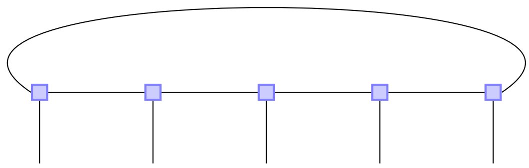
Tensor Calculus

Notes

Based on: many different sources



Sagnik Seth

Dept of Physical Sciences
IISER Kolkata

Contents

1 Introduction:

Hehe......😅 I am writing this as a way to understand tensors better and also to do something "productive". I don't know how much I will be able to complete but I intend to touch upon the basic aspects, albeit, in an extremely non-rigorous way....not going into heavy math (which I think is very bad 'coz math is great 😉). I will try to do some proofs (which I feel like doing) and skip others (since I don't care). I will definitely take $c = 1$ unless its necessary not to. I will use the convention $(+, -, -, -)$ for the metric tensor. I will also use some gen alpha slangs (which will indicate how chill I am!) and overall try to write in a fun way. A few references which I will use are listed below:

- *Tensor Calculus for Physics: A Concise Guide* by Dwight E Neuenschwander
- *The Poor Mans Introduction to Tensors* by Justin C. Feng
- Tensor Calculus YouTube video series by eigenchris and Andrew Dotson

I will continue to add the resources as I progress. I am writing this as my own personal notes and if it helps anyone else, I will be super happy. If any mistake is there, let me know! ✌️

2 Indices: the ultimate rizzler!

Indices make our lives easier when writing abstract quantitites having multiple components, like vectors. If we have a three-dimensional vector, we can write it as v^i where i can take the values 1, 2, or 3.

Why are the indices written as superscript? Well, these are contravariant indices which will be discussed later. For now, let just say that 'upstairs' indices are the 'normal thing'. Index placement is important and these are not powers...just the way we denote the components.

Consider the (in)famous equation:

$$\mathbf{F} = m\mathbf{a}$$

This can be written as $F^i = ma^i$, for each component i. Just remember that we should have the same kind of indices on both side of the equation finally. That is, if we have 'upstairs' index on the right, same should be on the left.

2.1 Einstein Convention

The OG rule...whenever you see two same indices, sum them. That's it! Let's make our hands dirty and look at some examples:

2.2 Examples

Matrix Multiplication:

Let us have the eigenvalue equation $M\mathbf{v} = \lambda\mathbf{v}$. We can write this as:

$$\sum_j M_{ij} v^j = \lambda v^i$$

Note two things here:

- The index j is summed over, so it does not come in the final expression (dummy index!).
- The index i occurs as a superscript on the right, so in the left also, the final expression should have the index i as a superscript.

Thus, using Einstein convention and correct index placement, the above equation can be written as:

$$M^i_j v^j = \lambda v^i$$

This can be visualised by treating each quantity as a 'box' with the indices as some 'hands' portruding out. When we sum, we just join these 'hands'. After taking the sum, the number of free hands decreases (index contraction)I A matrix has two hands and a vector has one hand. When we multiply a matrix with a vector, we obtain a vector, which should have one hand. This is represented in the diagram below:

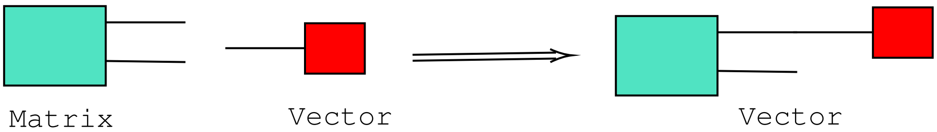


Figure 1: Matrix-vector multiplication. The final product has one free hand and is thus a vector.

The Scalar Product:

The dot product or scalar product of two vectors is a scalar (no hand). Then, there should be no free index in the expression. Thus, in the index notation:

$$\mathbf{v} \cdot \mathbf{v} = v^i v_i = v_i v^i$$

The last two expressions are same. The upstairs or downstairs indices do not matter, as these are summed over. Note that in this definition, we have used a *dual vector*, having a lower index. We can also define the dot product using the *regular* vector with an upper index but then a *metric* comes in.

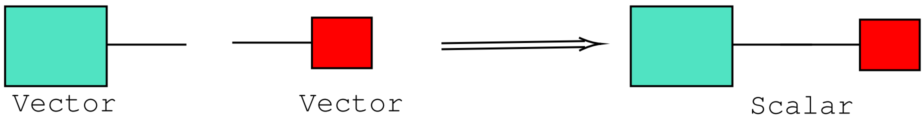


Figure 2: Dot product of two vectors. The final product is a scalar and has no free hands.

So basically a *scalar* is something that does not change under coordinate transformation, that is, if we go from a coordinate (x, y, z) to (x', y', z'), a scale $\lambda = \lambda'$.

Euclidean Vectors:

Any vector can be written as in terms of basis vectors: $\mathbf{A} = A^i \mathbf{e}_i$ where A^i are the components of the vector in the chosen basis. Now, we define

$$\mathbf{e}_n \cdot \mathbf{e}_m = g_{nm}$$

These g_{nm} are coefficients of metric tensor which will be discussed later. If these basis vectors are orthonormal, then the coefficients become the kronecker delta. Then we have the scalar product:

$$\begin{aligned}\mathbf{A} \cdot \mathbf{B} &= (A^m \mathbf{e}_m) \cdot B^n \mathbf{e}_n \\ &= (A^m B^n \mathbf{e}_m) \cdot \mathbf{e}_n \\ &= A^m B^n g_{nm}\end{aligned}$$

Note that we have used the *regular vector* with an upper index here, with the metric g. We can define a cross product of two vectors as:

$$(\mathbf{A} \times \mathbf{B})^i = \epsilon^{ijk} A^j B^k$$

where ϵ^{ijk} is the Levi-Civita symbol (cyclic permutation of i, j, k gives 1 and non-cyclic permutation gives -1 while repeated index in the symbol gives 0).

2.3 Some vector BS

Vectors as Directional Derivatives:

A vector can be thought of as a directional derivative. We define the directional derivative operator as:

$$\mathbf{v} \cdot \nabla = v^i \partial_i \equiv v^i \frac{\partial}{\partial x^i}$$

This is very similar to the vector expansion in terms of the basis vectors. Thus, the partial derivatives somewhat act like a basis. The basis of partial derivatives is indeed called a *coordinate basis*. Now let us calculate:

$$\mathbf{v} \cdot \nabla x^j = v^i \partial_i x^j = v^i \frac{\partial x^j}{\partial x^i} = v^i \delta_i^j = v^j$$

We have used the fact that coordinate components are independent of each other, that is, partial derivative of one component with respect to another gives a kronecker delta. Note that we have used the proper index placement here: $\partial_j \equiv \frac{\partial}{\partial x^j}$ has a lower index (by definition) while x^j has an upper index, thus kronecker delta has an upper as well as lower index.

Now, let us consider we have a position vector written in a basis $\{\mathbf{e}_i\}$, that is,

$$\mathbf{r} = x^i \mathbf{e}_i$$

Since $\mathbf{r}$ depends on the coordinate $\{x^i\}$, we can expand the differential displacement as:

$$\begin{aligned} d\mathbf{r} &= \frac{\partial \mathbf{r}}{\partial x^i} dx^i \\ &= \frac{\partial (x^j \mathbf{e}_j)}{\partial x^i} dx^i \\ &= \left(\frac{\partial x^j}{\partial x^i} \mathbf{e}_j + \frac{\partial \mathbf{e}_j}{\partial x^i} x^j \right) dx^i \end{aligned}$$

The second term is zero as the basis vectors are independent of the coordinates and the first term gives Kronecker delta, thus we have:

$$d\mathbf{r} = \mathbf{e_i} dx^i$$

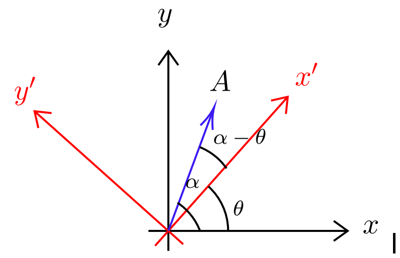
Comparing this with the first line of the previous expansion we have:

$$\boxed{\mathbf{e}_i = \frac{\partial \mathbf{r}}{\partial x^i}}$$

Thus any basis vector can be obtained from the partial derivative of the position vector with respect to the coordinates.

Vector Transformation:

Let us suppose we have a coordinate system (x, y, z) and we rotate it about the z-axis by an angle θ. The new coordinates are given by:



$$x' = x\cos\theta - y\sin\theta$$
$$y' = x\sin\theta + y\cos\theta$$
$$z' = z$$

These relations can be readily found out using the following: Suppose we have a point A making an angle α with the original system. Then we have $x = r\cos(\alpha), y = r\sin(\alpha)$. After rotation, the new coordinates are given by:

$$x' = r\cos(\alpha - \theta) = r\cos\alpha\cos\theta + r\sin\alpha\sin\theta$$
$$y' = r\sin(\alpha - \theta) = r\sin\alpha\cos\theta - r\cos\alpha\sin\theta$$

which gives the previous result. Note that the z coordinate does not change as the rotation is about the z-axis. Now we consider the infinitesimal displacement in the new coordinate frame:

$$\begin{aligned}
(ds')^2 &= (dx')^2 + (dy')^2 + (dz')^2 \\
&= (dx\cos\theta - dy\sin\theta)^2 + (dx\sin\theta + dy\cos\theta)^2 + dz^2 \\
&= dx^2(\cos^2\theta + \sin^2\theta) + dy^2(\cos^2\theta + \sin^2\theta) - \cancel{2dxdy\sin\theta\cos\theta} + \cancel{2dxdy\sin\theta\cos\theta} + dz^2 \\
&= dx^2 + dy^2 + dz^2 \\
&= ds^2
\end{aligned}$$

Thus we see that the infinitesimal displacement is invariant under coordinate transformation and is thus a scalar.

Now, note one thing: If we consider the new coordinates as a function of the old coordinate that is $x' \equiv x'(x, y, z)$, we can write:

$$(dx')^i = \frac{\partial x'^i}{\partial x^j} dx^j$$

Using this analogy, we can define the transformation of a vector as:

$$(v')^i = \frac{\partial x'^i}{\partial x^j} v^j$$

Thus a vector is a quantity which transform like this. The terms $\frac{\partial x'^i}{\partial x^j}$ are the components of the transformation matrix Λ^i_j. As we defined the transformation from x to x', we can also define the reverse transformation from x' to x as:

$$\begin{aligned} x^i &= \frac{\partial x^i}{\partial x'^j} x'^j \\ &= \left(\frac{\partial x^i}{\partial x'^j} \frac{\partial x'^j}{\partial x^k} \right) x^k \end{aligned}$$

Now in the above sum, j and k indices are summed over. We must obtain x^i from the right hand side also. Thus by observation, we can see that the term $\frac{\partial x^i}{\partial x'^j} \frac{\partial x'^j}{\partial x^k}$ must be equal to the kronecker delta δ^i_k.

3 Contravariant and Covariant: why the skibidi!

Let us suppose we have a vector space $\mathbf{V}$ and two bases $\{\mathbf{e}_i\}$ and $\{\mathbf{e}'_i\}$. We can the write the transformation of the basis into one another as:

$$\begin{aligned} \mathbf{e}_i &= \Lambda^j_i \mathbf{e}'_j \\ \mathbf{e}'_i &= (\Lambda^{-1})^j_i \mathbf{e}_j \end{aligned}$$

Now if we have a vector, we can write it in terms of the basis vectors as:

$$\mathbf{x} = (x')^j \mathbf{e}'_j = x^i \mathbf{e}_i = (x^i \Lambda^j_i) \mathbf{e}'_j$$

From this we get: $(x')^j = \Lambda^j_i x^i$. Well note that, in the transformation equation of the basis, if we have the primed basis in the left, then we had the inverse transformation matrix Λ^{-1} in the right, but here it is different (primed component in the left and Λ in the right). Thus, the basis vectors and the components transform in the "opposite" or "**contra**ry" way. Thus, these components are called the **contravariant** components of the vector.

Let us now consider the dual space $\mathbf{V}^*$[1] of the vector space $\mathbf{V}$. From the linearity property, we have:

$$f(x^i \mathbf{e}_i) = x^i f(\mathbf{e}_i) \equiv x^i f_i$$

Now, we use the basis transformation equation:

$$f_i = f(\mathbf{e}_i) = f(\Lambda^j_i \mathbf{e}'_j) = \Lambda^j_i f(\mathbf{e}'_j) = \Lambda^j_i f'_j$$

These f_i are the components of the "dual vector". Note that if we have unprimed things on the left, then we have the transformation matrix Λ on the right, which is similar to the transformation of the basis. Thus, we see that this transformation follows the same transformation as the basis vectors. Thus, these components are called the **covariant** components of the vector.

So, the components are named according to how the basis vectors transform. If they transform together, the are called **co**variant (and denoted by downstairs index) and if they transform in the opposite way, they are called **contra**variant (and denoted by upstairs index). The contravariant and the covariant components together form an 'invariance' like the scalar product (which do not change under coordinate transformation):

$$\mathbf{v} \cdot \mathbf{v} = v_i v^i$$

[1] The dual space is the set of all linear functionals, that is, linear maps $f : \mathbf{V} \to \mathbb{R}$.

We had earlier seen another definition of the inner product, using both contravariant components and the matric tensor, which was $\mathbf{v} \cdot \mathbf{v} = v^i v^j g_{ij}$. Comparing both these definitions, we can see a relation:

$$v_i = g_{ij} v^j$$

Thus when changing from contravariant to covariant, we just need to invoke the holy metric tensor (to be dicussed later further).

Note: In the Cartesian coordinates, the metric tensor is the Kronecker delta, that is, $g_{nm} = \delta_{nm}$ and hence the components of the vectors and dual vectors are the same, that is, $x^i = x_i$.

4 Why are tensors so sigma!

Nakahara defines tensor as:

> *A tensor T of type (p, q) is a multi-linear map that maps q vectors and p dual vectors to $\mathbb{R}$, that is:*
>
> $$T : \left(\bigotimes^{p} \mathbf{V}^* \right) \left(\bigotimes^{q} \mathbf{V} \right) \to \mathbb{R}$$

Dayummm!! 😑 Let us break this down. Consider a scalar which has no vector and no dual vector. Thus, it is a $(0, 0)$ type tensor. Now, let us consider a vector $\mathbf{v}$. This is a $(1, 0)$ tensor, that is, it maps a dual vector to a scalar. If we have a dual vector $\mathbf{f}$, then it is of type $(0, 1)$ and maps a vector to a scalar. This does not clear anything. Let us instead consider few examples:

Moment of Intertia Tensor:

Perhaps the first example of a tensor we had encountered during our classical mechanics course (which we had been told to understand just as a 'matrix').

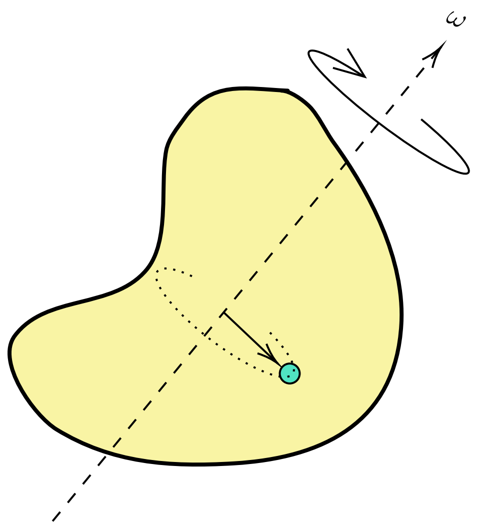


Figure 3: A rigid body rotating about an axis

Consider a rigid body made of tiny masses dm. Consider one such mass sitauted as a distance s from the fixed axis of rotation. It goes around a circle with speed $v = \omega s$. The angular momentum can be calculated as:

$$\mathbf{L} = \int (\mathbf{r} \times \mathbf{v}) dm = \int (\mathbf{r} \times (\boldsymbol{\omega} \times \mathbf{r})) dm = \int (\mathbf{r} \cdot \mathbf{r})\boldsymbol{\omega} - \cancelto{0}{(\mathbf{r} \cdot \boldsymbol{\omega})}\mathbf{r} dm = \boldsymbol{\omega} \underbrace{\int s^2 dm}_{I}$$

The integral is called the moment of inertia. In a more general case, where there is no fixed axis of rotation, we write:

$$\begin{aligned}\mathbf{L} &= \int (\mathbf{r} \times \mathbf{v}) dm \\ &= \int (\mathbf{r} \times (\boldsymbol{\omega} \times \mathbf{r})) dm \\ &= \int (\mathbf{r} \cdot \mathbf{r})\boldsymbol{\omega} - (\mathbf{r} \cdot \boldsymbol{\omega})\mathbf{r} dm\end{aligned}$$

We now write it in index notation, noting that $\omega^i = \delta^{ij}\omega_j$:

$$\begin{aligned}L^i &= \int (\mathbf{r} \cdot \mathbf{r})\delta^{ij}\omega_j - x^i(x^j\omega_j) dm \\ &= \omega_j \left(\int (\mathbf{r} \cdot \mathbf{r})\delta^{ij} - x^i x^j \ dm \right)\end{aligned}$$

The integral in the bracket is defined to be the inertia tensor:

$$I^{ij} = \int (\mathbf{r} \cdot \mathbf{r})\delta^{ij} - x^i x^j \ dm$$

Note that i and j goes from 1 to 3 and thus it has 9 components but since the expression is symmetric, we only have 6 independent components. This states that the angular momentum and the angular velocity are not necessarily parallel in some coordinate system where I have non-zero off-diagonal entries.

Electromagnetic Tensor:

The electromagnetic tensor is very useful in combining the electric field and magnetic field and finding their transformations. It is defined as:

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$$

where A_μ is the 4-potential. The indices μ and ν can take values from 0 to 3. The tensor has 16 components but only 6 of them are independent. The tensor is antisymmetric, that is, $F_{\mu\nu} = -F_{\nu\mu}$ and thus the diagonal entries are zero. We will discuss this later but Maxwell's equations can be written in a very compact form using the components of the electromagnetic tensor.

Electric-Susceptibility Tensor:

We had studied about polarisation in dielectrics in our classical electrodynamics course where we had often taken (for simplicity):

$$\mathbf{P} = \epsilon_0 \chi \mathbf{E}$$

Here we had taken the electric field to be parallel to the polarisation vector but in general, these are related by the susceptibility tensor as:

$$P^i = \epsilon_0 \chi^{ij} E^j$$

4.1 Matrices vs. Tensor? same same but different....

5 Metric Tensor: how yo mama's fatness is quantified!

Let us consider the spherical polar coordinates (r, θ, ϕ). Note that the coordinate displacement $d\phi$ does not have the dimension of length. So, while considering the displacement vector, we write $d\mathbf{r} \sim r \sin\theta d\phi \hat{\boldsymbol{\phi}}$. Thus, in general, for any displacement we write it in terms of the "metric tensor" g^{ij} as:

$$ds^2 = g_{ij} x^i x^j$$

In rectangular coordinates, we have $g_{ij} = \delta_{ij}$, that is, the metric tensor is just the identity matrix. In cylindrical coordinates where $dx^1 = d\rho, dx^2 = d\phi, dx^3 = z$, we have:

$$g_{ij} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \rho^2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Thus, the displacement is written as

$$ds^2 = g_{ij} dx^i dx^j = d\rho^2 + \rho^2 d\phi^2 + dz^2$$

In spherical polar coordinates, we have:

$$g_{ij} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2 \sin^2\theta \end{pmatrix}$$

Thus, the displacement is written as:

$$ds^2 = d\theta^2 + r^2 d\phi^2 + r^2 \sin^2\theta d\phi^2$$

If all of g_{ij} are non-negative we call that geometry "Riemannian" and if some of them are negative, it is termed "pseudo-Riemannian".

Now, how the hell do we calculate the components of the metric tensor?

Well we have previously seen how we could obtain the basis vectors using the partial derivatives of the coordinates. So, suppose we have to find the metric tensor for spherical polar coordinate system. For that, let us first write the position vector:

$$\mathbf{r} = x\mathbf{e}_x + y\mathbf{e}_y + z\mathbf{e}_z = r\sin\theta\cos\phi\ \mathbf{e}_x + r\sin\theta\sin\phi\ \mathbf{e}_y + r\cos\theta\ \mathbf{e}_z$$

From this we obtain:

$$\begin{aligned} \mathbf{e}_r &= \frac{\partial \mathbf{r}}{\partial r} = \sin\theta\cos\phi\ \mathbf{e}_x + \sin\theta\sin\phi\ \mathbf{e}_y + \cos\theta\ \mathbf{e}_z \\ \mathbf{e}_\theta &= \frac{\partial \mathbf{r}}{\partial \theta} = r\cos\theta\cos\phi\ \mathbf{e}_x + r\cos\theta\sin\phi\ \mathbf{e}_y - r\sin\theta\ \mathbf{e}_z \\ \mathbf{e}_\phi &= \frac{\partial \mathbf{r}}{\partial \phi} = -r\sin\theta\sin\phi\ \mathbf{e}_x + r\sin\theta\cos\phi\ \mathbf{e}_y \end{aligned}$$

Now, we had defined the metric tensor components to be the scalar product of the basis vectors. Also note that since the basis vectors are orthogonal, there will be no cross terms, so the tensor is diagonal. Using this we have:

$$\begin{aligned} g_{rr} &= \mathbf{e}_r \cdot \mathbf{e}_r = \sin^2\theta\cos^2\phi + \sin^2\theta\sin^2\phi + \cos^2\theta = 1 \\ g_{\theta\theta} &= \mathbf{e}_\theta \cdot \mathbf{e}_\theta = r^2\cos^2\theta\cos^2\phi + r^2\cos^2\theta\sin^2\phi + r^2\sin^2\theta = r^2 \\ g_{\phi\phi} &= \mathbf{e}_\phi \cdot \mathbf{e}_\phi = r^2\sin^2\theta\sin^2\phi + r^2\sin^2\theta\cos^2\phi = r^2\sin^2\theta \end{aligned}$$

This is exactly what we had written before. Thus using this procedure, we can easily find the components of the metric tensor and then just chill!

5.1 Metric in relativity:

We now consider the case of Minkowski space, where the coordinate displacements between two events are described by four component vector (**4 vector**):

$$dx^\mu = (dt, dx, dy, dz) \equiv (dt, d\mathbf{r})$$

The spacetime interval can be written in terms of the metric tensor $g_{\mu\nu}$ as:

$$ds^2 = dt^2 - dx^2 - dy^2 - dz^2 = g_{\mu\nu} dx^\mu dx^\nu$$

Thus in this case the metric tensor is:

$$g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 1 & 0 & -1 & 0 \\ 1 & 0 & 0 & -1 \end{pmatrix}$$

Note the convention. Other conventions include the obnoxious $(+,-,-,-)$ or the $(+,+,+,+)$ with an imaginary time. The overall thing is, spatial and temporal part should have some difference.

We also define the "proper time" as $d\tau = \frac{ds}{c}$. Since we take $c = 1$, then both are equivalent but let's take c to be c for once. Then,

$$c^2 d\tau^2 = c^2 dt^2 - dx^2 - dy^2 - dz^2 = c^2 dt^2 \left(1 - \frac{dx^2 + dy^2 + dz^2}{c^2 dt^2}\right) = c^2 dt^2 \left(1 - \frac{v^2}{c^2}\right)$$

From this we have

$$\boxed{\frac{\mathrm{d}\tau}{\mathrm{d}t} = \sqrt{1 - \frac{v^2}{c^2}} := \gamma}$$

6 Covariant Derivatives: levelling up the rizz!

7 Differential Forms: Dayuum!!

Definition. Suppose $C \subset \mathbb{R}^2$ be a curve and let $p \in C$ is a point. The tangent space to C at p is the set of all vectors tangent to C at p and is denoted by T_pC.