The Matrix Representation of a Three-Dimensional Rotation—Revisited

In a handout entitled *The Matrix Representation of a Three-Dimensional Rotation*, I provided a derivation of the explicit form for most general 3×3 rotation matrix, $R(\hat{\boldsymbol{n}}, \theta)$ that describes the counterclockwise rotation by an angle θ about a fixed axis $\hat{\boldsymbol{n}}$. For example, the matrix representation of the counterclockwise rotation by an angle θ about a fixed z-axis is given by [cf. eq. (7.18) on p. 129 of Boas]:

$$R(\boldsymbol{k}, \theta) \equiv \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} . \tag{1}$$

The general rotation matrix $R(\hat{\boldsymbol{n}}, \theta)$ satisfies the following relations:

$$R(\boldsymbol{n}, \theta + 2\pi k) = R(\boldsymbol{n}, \theta) , \qquad k = 0, \pm 1 \ \pm 2 \ldots ,$$

$$[R(\boldsymbol{n}, \theta)]^{-1} = R(\boldsymbol{n}, -\theta) = R(-\boldsymbol{n}, \theta) . \tag{2}$$

Combining these two results, it follows that

$$R(\boldsymbol{n}, 2\pi - \theta) = R(-\boldsymbol{n}, \theta) , \tag{3}$$

which implies that any three-dimensional rotation can be described by a counterclockwise rotation by θ about an arbitrary axis $\hat{\boldsymbol{n}}$, where $0 \leq \theta \leq \pi$. However, if we substitute $\theta = \pi$ in eq. (3), we conclude that

$$R(\boldsymbol{n}, \pi) = R(-\boldsymbol{n}, \pi) , \tag{4}$$

which means that for the special case of $\theta = \pi$, $R(\boldsymbol{n}, \pi)$ and $R(-\boldsymbol{n}, \pi)$ represent the *same* rotation. Finally, if $\theta = 0$, then $R(\boldsymbol{n}, 0) = \mathbf{I}$ is the identity operator, independently of the direction of $\hat{\boldsymbol{n}}$.

In these notes, I will provide a much simpler derivation of the explicit form for $R(\hat{\boldsymbol{n}}, \theta)$, based on the techniques of tensor algebra.

1. A derivation of the Rodriguez formula

The matrix elements of $R(\hat{\boldsymbol{n}}, \theta)$ will be denoted by R_{ij}. Since $R(\hat{\boldsymbol{n}}, \theta)$ describes a rotation by an angle θ about an axis $\hat{\boldsymbol{n}}$, the formula for R_{ij} that we seek will depend on θ and on the coordinates of $\hat{\boldsymbol{n}} = (n_1 , n_2 , n_3)$ with respect to a fixed Cartesian coordinate system. Note that since $\hat{\boldsymbol{n}}$ is a unit vector, it follows that:

$$n_1^2 + n_2^2 + n_3^2 = 1 . \tag{5}$$

Using the techniques of tensor algebra, we can derive the formula for R_{ij} in the following way. We can regard R_{ij} as the components of a second-rank tensor (see Appendix A). Likewise, the n_i are components of a vector (equivalently, a first-rank tensor). Two other important quantities for the analysis are the *invariant* tensors δ_{ij} (the Kronecker delta) and ϵ_{ijk} (the Levi-Civita tensor). If we invoke the covariance of Cartesian tensor equations, then one must be able to express R_{ij} in terms of a second-rank tensor composed of n_i, δ_{ij} and ϵ_{ijk}, as there are no other tensors in the problem that could provide a source of indices. Thus, the form of the formula for R_{ij} must be:

$$R_{ij} = a\delta_{ij} + bn_i n_j + c\epsilon_{ijk}n_k \,, \tag{6}$$

where there is an implicit sum over the index k in the third term of eq. (6).[1] The numbers a, b and c are real scalar quantities. As such, a, b and c are functions of θ, since the rotation angle is the only relevant scalar quantity in this problem.[2] It also follows that $\hat{\boldsymbol{n}}$ is an axial vector, in which case eq. (6) is covariant with respect to transformations between right-handed and left-handed orthonormal coordinate systems.[3]

We now propose to deduce conditions that are satisfied by a, b and c. The first condition is obtained by noting that

$$R(\hat{\boldsymbol{n}},\theta)\hat{\boldsymbol{n}} = \hat{\boldsymbol{n}} \,.$$

This is clearly true, since $R(\hat{\boldsymbol{n}},\theta)$, when acting on a vector, rotates the vector around the axis $\hat{\boldsymbol{n}}$, whereas any vector parallel to the axis of rotation is invariant under the action of $R(\hat{\boldsymbol{n}},\theta)$. In terms of components

$$R_{ij}n_j = n_i \,. \tag{7}$$

To determine the consequence of this equation, we insert eq. (6) into eq. (7) and make use of eq. (5). Noting that[4]

$$\delta_{ij}n_j = n_i \,, \qquad n_j n_j = 1 \qquad \epsilon_{ijk}n_j n_k = 0 \,, \tag{8}$$

[1]We follow the Einstein summation convention in these notes. That is, there is an implicit sum over any pair of repeated indices in the present and all subsequent formulae.

[2]One can also construct a scalar by taking the dot product of $\hat{\boldsymbol{n}}\cdot\hat{\boldsymbol{n}}$, but this quantity is equal to 1 [cf. eq. (5)], since $\hat{\boldsymbol{n}}$ is a unit vector. Thus, it does not add anything new.

[3]Under inversion of the coordinate system, $\theta \to -\theta$ and $\hat{\boldsymbol{n}} \to -\hat{\boldsymbol{n}}$. However, since $0 \leq \theta \leq \pi$ (by convention), we must then use eq. (2) to flip the signs of both θ and $\hat{\boldsymbol{n}}$ to represent the rotation $R(\hat{\boldsymbol{n}},\theta)$ in the new coordinate system. Hence, the signs of θ and $\hat{\boldsymbol{n}}$ effectively *do not change* under the inversion of the coordinate system. That is, θ is a scalar and $\hat{\boldsymbol{n}}$ is an axial (or pseudo-) vector.

[4]In the third equation of eq. (8), there is an implicit sum over j and k. Since $\epsilon_{ijk} = -\epsilon_{jik}$, when the sum $\epsilon_{ijk}n_j n_k$ is carried out, we find that for every positive term, there is an identical negative term to cancel it. The total sum is therefore equal to zero. This is an example of a very general rule. Namely, one *always* finds that the product of two tensor quantities, one symmetric under the interchange of a pair of summed indices and one antisymmetric under the interchange of a pair of summed indices, is equal to zero when summed over the two indices. In the present case, $n_j n_k$ is symmetric under the interchange of j and k, whereas ϵ_{ijk} is antisymmetric under the interchange of j and k. Hence their product, summed over j and k, is equal to zero.

it follows immediately that $n_i(a+b) = n_i$. Hence,

$$a + b = 1\,. \tag{9}$$

Since the formula for R_{ij} given by eq. (6) must be completely general, it must hold for any special case. In particular, consider the case where $\hat{\boldsymbol{n}} = \boldsymbol{k}$. In this case, eqs. (1) and (6) yields:

$$R(\boldsymbol{k},\theta)_{11} = \cos\theta = a\,, \qquad\qquad R(\boldsymbol{k},\theta)_{12} = -\sin\theta = c\,\epsilon_{123}n_3 = c\,. \tag{10}$$

Using eqs. (9) and (10) we conclude that,

$$a = \cos\theta\,, \qquad\qquad b = 1 - \cos\theta\,, \qquad\qquad c = -\sin\theta\,. \tag{11}$$

Inserting these results into eq. (6) yields the *Rodriguez formula*:

$$\boxed{R_{ij}(\hat{\boldsymbol{n}},\theta) = \cos\theta\,\delta_{ij} + (1-\cos\theta)n_i n_j - \sin\theta\,\epsilon_{ijk}n_k} \tag{12}$$

We can write the above quantity in 3×3 matrix form, although eq. (12) is more compact and convenient. For completeness, here is the explicit form for the general 3×3 rotation matrix:

$$R(\hat{\boldsymbol{n}},\theta) = \begin{pmatrix} \cos\theta + n_1^2(1-\cos\theta) & n_1n_2(1-\cos\theta) - n_3\sin\theta & n_1n_3(1-\cos\theta) + n_2\sin\theta \\ n_1n_2(1-\cos\theta) + n_3\sin\theta & \cos\theta + n_2^2(1-\cos\theta) & n_2n_3(1-\cos\theta) - n_1\sin\theta \\ n_1n_3(1-\cos\theta) - n_2\sin\theta & n_2n_3(1-\cos\theta) + n_1\sin\theta & \cos\theta + n_3^2(1-\cos\theta) \end{pmatrix}, \tag{13}$$

where $n_1^2 + n_2^2 + n_3^2 = 1$. Thus, we have reproduced the explicit form for the general 3×3 rotation matrix, which was derived in the previous handout, *The Matrix Representation of a Three-Dimensional Rotation*.

2. Determining the rotation axis and the rotation angle

In Section 3 of the previous handout, *The Matrix Representation of a Three-Dimensional Rotation*, I presented an algorithm for obtaining the direction of the rotation axis $\hat{\boldsymbol{n}}$ and the rotation angle θ if we are given an arbitrary 3×3 rotation matrix $R(\hat{\boldsymbol{n}},\theta)$. With some tensor algebra manipulations involving the Levi-Civita tensor, we can use eq. (12) to quickly obtain the desired results.

First, we compute the trace of $R(\hat{\boldsymbol{n}},\theta)$. In particular, using eq. (12) it follows that:

$$\boxed{\mathrm{Tr}\ R(\hat{\boldsymbol{n}},\theta) = R_{ii} = 1 + 2\cos\theta} \tag{14}$$

In deriving this result, we used the fact that $\delta_{ii} = \mathrm{Tr}\ \mathbf{I} = 3$ (since the indices run over $i = 1, 2, 3$ in three-dimensional space) and $\epsilon_{iik} = 0$ (the latter is a consequence of the fact that the Levi-Civita tensor is totally antisymmetric under the interchange of any two indices). By convention, $0 \leq \theta \leq \pi$, which implies that $\sin\theta \geq 0$. Thus,

$$\boxed{\cos\theta = \tfrac{1}{2}\left(R_{ii} - 1\right) \quad \text{and} \quad \sin\theta = (1-\cos^2\theta)^{1/2} = \tfrac{1}{2}\sqrt{(3-R_{ii})(1+R_{ii})}} \tag{15}$$

where $\cos\theta$ is determined from eq. (14). All that remains is to determine the axis of rotation $\hat{\boldsymbol{n}}$.

Let us multiply eq. (12) by ϵ_{ijm} and sum over i and j. Noting that[5]

$$\epsilon_{ijm}\delta_{ij} = \epsilon_{ijm}n_i n_j = 0\,, \qquad\qquad \epsilon_{ijk}\epsilon_{ijm} = 2\delta_{km}\,, \tag{16}$$

it follows that

$$2n_m \sin\theta = -R_{ij}\epsilon_{ijm}\,. \tag{17}$$

If R is a symmetric matrix (i.e. $R_{ij} = R_{ji}$), then $R_{ij}\epsilon_{ijm} = 0$ automatically (since ϵ_{ijk} is antisymmetric under the interchange of the indices i and j). In this case $\sin\theta = 0$ and we must seek other means to determine $\hat{\boldsymbol{n}}$. If $\sin\theta \neq 0$, then one can divide both sides of eq. (17) by $\sin\theta$. Using eq. (15), we obtain:[6]

$$\boxed{n_m = -\frac{R_{ij}\epsilon_{ijm}}{2\sin\theta} = \frac{-R_{ij}\epsilon_{ijm}}{\sqrt{(3-R_{ii})(1+R_{ii})}}\,, \qquad \sin\theta \neq 0} \tag{18}$$

More explicitly,

$$\hat{\boldsymbol{n}} = \frac{1}{\sqrt{(3-R_{ii})(1+R_{ii})}}\bigg(R_{32}-R_{23}\,,\,R_{13}-R_{31}\,,\,R_{21}-R_{12}\bigg)\,, \qquad R_{ii} \neq -1\,,\,3\,. \tag{19}$$

In Appendix B, we verify that $\hat{\boldsymbol{n}}$ as given by eq. (18) is a vector of unit length [as required by eq. (5)]. The overall sign of $\hat{\boldsymbol{n}}$ is fixed by eq. (18) due to our convention in which $\sin\theta \geq 0$.

If we multiply eq. (17) by n_m and sum over m, then

$$\sin\theta = -\tfrac{1}{2}\epsilon_{ijm}R_{ij}n_m\,, \tag{20}$$

after using $n_m n_m = 1$. This provides an additional check on the determination of the rotation angle.

As noted above, if R is a symmetric matrix (i.e. $R_{ij} = R_{ji}$), then $\sin\theta = 0$ and $\hat{\boldsymbol{n}}$ cannot be determined from eq. (18). In this case, eq. (14) determines whether $\cos\theta = +1$ or $\cos\theta = -1$. If $\cos\theta = +1$, then $R_{ij} = \delta_{ij}$ and the axis $\hat{\boldsymbol{n}}$ is undefined. If $\cos\theta = -1$, then according to eq. (12), $R_{ij} = 2n_i n_j - \delta_{ij}$, which determines the direction of $\hat{\boldsymbol{n}}$ up to an overall sign. That is,

$$\hat{\boldsymbol{n}} \text{ is undetermined if } \theta = 0\,,$$

$$\hat{\boldsymbol{n}} = \left(\epsilon_1\sqrt{\tfrac{1}{2}(1+R_{11})}\,,\,\epsilon_2\sqrt{\tfrac{1}{2}(1+R_{22})}\,,\,\epsilon_3\sqrt{\tfrac{1}{2}(1+R_{33})}\right)\,, \qquad \text{if } \theta = \pi\,, \tag{21}$$

[5]In regards to the first equation of eq. (16), see the footnote 4. The second equation of eq. (16) is given in eq. (5.8) on p. 511 of Boas.

[6]If $\sin\theta = 0$, both the numerator and denominator of eq. (18) vanish. We will show below [cf. eq. (21)] that $\hat{\boldsymbol{n}}$ is undefined if $R_{ii} = 3$, corresponding to the case of $R(\hat{\boldsymbol{n}}, 0) = \mathbf{I}$. When $R_{ii} = -1$, corresponding to the case of $R(\hat{\boldsymbol{n}}, \pi)$, $\hat{\boldsymbol{n}}$ can be determined directly from eq. (12).

where the individual signs $\epsilon_i = \pm 1$ are determined up to an overall sign via[7]

$$\epsilon_i \epsilon_j = \frac{R_{ij}}{\sqrt{(1+R_{ii})(1+R_{jj})}}\,, \qquad \text{for fixed } i \neq j\,,\; R_{ii} \neq -1\,,\; R_{jj} \neq -1\,. \tag{22}$$

The ambiguity of the overall sign of $\hat{\boldsymbol{n}}$ sign is not significant, since $R(\hat{\boldsymbol{n}}, \pi)$ and $R(-\hat{\boldsymbol{n}}, \pi)$ represent the same rotation [cf. eq. (4)].

One slightly inconvenient feature of the above analysis is that the case of $R_{ii} = -1$ (corresponding to $\theta = \pi$) requires a separate treatment in order to determine $\hat{\boldsymbol{n}}$. Moreover, for values of θ very close to π, the numerator and denominator of eq. (19) are very small, so that a very precise numerical evaluation of both the numerator and denominator is required to accurately determine the direction of $\hat{\boldsymbol{n}}$. Thus, we briefly mention another approach for determining $\hat{\boldsymbol{n}}$ that can be employed for all possible values of R_{ii} (except for $R_{ii} = 3$ corresponding to the identity rotation, where $\hat{\boldsymbol{n}}$ is not defined). This approach is based directly on the Rodriguez formula [eq. (12)]. Define the matrix

$$S = R + R^{\mathsf{T}} + (1 - R_{ii})\mathbf{I}\,. \tag{23}$$

Then, eq. (12) yields $S_{jk} = 2(1 - \cos\theta) n_j n_k = (3 - R_{ii}) n_j n_k$. Hence,[8]

$$\boxed{n_j n_k = \frac{S_{jk}}{3 - R_{ii}}\,, \qquad R_{ii} \neq 3} \tag{24}$$

Note that for θ close to π (which corresponds to $R_{ii} \simeq -1$), neither the numerator nor the denominator of eq. (24) is particularly small, and the direction of $\hat{\boldsymbol{n}}$ can be determined numerically without significant roundoff error.

To determine $\hat{\boldsymbol{n}}$ up to an overall sign, we simply set $j = k$ (no sum) in eq. (24), which fixes the value of n_j^2. If $\sin\theta \neq 0$, the overall sign of $\hat{\boldsymbol{n}}$ is fixed by eq. (17). If $\sin\theta = 0$ there are two cases. For $\theta = 0$ (corresponding to the identity rotation), the rotation axis $\hat{\boldsymbol{n}}$ is undefined. For $\theta = \pi$, the ambiguity in the overall sign of $\hat{\boldsymbol{n}}$ is immaterial, in light of eq. (4).

Thus, we have achieved our goal. Eqs. (15), (19) and (21) [or equivalently, eqs. (15), (17) and (24)] provide a simple algorithm for determining the rotation axis $\hat{\boldsymbol{n}}$ and the rotation angle θ for any rotation matrix $R(\hat{\boldsymbol{n}}, \theta) \neq \mathbf{I}$.

[7]If $R_{ii} = -1$ [no sum over i], then $n_i = 0$, in which case the corresponding ϵ_i is not well-defined.

[8]Eq. (23) yields $\operatorname{Tr} S \equiv S_{ii} = 3 - R_{ii}$. One can then use eq. (24) to verify that $\hat{\boldsymbol{n}}$ is a unit vector.

Appendix A: Matrix elements of matrices correspond to the components of second rank tensors

In the class handout entitled *Coordinates, matrix elements and changes of basis*, we examined how the matrix elements of linear operators change under a change of basis. Consider the matrix elements of a linear operator with respect to two different orthonormal bases, $\mathcal{B} = \{\hat{\boldsymbol{e}}_1\,,\,\hat{\boldsymbol{e}}_2\,,\,\hat{\boldsymbol{e}}_3\}$ and $\mathcal{B}' = \{\hat{\boldsymbol{e}}'_1\,,\,\hat{\boldsymbol{e}}'_2\,,\,\hat{\boldsymbol{e}}'_3\}$. Then, using the Einstein summation convention,

$$\hat{\boldsymbol{e}}'_j = P_{ij}\hat{\boldsymbol{e}}_i\,,$$

where P is an orthogonal matrix. Given any linear operator A with matrix elements a_{ij} with respect to the basis $\mathcal{B}$, the matrix elements a'_{ij} with respect to the basis $\mathcal{B}'$ are given by

$$a'_{k\ell} = (P^{-1})_{ki}a_{ij}P_{j\ell} = P_{ik}a_{ij}P_{j\ell}\,,$$

where we have used the fact that $P^{-1} = P^{\mathsf{T}}$ in the second step above. Finally, identifying $P = R^{-1}$, where R is also an orthogonal matrix, it follows that

$$a'_{k\ell} = R_{ki}R_{\ell j}a_{ij}\,,$$

which we recognize as the transformation law for the components of a second rank Cartesian tensor.

Appendix B: Verifying that $\hat{\boldsymbol{n}}$ as determined from eq. (18) is a unit vector

We first need some preliminary results. Using the results from the handout entitled *The Characteristic Polynomial*, the characteristic equation of an arbitrary 3×3 matrix R is given by:

$$p(\lambda) = -\left[\lambda^3 - \lambda^2\mathrm{Tr}\ R + c_2\lambda^2 - \det\ R\right]\,,$$

where

$$c_2 = \tfrac{1}{2}\left[(\mathrm{Tr}\ R)^2 - \mathrm{Tr}(R^2)\right]\,.$$

For a special orthogonal matrix, $\det R = 1$. Hence,

$$p(\lambda) = -\left[\lambda^3 - \lambda^2\mathrm{Tr}\ R + \tfrac{1}{2}\lambda\left[(\mathrm{Tr}\ R)^2 - \mathrm{Tr}(R^2)\right] - 1\right]\,.$$

We now employ the Cayley-Hamilton theorem, which states that a matrix satisfies its own characteristic equation, i.e. $p(R) = 0$. Hence,

$$R^3 - R^2\,\mathrm{Tr}\ R + \tfrac{1}{2}R\left[(\mathrm{Tr}\ R)^2 - \mathrm{Tr}(R^2)\right] - \mathbf{I} = 0\,.$$

Multiplying the above equation by R^{-1}, and using the fact that $R^{-1} = R^{\mathsf{T}}$ for an orthogonal matrix,

$$R^2 - R\,\mathrm{Tr}\ R + \tfrac{1}{2}\mathbf{I}\left[(\mathrm{Tr}\ R)^2 - \mathrm{Tr}(R^2)\right] - R^{\mathsf{T}} = 0\,.$$

Finally, we take the trace of the above equation. Using $\mathrm{Tr}(R^{\mathsf{T}}) = \mathrm{Tr}\ R$, we can solve for $\mathrm{Tr}(R^2)$. Using $\mathrm{Tr}\ \mathbf{I} = 3$, the end result is given by:

$$\mathrm{Tr}(R^2) = (\mathrm{Tr}\ R)^2 - 2\,\mathrm{Tr}\,R\,, \tag{25}$$

which is satisfied by all 3×3 special orthogonal matrices.

We now verify that $\hat{\boldsymbol{n}}$ as determined from eq. (18) is a unit vector. For convenience, we repeat eq. (18) here:

$$n_m = -\tfrac{1}{2}\frac{R_{ij}\epsilon_{ijm}}{\sin\theta} = \frac{-R_{ij}\epsilon_{ijm}}{\sqrt{(3-R_{ii})(1+R_{ii})}}\,, \qquad \sin\theta \neq 0\,,$$

where $R_{ii} \equiv \mathrm{Tr}\ R$. We evaluate $\hat{\boldsymbol{n}}\cdot\hat{\boldsymbol{n}} = n_m n_m$ as follows:

$$n_m n_m = \frac{R_{ij}\epsilon_{ijm}R_{k\ell}\epsilon_{klm}}{(3-R_{ii})(1+R_{ii})} = \frac{R_{ij}R_{k\ell}(\delta_{ik}\delta_{j\ell} - \delta_{i\ell}\delta_{jk})}{(3-R_{ii})(1+R_{ii})} = \frac{R_{ij}R_{ij} - R_{ij}R_{ji}}{(3-R_{ii})(1+R_{ii})}\,.$$

The numerator of the above expression is equal to:

$$\begin{aligned} R_{ij}R_{ij} - R_{ij}R_{ji} &= \mathrm{Tr}(R^{\mathsf{T}}R) - \mathrm{Tr}(R^2) = \mathrm{Tr}\ \mathbf{I} - \mathrm{Tr}(R^2) \\ &= 3 - \mathrm{Tr}(R^2) = 3 - (\mathrm{Tr}\,R)^2 + 2\,\mathrm{Tr}\,R = (3-R_{ii})(1+R_{ii})\,, \end{aligned} \tag{26}$$

after using eq. (25) for $\mathrm{Tr}(R^2)$. Hence, employing eq. (26) yields

$$\hat{\boldsymbol{n}}\cdot\hat{\boldsymbol{n}} = n_m n_m = \frac{R_{ij}R_{ij} - R_{ij}R_{ji}}{(3-R_{ii})(1+R_{ii})} = 1\,,$$

and the proof is complete.