Notation for vectors

October 12, 2014

To understand rotations, we first distinguish between *space* and a *vector space*. Euclidean 3-space can be thought of as either, but this is not always true. For example, if we wish to discuss the two dimensional sphere, $\mathbb{S}^2$ (the surface of a 3-dim ball) it is clear that it is a space but not a vector space. If we want to think of vectors associated with this 2-dim space, we need to imagine a plane tangent at each point of the sphere. Each plane is a vector space. Then a *vector field* on the sphere is a selection of one vector from each of these tangent planes.

We may choose a basis, called a *frame*, for the vectors in each of these tangent planes, and it is convenient to choose a pair of orthonormal vectors for this basis. Any vector in a given tangent plane may be written as a linear combination of the vectors in the frame at that point. The complete collection of orthonormal frames for the whole sphere is called an *orthonormal frame field*.

Now consider Euclidean 3-space, $\mathbb{R}^3$. We imagine a 3-dim vector space attached to each point of this space. In this case the Euclidean 3-space may also be thought of as a vector space, with each point $P = (x, y, z)$ having an associated vector, $\mathbf{x}$. This allows the possibility of using the same basis at each point $\mathbf{x}$ by simply adding $\mathbf{x}$ to each basis vector at the origin. For the frame at the origin we may choose any basis for the vector space – any three linearly independent vectors. For an *orthonormal frame* we choose three orthogonal unit vectors $(\mathbf{i}, \mathbf{j}, \mathbf{k})$:

$$\begin{aligned} \mathbf{i}\cdot\mathbf{i} = \mathbf{j}\cdot\mathbf{j} = \mathbf{k}\cdot\mathbf{k} &= 1 \\ \mathbf{i}\cdot\mathbf{j} = \mathbf{j}\cdot\mathbf{k} = \mathbf{k}\cdot\mathbf{i} &= 0 \end{aligned}$$

If we add this to each vector $\mathbf{x}$ to get an orthonormal frame field for Euclidean 3-space, we have the very special Cartesian frame field. A Cartesian frame field exists only if the *space* is also a *vector space*.

We do not have to choose the Cartesian frame field. For example, if we use spherical coordinates (r, θ, φ) to label points of the space, then at any point we may choose a set of orthonormal basis vectors $\left(\hat{\boldsymbol{r}}, \hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\varphi}}\right)$ with each unit vector pointing in the direction of increase of the corresponding coordinate. Clearly this basis will not be the same at each point, with $\hat{\boldsymbol{r}}$, for example, pointing in different directions at each different value of θ and φ.

Next, fix an orthonormal frame $\mathbf{i}, \mathbf{j}, \mathbf{k}$, at the origin of Euclidean 3-space, and a second orthonormal frame, $\mathbf{i}', \mathbf{j}', \mathbf{k}'$, at any point $\mathbf{x}$. Any vector $\mathbf{v}$ at the origin may be written as a linear combination

$$\mathbf{v} = v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k}$$

and because $\mathbb{R}^3$ is also a vector space, $\mathbf{v}+\mathbf{x}$ is the same vector $\mathbf{v}$ transplanted to the point $\mathbf{x}$. Then $(\mathbf{v}+\mathbf{x}) - \mathbf{x}$ is the same vector relocated to $\mathbf{x}$, and it may be expressed in the secod frame,

$$\mathbf{v} = v_1'\mathbf{i}' + v_2'\mathbf{j}' + v_3'\mathbf{k}'$$

The frames $\mathbf{i}, \mathbf{j}, \mathbf{k}$ and $\mathbf{i}', \mathbf{j}', \mathbf{k}'$ thus differ in two ways: first, they are *translated* by the vector $\mathbf{x}$, and second, they may be *rotated* relative to one another.

The combined transformations, translation and rotation, comprise the Euclidean group in 3-dimensions.

We now turn to a study of the rotations, which comprise the orthogonal group in 3-dimensions. The orthogonal group relates any two orthonormal frames at the same point.

1 Basis transformations and the Einstein convention

We now seek the relationship between two orthonormal bases with a common origin. The first key fact is that the transformation is *linear*, and this is immediate by the definition of a vector basis.

1.1 Passive and active transformations

There are two ways to think of a rotation: active and passive.

A passive transformation is one where we think of all vectors as staying fixed while the basis changes. To accomplish this, we note that the new basis is comprised of vectors, so the vectors of the new basis $(\mathbf{i}', \mathbf{j}', \mathbf{k}')$ may be expanded in the old:

$$\begin{aligned}
\mathbf{i}' &= a_{11}\mathbf{i} + a_{21}\mathbf{j} + a_{31}\mathbf{k} \\
\mathbf{j}' &= a_{12}\mathbf{i} + a_{22}\mathbf{j} + a_{32}\mathbf{k} \\
\mathbf{k}' &= a_{13}\mathbf{i} + a_{23}\mathbf{j} + a_{33}\mathbf{k}
\end{aligned}$$

or as a matrix equation,

$$\begin{pmatrix} \mathbf{i}' & \mathbf{j}' & \mathbf{k}' \end{pmatrix} = \begin{pmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

This transformation must be invertible,

$$\begin{pmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \end{pmatrix} = \begin{pmatrix} \mathbf{i}' & \mathbf{j}' & \mathbf{k}' \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix}$$

$$\begin{pmatrix} \mathbf{i} \\ \mathbf{j} \\ \mathbf{k} \end{pmatrix} = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix} \begin{pmatrix} \mathbf{i}' \\ \mathbf{j}' \\ \mathbf{k}' \end{pmatrix}$$

Then, if a vector $\mathbf{v}$ is expanded as a linear combination

$$\mathbf{v} = v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k}$$

in the original basis, its new components will be found by substituting the inverse transformation for $\mathbf{i}, \mathbf{j}, \mathbf{k}$. The same vector $\mathbf{v}$ will have new components in the new basis, $\mathbf{v} = v_1'\mathbf{i}' + v_2'\mathbf{j}' + v_3'\mathbf{k}'$. Carrying out this substitution, v_1' is given by

$$v_1'\mathbf{i}' = (b_{11}v_1 + b_{12}v_2 + b_{13}v_3)\,\mathbf{i}'$$

and similarly for the other two components, leading to

$$\begin{pmatrix} v_1' \\ v_2' \\ v_3' \end{pmatrix} = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$$

An active transformation is one in which we leave the basis fixed, but transform all vectors. In this case it is the vector $\mathbf{v}$ which which changes to a *different* vector $\mathbf{v}'$, with different components given by

$$\begin{pmatrix} v_1' \\ v_2' \\ v_3' \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$$

but both vectors expanded in the *same* basis.

1.2 Einstein summation convention

All of this is much easier in index notation. Let the three basis vectors be denoted by $\hat{\mathbf{e}}_i$, $i = 1, 2, 3$, so that

$$\begin{aligned} \hat{\mathbf{e}}_1 &= \hat{\mathbf{i}} \\ \hat{\mathbf{e}}_2 &= \hat{\mathbf{j}} \\ \hat{\mathbf{e}}_3 &= \hat{\mathbf{k}} \end{aligned}$$

and similarly for $\hat{\mathbf{e}}'_i$. Then the passive basis transformations above may be written as

$$\begin{aligned} \hat{\mathbf{e}}'_i &= \sum_{j=1}^{3} \hat{\mathbf{e}}_j a_{ji} \\ \hat{\mathbf{e}}_i &= \sum_{j=1}^{3} \hat{\mathbf{e}}'_j b_{ji} \end{aligned}$$

and the vector expansions as

$$\mathbf{v} = \sum_{j=1}^{3} v_j \hat{\mathbf{e}}_j$$

It is easy to see that this will lead us to write $\sum_{j=1}^{3}$ millions of times. The Einstein convention avoids this by noting that when there is a sum there is also a repeated index – j, in the cases above. Also, we almost never repeat an index that we do not sum, so we may drop the summation sign. Thus,

$$\begin{aligned} \sum_{j=1}^{3} \hat{\mathbf{e}}_j a_{ji} &\implies \hat{\mathbf{e}}_j a_{ji} \\ \sum_{j=1}^{3} \hat{\mathbf{e}}_j^{'} b_{ji} &\implies \hat{\mathbf{e}}_j^{'} b_{ji} \\ \sum_{j=1}^{3} v_j \hat{\mathbf{e}}_j &\implies v_j \hat{\mathbf{e}}_j \end{aligned}$$

The repeated index is called a dummy index, and it does not matter what letter we choose for it,

$$v_j \hat{\mathbf{e}}_j = v_k \hat{\mathbf{e}}_k$$

as long as we do not use an index that we have used elsewhere in the same expression. Thus, in the basis change examples above, we cannot use i as the dummy index because it is used to distinguish three independent equations:

$$\begin{aligned} \hat{\mathbf{e}}_1^{'} &= \hat{\mathbf{e}}_j a_{k1} \\ \hat{\mathbf{e}}_2^{'} &= \hat{\mathbf{e}}_j a_{j2} \\ \hat{\mathbf{e}}_3^{'} &= \hat{\mathbf{e}}_j a_{j3} \end{aligned}$$

Such an index is called a free index. Free indices must match in every term of an expression.

Since the basis is orthonormal, we know that the dot product is given by

$$\hat{\mathbf{e}}_i \cdot \hat{\mathbf{e}}_j = \delta_{ij}$$

where δ_{ij} is the Kronecker delta, equal to 1 if $i = j$ and to 0 if $i \neq j$. Notice that the expression above represents nine separate equations. If we repeat the index, we have a single equation

$$\begin{aligned} \hat{\mathbf{e}}_k \cdot \hat{\mathbf{e}}_k &= \delta_{kk} \\ &= 3 \end{aligned}$$

Be sure you understand why the result is 3.

We can find the relationship between the matrices a_{ij} and b_{ij}, since, substituting one basis change into the other,

$$\begin{aligned} \hat{\mathbf{e}}_i^{'} &= \hat{\mathbf{e}}_j a_{ji} \\ &= \left(\hat{\mathbf{e}}_k^{'} b_{kj}\right) a_{ji} \\ &= \hat{\mathbf{e}}_k^{'} \left(b_{kj} a_{ji}\right) \end{aligned}$$

Taking the dot product with $\hat{\mathbf{e}}_m^{'}$ (notice that we cannot use i, j or k), we have

$$\begin{aligned} \hat{\mathbf{e}}_m^{'} \cdot \hat{\mathbf{e}}_i^{\prime} &= \hat{\mathbf{e}}_m^{'} \cdot \hat{\mathbf{e}}_k^{'} \left(b_{kj} a_{ji}\right) \\ \delta_{mi} &= \delta_{mk} \left(b_{kj} a_{ji}\right) \\ &= b_{mj} a_{ji} \end{aligned}$$

and since $\delta_{mi} = \delta_{im}$ is the identity matrix, 1, this shows that

$$BA = 1$$

so that the matrix B with components b_{ij} is inverse to A,

$$B = A^{-1}$$

2 The metric, and index position

Notice that active and passive transformations differ in the way the components of vectors transform,

$$\begin{aligned} v_i^{\prime} &= A_{ij} v_j \quad \textit{active} \\ v_i^{\prime} &= A_{ij}^{-1} v_j \quad \textit{passive} \end{aligned}$$

This property has its origin in the fact that there are two distinct ways to associate vectors spaces with spaces involving tangent vectors, and integrals along curves. Instead of digressing to talk about this in detail, we will introduce the resulting notation.

The essential feature is that vectors that transform with A will be written with raised indices, v^i, while vectors which transform with A^{-1} will be written with lowered indices. A new object, the metric, will map these to each other in a $1-1$ way so that we may regard v^i and v_i as different ways of representing the same vector.

We may define a rotation as any linear transformation of generalized coordinates x^i which preserves the length. In Cartesian coordinates, an infinitesimal length interval is given by

$$\begin{aligned} ds^2 &= \sum_{i=1}^{3} dx^i dx^i \\ &= \sum_{i,j=1}^{3} \delta_{ij} dx^i dx^j \\ &= \delta_{ij} dx^i dx^j \end{aligned}$$

where

$$\delta_{ij} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

and in the final line we employ the summation convention. The matrix δ_{ij} is called the *metric*. From this we may find the length interval in any other coordinate system. Let the new coordinates be y^i so that the coordinate transformation and its inverse are given by

$$y^i(\mathbf{x}), x^i(\mathbf{y})$$

Then taking the differential,

$$\begin{aligned} dx^i &= \frac{\partial x^i}{\partial y^j} dy^j \\ dy^i &= \frac{\partial y^i}{\partial x^j} dx^j \end{aligned}$$

we substitute to find

$$\begin{aligned} ds^2 &= \delta_{ij} dx^i dx^j \\ &= \delta_{ij} \left(\frac{\partial x^i}{\partial y^k} dy^k \right) \left(\frac{\partial x^j}{\partial y^m} dy^m \right) \\ &= \left(\delta_{ij} \frac{\partial x^i}{\partial y^k} \frac{\partial x^j}{\partial y^m} \right) dy^k dy^m \end{aligned}$$

We define the metric in the new coordinate system to be

$$g_{km} = \delta_{ij} \frac{\partial x^i}{\partial y^k} \frac{\partial x^j}{\partial y^m}$$

so that we always write

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

In the special case of Cartesian coordinates, g_{ij} takes the special form δ_{ij}.

The length of any vector, v^i, is given using the metric,

$$\|\mathbf{v}\|^2 = g_{ij} v^i v^j$$

For any vector, v^i, we define a related object v_i by

$$v_i \equiv g_{ij} v^j$$

so that we may write the length of the vector as simply

$$\|\mathbf{v}\|^2 = v^i v_i$$

Because the metric is symmetric, we have $g_{ij} v^i v^j = g_{ji} v^i v^j = v_j v^j$, so

$$v^i v_i = v_i v^i$$

Notice how each of these objects changes when we perform a coordinate transformation. Let $x^i(\lambda)$ be a curve expressed in x^i coordinates, so that $\frac{dx^i}{d\lambda}$ is a tangent vector to the curve. Then from $dy^i = \frac{\partial y^i}{\partial x^j} dx^j$ we see that the same tangent vector in y^i coordinates must be given by

$$\begin{aligned} \frac{dy^i}{d\lambda} &= \frac{\partial y^i}{\partial x^j} \frac{dx^j}{d\lambda} \\ &= J^i_{\ j} \frac{dx^j}{d\lambda} \end{aligned}$$

where $J^i_{\ j} = \frac{\partial y^i}{\partial x^j}$ is the Jacobian matrix of the transformation. Any vector which transforms this way is called *contravariant vector* and is written with the index up.

Next look at the metric. We have $g_{km} = \delta_{ij} \frac{\partial x^i}{\partial y^k} \frac{\partial x^j}{\partial y^m}$ and the same would be true even if the initial metric hadn't been Cartesian. Letting $\bar{J}^i_{\ j} = \frac{\partial x^i}{\partial y^j}$ denote the inverse to $J^i_{\ j}$, where x^i and y^i are two arbitrary coordinate systems, and $\tilde{g}_{km} = g_{km}(y)$,

$$\tilde{g}_{km} = g_{ij} \bar{J}^i_{\ k} \bar{J}^j_{\ m}$$

The metric g_{ij} is an example of a covariant matrix. Using this transformation together with the definition $v_i = g_{ij} v^j$ we see that v_i must transform as

$$\begin{aligned} v_i(y) = \tilde{v}_i &= \tilde{g}_{ij} \tilde{v}^j \\ &= \left(g_{km} \bar{J}^k_{\ i} \bar{J}^m_{\ j} \right) \bar{J}^j_{\ l} v^l \\ &= g_{km} \bar{J}^k_{\ i} \left(\bar{J}^m_{\ j} \bar{J}^j_{\ l} \right) v^l \\ &= g_{km} \bar{J}^k_{\ i} \delta^m_l v^l \\ &= g_{km} v^m \bar{J}^k_{\ i} \\ &= v_k \bar{J}^k_{\ i} \end{aligned}$$

This is the transformation law for a *covariant vector*. The inverse Jacobian is applied to all lowered indices.

Covariant and contravariant vectors underlie the active and passive transformation laws. From the behavior of each, we immediately see that a sum over one raised and one lowered index – exactly the demand of the Einstein convention – is *invariant*:

$$\begin{aligned} \tilde{v}^i \tilde{v}_i &= \left(\bar{J}^i_{\ m} v^m \right) \left(v_k \bar{J}^k_{\ i} \right) \\ &= \bar{J}^k_{\ i} \bar{J}^i_{\ m} v^m v_k \\ &= \delta^k_m v^m v_k \\ &= v^m v_m \end{aligned}$$

so the length of a vector is unchanged by coordinate transformations.