ORTHOGONAL MATRICES

Informally, an orthogonal $n \times n$ matrix is the n-dimensional analogue of the rotation matrices R_θ in $\mathbb{R}^2$. When does a transformation of $\mathbb{R}^3$ (or $\mathbb{R}^n$) deserve to be called a rotation? We expect it to fix the origin, and to have the property that applying the rotation to the sum of two vectors is the same as adding the rotated vectors. Most important, rotations are 'rigid motions', in the geometric sense of preserving the lengths of vectors and the angle between vectors.

An efficient way to get the first two properties is to require the transformation to be *linear.* Consider now lengths and angles. The cosine of the angle formed by two vectors is given by their inner product (or 'dot product') divided by the products of their lengths. Thus if our linear transformation preserves lengths of vectors and also the inner product of two vectors, it will automatically be a 'rigid motion'. In fact, preservation of inner products already implies preservation of lengths, since $||v||^2 = \langle v, v \rangle$, for any v.

Thus we want our rotation matrices to satisfy:

$$\langle Av, Aw \rangle = \langle v, w \rangle, \forall v, w \in \mathbb{R}^n.$$

Recall the basic property of the transpose (for *any* A):

$$\langle Av, w \rangle = \langle v, A^T w \rangle, \forall v, w \in \mathbb{R}^n.$$

It implies that requiring A to have the property above is the same as requiring:

$$\langle v, A^T A w \rangle = \langle v, w \rangle \forall v, w \in \mathbb{R}^n.$$

The only way this is possible is if $A^T A$ is the identity matrix I_n. This means the following is a reasonable definition of 'rotation matrix':

Definition. An $n \times n$ matrix is *orthogonal* if $A^T A = I_n$.

Basic properties. (1) A matrix is orthogonal exactly when its column vectors have length one, and are pairwise orthogonal; likewise for the column vectors. In short, the columns (or the rows) of an orthogonal matrix are an *orthonormal basis* of $\mathbb{R}^n$, and any orthonormal basis gives rise to a number of orthogonal matrices.

(2) Any orthogonal matrix is invertible, with $A^{-1} = A^T$. If A is orthogonal, so are A^T and A^{-1}.

(3) The product of orthogonal matrices is orthogonal: if $A^T A = I_n$ and $B^T B = I_n$,

$$(AB)^T(AB) = (B^T A^T)AB = B^T(A^T A)B = B^T B = I_n.$$

(4) The 2×2 rotation matrices R_θ are orthogonal. Recall:

$$R_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$

(R_θ rotates vectors by θ radians, counterclockwise.)

(5) The determinant of an orthogonal matrix is equal to 1 or -1. The reason is that, since $det(A) = det(A^T)$ for any A, and the determinant of the product is the product of the determinants, we have, for A orthogonal:

$$1 = det(I_n) = det(A^T A) = det(A^{(T)})det(A) = (detA)^2.$$

(6) Any real eigenvalue of an orthogonal matrix has absolute value 1. To see this, consider that $||Rv|| = ||v||$ for any v, if R is orthogonal. But if $v \neq 0$ is an eigenvector with eigenvalue λ:

$$Rv = \lambda v \quad \Rightarrow ||v|| = ||Rv|| = |\lambda|||v||,$$

hence $|\lambda| = 1$. (Actually, it is also true that each complex eigenvalue must have modulus 1, and the argument is similar).

All seems to be working, except for a slight problem. *Reflections* have all the properties we listed earlier for rotations- they preserve lengths of vectors and angles between vectors. So we should expect some orthogonal matrices to represent reflections (about a line through the origin in $\mathbb{R}^2$, or a plane through the origin in $\mathbb{R}^3$.) How can we distinguish between rotations and reflections? First, we have to agree on what is meant by 'reflection':

A linear transformation T of $\mathbb{R}^n$ is a *reflection* if there is a one-dimensional subspace L (a line through 0) so that $Tv = -v$ for $v \in L$ and $Tv = v$ for v in the orthogonal complement $L^\perp$.

This agrees with the usual geometric notion of reflection in $\mathbb{R}^2$ and $\mathbb{R}^3$. Like rotations, reflections are rigid motions of space. *How are they different?* (Not at all a trivial question!) The following distinction is easy to work with: a rotation R is really the end result of a 'continuous motion', where we start by 'doing nothing' (the identity map), and gradually move all vectors by rotations $R(t)$ (where t denotes 'time'), until we arrive at the desired rotation $R = R(1)$. Since $det(I_n) = 1$ and $det(R(t))$ depends continuously

on t (and is always 1 or -1), we must have $\det R = 1$ if R is a rotation. On the other hand, for a reflection (as defined above) if we take an orthonormal basis $\mathcal{B}$ of $\mathbb{R}^n$ with the first vector in L, the matrix of the reflection T in this basis will be (say, if $n = 3$):

$$[T]_\mathcal{B} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

This clearly has determinant -1. This discussion justifies the following definition:

Algebraic definition. An orthogonal transformation of $\mathbb{R}^n$ is a *rotation* if it has determinant 1, a *reflection* if it has determinant -1.

Examples.

1. *Reflection on a given plane in $\mathbb{R}^3$.* Find the orthogonal matrix (in the standard basis) that implements reflection on the plane with equation:

$$2x_1 + 3x_2 + x_3 = 0$$

Solution. The orthogonal line L is spanned by the unit vector:

$$n = \frac{1}{\sqrt{14}}(2, 3, 1).$$

Any vector $v \in \mathbb{R}^3$ can be written as the sum of its orthogonal projections on L and on the given plane $L^\perp$:

$$v = \langle v, n \rangle n + w, \quad w = v - \langle v, n \rangle n \in L^\perp.$$

The reflection T flips n ($Tn = -n$) and fixes every $w \in L^\perp$ (that is, $Tw = w$). By linearity of T:

$$\begin{aligned} Tv &= \langle v, n \rangle Tn + Tw \\ &= \langle v, n \rangle(-n) + w \\ &= -\langle v, n \rangle n + (v - \langle v, n \rangle n) \\ &= v - 2\langle v, n \rangle \end{aligned}.$$

(Incidentally, this calculation gives a general formula for the reflection on a plane with unit normal n.) In the present case, this reads:

$$Tv = v - 2\frac{1}{14}\langle v, (2, 1, 3) \rangle (2, 1, 3),$$

3

in particular:

$$\begin{array}{rcl} Te_1 &=& e_1 - (4/14)(2,1,3) = (3/7, -2/7, -6/7) \\ Te_2 &=& e_2 - (2/14)(2,1,3) = (-2/7, 6/7, -3/7) \\ Te_3 &=& e_3 - (6/14)(2,1,3) = (-6/7, -3/7, -2/7) \end{array}.$$

These are the columns of the matrix of T in the standard basis:

$$[T]_{\mathcal{B}_0} = \frac{1}{7} \begin{bmatrix} 3 & -2 & -6 \\ -2 & 6 & -3 \\ -6 & -3 & -2 \end{bmatrix}.$$

2. *Rotation in $\mathbb{R}^3$ with given axis and given angle.* Find the 3×3 orthogonal matrix that implements the rotation R in $\mathbb{R}^3$ with axis the subspace spanned by $(1, 2, 3)$ (a line L), by an angle of $\pi/6$ radians (counterclockwise when looking down the axis).

Solution. The first step is to find a positive orthonormal basis of $\mathbb{R}^3$, $\mathcal{B} = \{u, v_1, v_2\}$, where u is a unit vector spanning L and v_1, v_2 are an orthonormal basis for the orthogonal plane $L^\perp$. 'Positive' means $det[u|v_1|v_2] = 1$. We take $u = (1/\sqrt{14})(1, 2, 3)$. Start with any basis of of $L^\perp$, say (by 'trial and error': take any two l.i. vectors satisfying the equation $x_1 + 2x_2 + 3x_3 = 0$ of $L^\perp$):

$$w_1 = (3, 0, -1), \quad w_2 = (0, 3, -2).$$

Since:

$$det \begin{bmatrix} 1 & 3 & 0 \\ 2 & 0 & 3 \\ 3 & -1 & -2 \end{bmatrix} = 33 > 0,$$

the basis $\{u, w_1, w_2\}$ of $\mathbb{R}^3$ is positive. Applying Gram-Schmidt to w_1, w_2, we obtain an orthonormal basis of the plane $L^\perp$:

$$v_1 = (1/\sqrt{10})(3, 0, -1), \quad v_2 = (1/\sqrt{35})(-1, 5, -3).$$

(The basis $\{u, v_1, v_2\}$ is still positive.)

Second step: computing the action of R on the basis vectors. Since u is a vector on the axis, it is fixed by R: $Ru = u$. In the plane $L^\perp$, R acts exactly like the counterclockwise rotation by $\pi/6$ radians in $\mathbb{R}^2$, so we know what R does to an orthonormal basis. We obtain:

$$\begin{array}{rcl} Ru &=& (1/\sqrt{14})(1, 2, 3) \\ Rv_1 &=& \cos(\pi/6)v_1 + \sin(\pi/6)v_2 = (\sqrt{3}/2)v_1 + (1/2)v_2 \\ Rv_2 &=& \sin(\pi/6)v_1 + \cos(\pi/6)v_2 = (-1/2)v_1 + (\sqrt{3}/2)v_2. \end{array}$$

This means that the matrix of R in the 'adapted' basis $\mathcal{B}$ is:

$$[R]_\mathcal{B} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \sqrt{3}/2 & -1/2 \\ 0 & 1/2 & \sqrt{3}/2 \end{bmatrix}.$$

This is orthogonal; in fact, in block form, it has a 1 (corresponding to the axis u, an eigenvector of R with eigenvalue 1) and a 2×2 block given by the rotation $R_{\pi/6}$ in the plane.

Third step. Now use the change of basis formula. We have:

$$[R]_{\mathcal{B}_0} = P[R]_\mathcal{B}P^{-1} = P[R]_\mathcal{B}P^T,$$

since P is orthogonal. The change of basis matrix P is given by the calculation in the first step:

$$P = \begin{bmatrix} 1/\sqrt{14} & 3/\sqrt{10} & -1/\sqrt{35} \\ 2/\sqrt{14} & 0 & 5/\sqrt{35} \\ 3/\sqrt{14} & -1/\sqrt{10} & -3/\sqrt{35} \end{bmatrix}.$$

And that's it. Anyone so inclined can go ahead and multiply out the matrices, but I don't see the point in doing this.

Remark on eigenvalues. As remarked above, any vector on the axis of a rotation in $\mathbb{R}^3$ is an eigenvector for the eigenvalue $\lambda = 1$. In fact, the eigenspace $E(1)$ *is* the axis. Is it possible to see algebraically that a rotation A in $\mathbb{R}^3$ (as defined above: an orthogonal matrix with determinant 1) has 1 as an eigenvalue? Well, since the determinant is 1, the characteristic equation has the form:

$$\lambda^3 + a_2\lambda^2 + a_1\lambda + 1 = 0.$$

A polynomial of degree three has at least one real root, and each real root is either 1 or -1 (since A is orthogonal- property (6) above). If all the roots are real, at least one must be 1 (since their product is the determinant, namely 1). If there are two complex conjugate roots $(a \pm ib)$ and the real root is -1, the product of the three roots is $(a+ib)(a-ib)(-1) = -(a^2+b^2)$, which can't possibly be equal to 1. This shows that at least one of the roots is 1, pictorially: *every rotation in $\mathbb{R}^3$ has an axis.* From the *geometric* construction of a rotation, we see that, in fact, the other two eigenvalues must be complex (since in the plane orthogonal to the axis there are no fixed directions, hence no real eigenvectors.)

What about reflections? Now the product of the roots (=the determinant) is -1, so if all three are real, one must be -1 (and the other 2 roots are 1). If there is a complex conjugate pair of roots $a \pm ib$, their product is $a^2 + b^2$, which is positive, so the only way the product of all three can be -1 is if the real root is -1; thus a reflection in $\mathbb{R}^3$ always has -1 as an eigenvalue: *every reflection flips a direction* (defining reflections algebraically, as above.) Actually, from the *geometric* construction of a general reflection it is easy to see that the other two eigenvalues are 1, since every vector in the plane orthogonal to the 'flipped direction' is *fixed* by the reflection (hence is an eigenvector with eigenvalue 1).