ROTATIONS IN $\mathbb{R}^3$

TIM SMITS

1. Introduction

It is a standard example in linear algebra courses that matrices of the form

$$R_\theta = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix}$$

correspond to a linear transformation that rotates a vector $v \in \mathbb{R}^2$ counterclockwise around the origin by an angle of $\theta \in [0, 2\pi)$.

It is evident that all such matrices R_θ are orthogonal of determinant 1, and conversely, a simple algebraic computation shows that all orthogonal 2×2 matrices of determinant 1 can be written as R_θ for some θ. Therefore in $\mathbb{R}^2$, there is no difference between rotations around the origin and determinant 1 orthogonal matrices.

The purpose of this handout is to prove an analogous result in $\mathbb{R}^3$.

2. Rotations in $\mathbb{R}^3$

Definition 2.1. The *real special orthogonal group of dimension* 3 is denoted as $\mathrm{SO}_3(\mathbb{R})$ and defined by

$$\mathrm{SO}_3(\mathbb{R}) = \{A \in M_3(\mathbb{R}) : A \text{ is orthogonal and } \det(A) = 1\}.$$

Our goal is to prove the following theorem:

Theorem 2.2. *Let $A \in SO_3(\mathbb{R})$. There there is $\theta \in [0, 2\pi)$ such that*

$$A \sim \begin{pmatrix} \cos(\theta) & \sin(\theta) & 0 \\ -\sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Before we prove this, we'd first like to comment on what this is actually saying. Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be the linear transformation associated to A. The above theorem says there is a basis β of $\mathbb{R}^3$ such that

$$[T]_\beta = \begin{pmatrix} \cos(\theta) & \sin(\theta) & 0 \\ -\sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

If we let $\beta = \{v_1, v_2, v_3\}$, then the vectors v_1, v_2, v_3 form orthonormal coordinate axes for a coordinate system. Geometrically, T is a rotation in the v_1v_2-plane around the v_3-axis in the counterclockwise direction (relative to v_3). Hence, all elements of $\mathrm{SO}_3(\mathbb{R})$ represent a rotation around some axis, and so it makes sense to talk about $\mathrm{SO}_3(\mathbb{R})$ as the set of rotations in $\mathbb{R}^3$.

The proof of theorem 2.2 is rather elementary, but is interesting because of it's synthesis of all the standard topics in a first year linear algebra course. First, we show that any matrix in $\mathrm{SO}_3(\mathbb{R})$ fixes some line.

Lemma 2.3. *Let $A \in SO_3(\mathbb{R})$. Then 1 is an eigenvalue of A.*

Proof. Writing $p_A(x) = \det(A - xI)$, then we see that $p_A(x)$ may be written as

$$p_A(x) = -x^3 + \mathrm{Tr}(A)x^2 - cx + \det(A)$$

for some $c \in \mathbb{R}$. Since $\det(A) = 1$, this means

$$p_A(x) = -x^3 + \mathrm{Tr}(A)x^2 - cx + 1$$

for some $c \in \mathbb{R}$. As $p_A(0) = 1$ and $\lim_{x\to\infty} p_A(x) = -\infty$, this means there is some $\lambda \in (0, \infty)$ with $p_A(\lambda) = 0$. As A is orthogonal, the only possible real eigenvalues of A are ± 1, and therefore this forces $\lambda = 1$ as desired. □

Next, we need a general fact about linear independence of eigenvectors corresponding to distinct eigenvalues. We prove this in the special case when $A \in M_3(\mathbb{R})$, although we remark this is true regardless of the dimension and can be found in any linear algebra textbook.

Lemma 2.4. *Suppose that $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{C}$ are distinct eigenvalues of a matrix $A \in M_3(\mathbb{R})$. Then A is diagonalizable over $\mathbb{C}$.*

Proof. Let $v_1, v_2, v_3 \in \mathbb{C}^3$ denote eigenvectors for $\lambda_1, \lambda_2, \lambda_3$ respectively. We first show that v_1, v_2 are linearly independent. Suppose that

$$c_1 v_1 + c_2 v_2 = 0$$

for some $c_i \in \mathbb{C}$. Applying A to both sides, we obtain

$$c_1\lambda_1 v_1 + c_2\lambda_2 v_2 = 0.$$

Similarly, multiplying the first equation by λ_1 we obtain

$$c_1\lambda_1 v_1 + c_2\lambda_1 v_2 = 0.$$

Subtracting the two equations, we find

$$c_2(\lambda_2 - \lambda_1)v_2 = 0.$$

This then means that $c_2 = 0$, and therefore that $c_1 = 0$ as well, which shows that v_1, v_2 are linearly independent.

Now, suppose that

$$c_1 v_1 + c_2 v_2 + c_3 v_3 = 0$$

for some $c_i \in \mathbb{C}$.

As before, applying A, multiplying by λ_3, and subtracting, we find

$$c_1(\lambda_3 - \lambda_1)v_1 + c_2(\lambda_3 - \lambda_2)v_2 = 0.$$

Both these coefficients must be 0 because v_1, v_2 are linearly independent, and therefore $c_1 = c_2 = c_3 = 0$ as desired.

□

We're now ready to begin the proof of theorem 2.2.

Proof of Theorem 2.2. Let $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{C}$ be the eigenvalues of A. By lemma 2.3, we know $\lambda_1 = 1$ is an eigenvalue, and because $\det(A) = \lambda_1\lambda_2\lambda_3 = 1$ and complex roots come in conjugate pairs, this means $\lambda_3 = \overline{\lambda_2}$. For brevity, we write $\lambda_2 := \lambda$. Therefore, we see that $|\lambda| = 1$, and so writing λ in polar coordinates, there is $\theta \in [0, 2\pi)$ such that $\lambda = e^{i\theta}$.

We first deal with the case where λ is not real. Let $v \in \mathbb{C}^3$ be an eigenvector for λ. Write $v = u + iw$ for $u, w \in \mathbb{R}^3$ as a sum of its real and imaginary parts. Writing out $\lambda = \cos(\theta) + i\sin(\theta)$, and using $Av = \lambda v$, we therefore find

$$Au = \cos(\theta)u - \sin(\theta)w$$

and

$$Aw = \sin(\theta)u + \cos(\theta)w.$$

Let x be an eigenvector for 1. The claim is that $\beta = \{u, w, x\}$ is a basis of $\mathbb{R}^3$, from which the theorem then follows from the change of basis formula. To that end, note that

$$\overline{v} = u - iw$$

is an eigenvector for A of eigenvalue $\overline{\lambda}$. Suppose for some $c_1, c_2, c_3 \in \mathbb{R}$ that

$$c_1x + c_2u + c_3w = 0.$$

Writing $u = \frac{1}{2}(v + \overline{v})$ and $w = \frac{1}{2}(v - \overline{v})$, this means that

$$c_1x + (\frac{1}{2}c_2 + \frac{1}{2}c_3)v + (\frac{1}{2}c_2 - \frac{1}{2}c_3)\overline{v} = 0.$$

By lemma 2.4, the set of vectors $\{x, v, \overline{v}\}$ is linearly independent because they are eigenvectors for distinct eigenvalues. Therefore, all coefficients in the previous linear combination must be 0, which means $c_1 = c_2 = c_3 = 0$ as desired.

It remains to see what happens when λ is real – surprisingly, this is much harder. In this case, as noted in lemma 2.3, the only possibilities for λ are $\lambda = \pm 1$. We'll first start with the case $\lambda = -1$, as this is easier. In this situation, we have two different eigenvalues, 1 and -1, of algebraic multiplicities 1 and 2 respectively. We will show that $\dim(E_{-1}) = 2$, which would mean that A is diagonalizable, and therefore has the desired form (taking $\theta = \pi$).

Let v be an eigenvector for 1 and let w be an eigenvector for -1. Let $V = \mathrm{Span}\{v, w\}$, and note that $\dim(V) = 2$ by the argument of lemma 2.4. For any $u \in V^\perp$, we note that $Au \in V^\perp$ too: this is because by orthogonality of A, we must have

$$Au \cdot v = Au \cdot Av = u \cdot v = 0$$

and

$$Au \cdot w = Au \cdot A(-w) = u \cdot (-w) = 0.$$

Now, because $\dim(V^\perp) = 1$, this means that $Au = cu$ for some $c \in \mathbb{R}$, and therefore u is an eigenvector of A. Certainly $u \notin E_1$ because $\dim(E_1) = 1$, and therefore this means $u \in E_{-1}$. As $u \cdot w = 0$, $\{u, w\}$ is linearly independent, which then forces $\dim(E_{-1}) = 2$ as desired.

The only remaining case is $\lambda = 1$. Here, there is a single eigenvalue of algebraic multiplicity 3.

Our goal is to show that $E_1^\perp = \{0\}$. It then follows that $E_1 = \mathbb{R}^3$, and therefore this forces $A = I$, so A has the desired form. To see this, let $v \in E_1$ and $w \in E_1^\perp$. By definition, this means $w \cdot v = 0$, and because A is orthogonal,

$$Aw \cdot Av = Aw \cdot v = 0,$$

which means $Aw \in E_1^\perp$ as well. Now, consider the vectors w, Aw. If $Aw = w$, then $w \in E_1$, so $w \in E_1 \cap E_1^\perp = \{0\}$ means $w = 0$ and we're done. Therefore, we may assume that $Aw \neq w$. It then follows that $\{w, Aw\}$ must be linearly independent (we said the only eigenvalue of A is 1, so these can never be scalar multiples of each other!). Therefore, this means $\beta = \{v, w, Aw\}$ is a basis of $\mathbb{R}^3$, and in particular, $\{w, Aw\}$ forms a basis of $E_1^\perp$.

Now, as $A^2 w \in E_1^\perp$, we may write

$$A^2 w = c_1 w + c_2 Aw$$

for some $c_1, c_2 \in \mathbb{R}$, and therefore we find that

$$(A^2 - c_2 A - c_1 I)w = 0.$$

Let r_1, r_2 denote the roots of the polynomial $x^2 - c_2 x - c_1$. Then factoring, this means

$$(A - r_2 I)(A - r_1 I)w = (A - r_1 I)(A - r_2 I)w = 0.$$

If one of $r_1, r_2 \neq 1$, then one of $\det(A - r_1 I), \det(A - r_2 I) \neq 0$ (because the only eigenvalue is assumed to be 1!), and therefore the corresponding matrix is invertible over $\mathbb{C}$. This then means either

$$(A - r_1 I)w = 0$$

or

$$(A - r_2 I)w = 0,$$

which means that w is an eigenvector. This would then force $w \in E_1 \cap E_1^\perp = \{0\}$, which contradicts our assumption that $Aw \neq w$. Therefore, we may assume that $r_1 = r_2 = 1$. In this case, we must have

$$(A - I)^2 w = 0.$$

However, this means $(A - I)(Aw - w) = 0$, and therefore $Aw - w \in E_1$. This contradicts that $\{v, w, Aw\}$ is a basis, and so the only possibility is that $E_1^\perp = \{0\}$ as desired.

□

Example 2.5. Consider

$$A = \begin{pmatrix} 1/3 & 2/3 & -2/3 \\ 2/3 & 1/3 & 2/3 \\ 2/3 & -2/3 & -1/3 \end{pmatrix},$$

which is easily seen to be orthogonal and has determinant 1. The characteristic polynomial of A is given by

$$p_A(x) = -x^3 + \frac{1}{3}(x^2 - x) + 1,$$

and so the eigenvalues of A are

$$1, \frac{1}{3}(-1 \pm 2i\sqrt{2}).$$

We take

$$\lambda = \frac{1}{3}(-1 + 2i\sqrt{2}) = -\frac{1}{3} + \frac{2\sqrt{2}}{3}i = \cos(\theta) + i\sin(\theta).$$

One may compute that an eigenvector for λ is given by

$$v = u + iw = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + i \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \\ 0 \end{pmatrix},$$

and an eigenvector for 1 is

$$x = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}.$$

Then

$$\beta = \left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \right\}$$

is a basis of $\mathbb{R}^3$, and relative to this basis the matrix of the linear transformation $T(x) = Ax$ is

$$[T]_\beta = \begin{pmatrix} -1/3 & 2\sqrt{2}/3 & 0 \\ -2\sqrt{2}/3 & -1/3 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

The axis of rotation is the the line spanned by

$$\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix},$$

and the angle of rotation is $\theta = \pi - \arctan(2\sqrt{2}) \approx 1.91$.

Remark 2.6. In dimension n, we write $\mathrm{SO}_n(\mathbb{R})$ to mean the real $n \times n$ orthogonal matrices of determinant 1. For $n > 3$ odd, the same argument as in lemma 2.3 shows that A must have 1 as an eigenvalue, and therefore must fix a line. However, this is *not* true in general! When $n = 4$, the matrix

$$A = \begin{pmatrix} 1/2 & \sqrt{3}/2 & 0 & 0 \\ -\sqrt{3}/2 & 1/2 & 0 & 0 \\ 0 & 0 & 1/\sqrt{2} & 1/\sqrt{2} \\ 0 & 0 & -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \in \mathrm{SO}_4(\mathbb{R})$$

does not have 1 as an eigenvalue, and therefore does not fix any line!