Orthogonal transformations

October 12, 2014

1 Defining property

The squared length of a vector is given by taking the dot product of a vector with itself,

$$\begin{aligned} v^2 &= \mathbf{v} \cdot \mathbf{v} \\ &= g_{ij} v^i v^j \end{aligned}$$

An *orthogonal transformation* is a linear transformation of a vector space that preserves lengths of vectors. This defining property may therefore be written as a linear transformation,

$$\mathbf{v}' = O\mathbf{v}$$

such that

$$\mathbf{v}' \cdot \mathbf{v}' = \mathbf{v} \cdot \mathbf{v}$$

Write this definition in terms of components using index notation. The transformation O must have mixed indices so that the sum is over one index of each type and the free indices match,

$$v'^i = O^i_{\ j} v^j$$

we have

$$\begin{aligned} g_{ij} v'^i v'^j &= g_{ij} v^i v^j \\ g_{ij} \left(O^i_{\ k} v^k \right) \left(O^j_{\ m} v^m \right) &= g_{ij} v^i v^j \\ \left(g_{ij} O^i_{\ k} O^j_{\ m} \right) v^k v^m &= g_{km} v^k v^m \end{aligned}$$

or, bringing both terms to the same side,

$$\left(g_{ij} O^i_{\ k} O^j_{\ m} - g_{km} \right) v^k v^m = 0$$

Because v^k is arbitrary, we can easily make six independent choices so that the resulting six matrices $v^k v^m$ span the space of all symmetric matrices. The only way for all six of these to vanish when contracted on $\left(g_{ij} O^i_{\ k} O^j_{\ m} - g_{km} \right)$ is if

$$g_{ij} O^i_{\ k} O^j_{\ m} - g_{km} = 0$$

This is the defining property of an orthogonal transformation.

In Cartesian coordinates, $g_{ij} = \delta_{ij}$, and this condition may be written as

$$O^t O = 1$$

where O^t is the transpose of O. This is equivalent to $O^t = O^{-1}$.

Notice that nothing in the preceeding arguments depends on the dimension of $\mathbf{v}$ being three. We conclude that orthogonal transformations in any dimension n must satisfy $O^t O = 1$, where O is an $n \times n$ matrix. Consider the determinant of the defining condition,

$$\begin{aligned} \det 1 &= \det\left(O^t O\right) \\ 1 &= \det O^t \det O \end{aligned}$$

and since the determinant of the transpose of a matrix equals the determinant of the original matrix, we have $\det O = \pm 1$. Any orthogonal transformation with $\det O = -1$ is just the parity operator, $P = \begin{pmatrix} -1 & & \\ & -1 & \\ & & -1 \end{pmatrix}$ times an orthogonal transformation with $\det O = +1$, so if we treat parity independently we have the *special orthogonal group*, $SO(n)$, of unit determinant orthogonal transformations.

1.1 More detail

If the argument above is not already clear, it is easy to see what is happening in 2-dimensions. We have

$$v^k v^m = \begin{pmatrix} v^1 v^1 & v^1 v^2 \\ v^2 v^1 & v^2 v^2 \end{pmatrix}$$

which is a symmetric matrix. But $\left(g_{ij} O^i{}_k O^j{}_m - g_{km}\right) v^k v^m = 0$ must hold for *all* choices of v^i. If we make the choice $v^i = (1, 0)$ then

$$v^k v^m = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

and $\left(g_{ij} O^i{}_k O^j{}_m - g_{km}\right) v^k v^m = 0$ implies $g_{ij} O^i{}_1 O^j{}_1 = g_{11}$. Now choose $v^i = (0, 1)$ so that $v^k v^m = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ and we see we must also have $g_{ij} O^i{}_2 O^j{}_2 = g_{22}$. Finally, let $v^i = (1, 1)$ so that $v^k v^m = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$. This gives

$$\left(g_{ij} O^i{}_1 O^j{}_1 - g_{11}\right) + \left(g_{ij} O^i{}_1 O^j{}_2 - g_{12}\right) + \left(g_{ij} O^i{}_2 O^j{}_1 - g_{21}\right) + \left(g_{ij} O^i{}_2 O^j{}_2 - g_{22}\right) = 0$$

but by the previous two relations, the first and fourth term already vanish and we have

$$\left(g_{ij} O^i{}_1 O^j{}_2 - g_{12}\right) + \left(g_{ij} O^i{}_2 O^j{}_1 - g_{21}\right) = 0$$

Because $g_{ij} = g_{ji}$, these two terms are the same, and we conclude

$$g_{ij} O^i{}_k O^j{}_m = g_{km}$$

for all km. In 3-dimensions, six different choices for v^i will be required.

2 Infinitesimal generators

We work in Cartesian coordinates, where we can use matrix notation. Then, for the real, 3-dimensional representation of rotations, we require

$$O^t = O^{-1}$$

Notice that the identity satisfies this condition, so we may consider linear transformations near the identity which also satisfy the condition. Let

$$O = 1 + \varepsilon$$

where $[\varepsilon]_{ij} = \varepsilon_{ij}$ are all small, $|\varepsilon_{ij}| \ll 1$ for all i, j. Keeping only terms to first order in ε_{ij}, we have:

$$\begin{aligned} O^t &= 1 + \varepsilon^t \\ O^{-1} &= 1 - \varepsilon \end{aligned}$$

where we see that we have the right form for O^{-1} by computing

$$\begin{aligned} OO^{-1} &= (1+\varepsilon)(1-\varepsilon) \\ &= 1 - \varepsilon^2 \\ &\approx 1 \end{aligned}$$

correct to first order in ε. Now we impose our condition,

$$\begin{aligned} O^t &= O^{-1} \\ 1 + \varepsilon^t &= 1 - \varepsilon \\ \varepsilon^t &= -\varepsilon \end{aligned}$$

so that the matrix ε must be antisymmetric.

Next, we write the most general antisymmetric 3×3 matrix as a linear combination of a convenient basis,

$$\begin{aligned} \varepsilon &= w^i J_i \\ &= w^1 \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} + w^2 \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} + w^3 \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & w^2 & -w^3 \\ -w^2 & 0 & w^1 \\ w^3 & -w^1 & 0 \end{pmatrix} \end{aligned}$$

where $\left|w^i\right| \ll 1$. Notice that the components of the three matrices J_i are neatly summarized by

$$[J_i]_{jk} = \varepsilon_{ijk}$$

where ε_{ijk} is the totally antisymmetric Levi-Civita tensor. For example, $[J_1]_{ij} = \varepsilon_{1ij} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}$. The matrices J_i are called the *generators* of the transformations. The most general antisymmetric is then a linear combination of the three generators.

Knowing the generators is enough to recover an arbitrary rotation. Starting with

$$O = 1 + \varepsilon$$

we may apply O repeatedly, taking the limit

$$\begin{aligned} O(\theta) &= \lim_{n \longrightarrow \infty} O^n \\ &= \lim_{n \longrightarrow \infty} (1+\varepsilon)^n \\ &= \lim_{n \longrightarrow \infty} \left(1 + w^i J_i\right)^n \end{aligned}$$

Let w be the length of the infinitesmal vector w_i, so that $w_i = w n_i$, where n_i is a unit vector. Then the limit is taken in such a way that

$$\lim_{n \longrightarrow \infty} nw = \theta$$

where θ is finite. Using the binomial expansion, $(a+b)^n = \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} a^{n-k} b^k$ we have

$$\begin{aligned}
\lim_{n\longrightarrow\infty} O^n &= \lim_{n\longrightarrow\infty} \left(1 + w^i J_i\right)^n \\
&= \lim_{n\longrightarrow\infty} \sum_{k=0}^{n} \frac{n!}{k!\,(n-k)!} (1)^{n-k} \left(w^i J_i\right)^k \\
&= \lim_{n\longrightarrow\infty} \sum_{k=0}^{n} \frac{n\,(n-1)\,(n-2)\cdots(n-k+1)}{k!} \frac{1}{n^k} \left((nw)\, n^i J_i\right)^k \\
&= \lim_{n\longrightarrow\infty} \sum_{k=0}^{n} \frac{\frac{n}{n}\left(\frac{n-1}{n}\right)\left(\frac{n-2}{n}\right)\cdots\left(\frac{n-k+1}{n}\right)}{k!} \left(\theta n^i J_i\right)^k \\
&= \sum_{k=0}^{\infty} \frac{1}{k!} \left(\theta n^i J_i\right)^k \\
&\equiv \exp\left(\theta n^i J_i\right)
\end{aligned}$$

We define the exponential of a matrix by the power series for the exponential, applied using powers of the matrix. This is the matrix for a rotation through an angle θ around an axis in the direction of $\mathbf{n}$.

To find the detailed form of a general rotation, we now need to find powers of $n^i J_i$. This turns out to be straightforward:

$$\begin{aligned}
\left[n^i J_i\right]^j_{\ k} &= n^i \varepsilon_{i\ k}^{\ j} \\
\left[\left(n^i J_i\right)^2\right]^m_{\ n} &= \left(n^i \varepsilon_{i\ k}^{\ m}\right)\left(n^j \varepsilon_{j\ n}^{\ k}\right) \\
&= n^i n^j \varepsilon_{i\ k}^{\ m} \varepsilon_{j\ n}^{\ k} \\
&= -n^i n^j \varepsilon_{i\ k}^{\ m} \varepsilon_{jn}^{\ \ \ k} \\
&= -n^i n^j \left(\delta_{ij}\delta^m_n - \delta_{in}\delta^m_j\right) \\
&= -\left(\delta_{ij} n^i n^j\right)\delta^m_n + \delta_{in} n^i \delta^m_j n^j \\
&= -\delta^m_n + n^m n_n \\
&= -\left(\delta^m_n - n^m n_n\right) \\
\left[\left(n^i J_i\right)^3\right]^m_{\ n} &= -\left(\delta^m_k - n^m n_k\right) n^i \varepsilon_{i\ n}^{\ k} \\
&= -\delta^m_k n^i \varepsilon_{i\ n}^{\ k} + n^m n_k n^i \varepsilon_{i\ n}^{\ k} \\
&= -n^i \varepsilon_{i\ n}^{\ m} + n^m n^k n^i \varepsilon_{ikn} \\
&= -\left[n^i J_i\right]^m_{\ n}
\end{aligned}$$

where $n^m n^k n^i \varepsilon_{ikn} = n^m \left(\mathbf{n} \times \mathbf{n}\right) = 0$. The powers come back to $n^i J_i$ with only a sign change, so we can divide the series into even and odd powers. For all $k > 0$,

$$\begin{aligned}
\left[\left(n^i J_i\right)^{2k}\right]^m_{\ n} &= (-1)^k \left(\delta^m_n - n^m n_n\right) \\
\left[\left(n^i J_i\right)^{2k+1}\right]^m_{\ n} &= (-1)^k \left[n^i J_i\right]^m_{\ n}
\end{aligned}$$

For $k = 0$ we have the identity, $\left[\left(n^i J_i\right)^0\right]^m_{\ n} = \delta^m_n$.

We can now compute the exponential explicitly:

$$\left[O\left(\theta, \hat{\mathbf{n}}\right)\right]^m_{\ n} = \left[\exp\left(\theta n^i J_i\right)\right]^m_{\ n}$$

$$
\begin{aligned}
&= \left[\sum_{k=0}^{\infty} \frac{1}{k!} \left(\theta n^i J_i\right)^k\right]^m_{\ n} \\
&= \left[\sum_{l=0}^{\infty} \frac{1}{(2l)!} \left(\theta n^i J_i\right)^{2l}\right]^m_{\ n} + \left[\sum_{l=0}^{\infty} \frac{1}{(2l+1)!} \left(\theta n^i J_i\right)^{2l+1}\right]^m_{\ n} \\
&= \left[1 + \sum_{l=1}^{\infty} \frac{1}{(2l)!} \left(\theta n^i J_i\right)^{2l}\right]^m_{\ n} + \left[\sum_{l=0}^{\infty} \frac{1}{(2l+1)!} \left(\theta n^i J_i\right)^{2l+1}\right]^m_{\ n} \\
&= \delta^m_n + \sum_{l=1}^{\infty} \frac{(-1)^l}{(2l)!} \theta^{2l} \left(\delta^m_n - n^m n_n\right) + \sum_{l=0}^{\infty} \frac{(-1)^l}{(2l+1)!} \theta^{2l+1} \left[n^i J_i\right]^m_{\ n} \\
&= \delta^m_n + (\cos\theta - 1)\left(\delta^m_n - n^m n_n\right) + \sin\theta\, n^i \varepsilon_{i\ n}^{\ m}
\end{aligned}
$$

where we get $(\cos\theta - 1)$ because the $l = 0$ term is missing from the sum.

To see what this means, let O act on an arbitrary vector $\mathbf{v}$, and write the result in normal vector notation,

$$
\begin{aligned}
\left[O\left(\theta, \hat{\mathbf{n}}\right)\right]^m_{\ n} v^n &= \delta^m_n v^n + (\cos\theta - 1)\left(\delta^m_n - n^m n_n\right) v^n + \sin\theta\, n^i \varepsilon_{i\ n}^{\ m} v^n \\
&= v^m + (\cos\theta - 1)\left(\delta^m_n v^n - n^m n_n v^n\right) - \sin\theta\, n^i v^n \varepsilon_{in}^{\quad m} \\
&= v^m + (\cos\theta - 1)\left(v^m - (\mathbf{n}\cdot\mathbf{v})\, n^m\right) - \left[\mathbf{n}\times\mathbf{v}\right]^m \sin\theta
\end{aligned}
$$

Going fully to vector notation,

$$
O\left(\theta, \hat{\mathbf{n}}\right)\mathbf{v} = \mathbf{v} + (\cos\theta - 1)\left(\mathbf{v} - (\mathbf{n}\cdot\mathbf{v})\,\mathbf{n}\right) - (\mathbf{n}\times\mathbf{v})\sin\theta
$$

Finally, define the components of $\mathbf{v}$ parallel and perpendicular to the unit vector $\mathbf{n}$:

$$
\begin{aligned}
\mathbf{v}_{\parallel} &= (\mathbf{v}\cdot\mathbf{n})\,\mathbf{n} \\
\mathbf{v}_{\perp} &= \mathbf{v} - (\mathbf{v}\cdot\mathbf{n})\,\mathbf{n}
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
O\left(\theta, \hat{\mathbf{n}}\right)\mathbf{v} &= \mathbf{v} - \left(\mathbf{v} - (\mathbf{v}\cdot\mathbf{n})\,\mathbf{n}\right) + \left(\mathbf{v} - (\mathbf{v}\cdot\mathbf{n})\,\mathbf{n}\right)\cos\theta - \sin\theta\,(\mathbf{n}\times\mathbf{v}) \\
&= \mathbf{v}_{\parallel} + \mathbf{v}_{\perp}\cos\theta - \sin\theta\,(\mathbf{n}\times\mathbf{v})
\end{aligned}
$$

This expresses the rotated vector in terms of three mutually perpendicular vectors, $\mathbf{v}_{\parallel}, \mathbf{v}_{\perp}, (\mathbf{n}\times\mathbf{v})$. The direction $\mathbf{n}$ is the axis of the rotation. The part of $\mathbf{v}$ parallel to $\mathbf{n}$ is therefore unchanged. The rotation takes place in the plane perpendicular to $\mathbf{n}$, and this plane is spanned by $\mathbf{v}_{\perp}, (\mathbf{n}\times\mathbf{v})$. The rotation in this plane takes $\mathbf{v}_{\perp}$ into the linear combination $\mathbf{v}_{\perp}\cos\theta - (\mathbf{n}\times\mathbf{v})\sin\theta$, which is exactly what we expect for a rotation of $\mathbf{v}_{\perp}$ through an angle θ. The rotation $O\left(\theta, \hat{\mathbf{n}}\right)$ is therefore a rotation by θ around the axis $\hat{\mathbf{n}}$.