_{ij} is a special 2nd-order tensor that has the property that the three components where $i = j$ are equal to 1; the rest six where $i \neq j$ are 0.
- The Kronecker delta is very useful to substitute indices. Thus, for instance

$$\delta_{ij}a_i = a_j \qquad \text{and} \qquad \delta_{jk}T_{ik} = T_{ij}$$