12 Orthogonal Matrices

In this lecture, we start formally studying the symmetry of shapes, combining group theory with linear algebra. The matrices considered will be over $\mathbb{R}$, the field of real numbers, rather than $\mathbb{C}$.

12.1 Dot Products and Orthogonal Matrices

Recall the following definitions.

Definition 12.1
Given column vectors $x, y \in \mathbb{R}^n$, the **dot product** is defined as $x \cdot y = x^T y = \sum_{i=1}^n x_i y_i$. The **length** of a vector v is $|v| = \sqrt{v \cdot v}$.

The dot product is defined algebraically, but also carries geometric information about two vectors:

$$x \cdot y = |x||y| \cos\theta.$$

Moreover, if $x \cdot y = 0$, then x and y will be perpendicular vectors in $\mathbb{R}^n$.

To start out with, consider bases for which the pairwise dot products are as simple as possible.

Definition 12.2
A basis $\{v_1, \cdots, v_n\}$ is called *orthonormal* if $|v_i| = 1$ and $v_i \cdot v_j = 0$ for $i \neq j$. That is, since $|v_i| = \sqrt{v_i \cdot v_i}$,

$$v_i \cdot v_j = \delta_{ij},$$

where δ_{ij} denotes the Kronecker delta.[a]

[a] The Kronecker delta δ_{ij} is equal to 0 if $i \neq j$ and 1 if $i = j$.

Now, $\mathbb{R}^n$ not only has a vector space structure, but it also has some extra structure provided by the dot product. Since $|v| = \sqrt{v \cdot v}$, the dot product produces some notion of "length" or "distance."

Guiding Question
What kinds of matrices interact well with this notion of distance?

Orthogonal matrices are those preserving the dot product.

Definition 12.3
A matrix $A \in GL_n(\mathbb{R})$ is *orthogonal* if $Av \cdot Aw = v \cdot w$ for all vectors v and w.

In particular, taking $v = w$ means that lengths are preserved by orthogonal matrices. There are many equivalent characterizations for orthogonal matrices.

Theorem 12.4
The following conditions are all equivalent:

1. The matrix A is orthogonal.
2. For all vectors $v \in \mathbb{R}^n$, $|Av| = |v|$. That is, A preserves lengths.
3. For an n-dimensional matrix A, $A^T A = I_n$.
4. The columns of A form an orthonormal basis.[a]

[a] Since A^T also satisfies the third condition, this means that the rows of A, which are the columns of A^T, will also form an orthonormal basis.

Proof. All the conditions will end up equivalent.

- Condition (1) implies (2). Because A preserves dot products, $|Av| = \sqrt{Av \cdot Av} = \sqrt{v \cdot v} = |v|$, and so A also preserves lengths.
- Condition (2) implies (1) because

$$\begin{aligned} Av \cdot Aw &= \frac{1}{2}\left(|Av + Aw|^2 - |Av|^2 - |Aw|^2\right) \\ &= \frac{1}{2}\left(|v + w|^2 - |v|^2 - |w|^2\right) \\ &= v \cdot w. \end{aligned}$$

  Namely, dot products can be written in terms of lengths, and lengths can be written in terms of dot products, so preserving one is equivalent to preserving the other.
- Condition (1) states that $Av \cdot Aw = v \cdot w$; unwinding the dot product in terms of matrix multiplication, this equation is $v^T A^T Aw = v^T w$ for all $v, w \in \mathbb{R}^n$. Evidently, (3) implies (1), since if $A^T A = I_n$, $v^T A^T Aw = v^T w$.

  By calculation, it can be seen that for e_i and e_j the ith and jth standard basis vectors, $e_i^T M e_j = M_{ij}$, which is the (i, j)th component of the matrix M. If (1) is true, taking $v = e_i$ and $w = e_j$ over all i and j gives us that the (i, j)th component of $A^T A$ is 1 when $i = j$ and 0 otherwise.
- Condition (4) is equivalent to (3) from simply computing the matrix product: the (i, j)th entry of $A^T A$ is the dot product of the ith column of A with the jth column of A, which is 1 when $i = j$ and 0 otherwise.

□

Orthogonal matrices preserve lengths, as well as preserving angles up to sign. In general, a set of matrices satisfying some well-behaved properties of a set of matrices generally form a subgroup, and this principle does hold true in the case of orthogonal matrices.

Proposition 12.5
The orthogonal matrices form a subgroup O_n of GL_n.

Proof. Using condition (3), if for two orthogonal matrices A and B, $A^T A = B^T B = I_n$, it is clear that $(AB)^T AB = B^T A^T AB = B^T B = I_n$. The other subgroup properties are not difficult to verify. □

12.2 The Special Orthogonal Group

Given an orthogonal matrix A, $A^T A = I_n$, and so $\det(A^T A) = \det(A^T)\det(A) = \det(A)^2 = \det(I_n) = 1$. As a result, $\det(A) = \pm 1$. The determinant is a homomorphism from $\det : GL_n \longrightarrow \mathbb{R}$, and the restriction to O_n is a homomorphism $\det : O_n \longrightarrow \{\pm 1\}$. The kernel forms a subgroup of O_n.

Definition 12.6 (Special Orthogonal Group)
The orthogonal matrices with determinant 1 form a subgroup $SO_n \subset O_n \subset GL_n$ called the **special orthogonal group**.

Because the determinant is surjective[41], the kernel, SO_n, is an index 2 subgroup inside of O_n. The two cosets are SO_n itself and all the matrices with determinant -1.

To gain some intuition for orthogonal matrices, we will look at some examples! For $n = 1$, the orthogonal group has two elements, $[1]$ and $[-1]$, which is not too interesting.

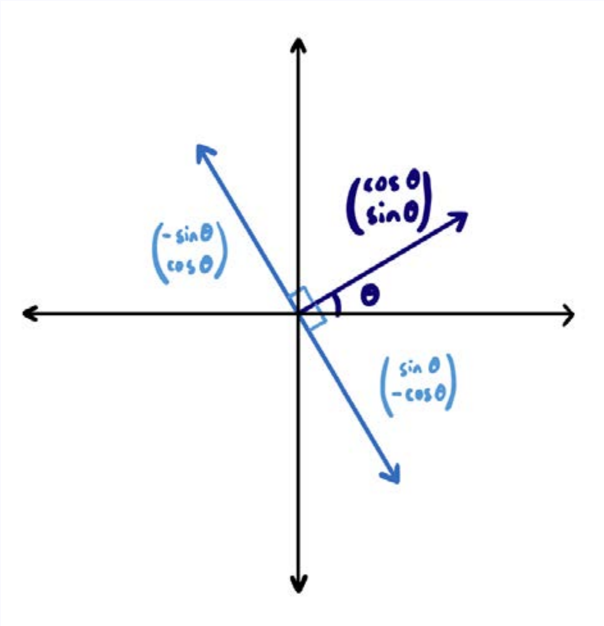
12.3 Orthogonal Matrices in Two Dimensions

What are the orthogonal matrices in two dimensions?

[41] For example, the identity matrix is always orthogonal and has determinant 1, and the diagonal matrix with -1 in the first row and column and 1 down the rest of the diagonal is also orthogonal and has determinant -1.

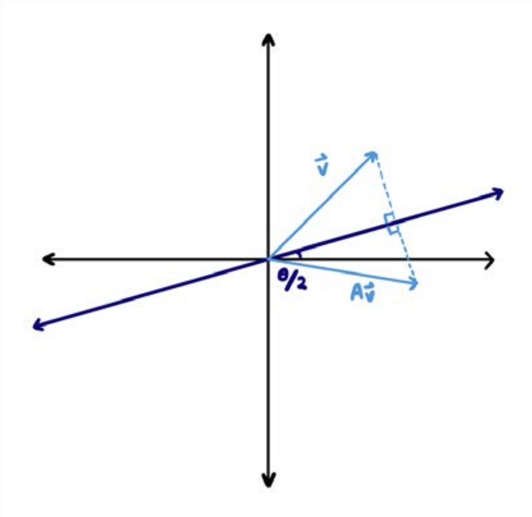
Example 12.7 (O_2)
Describing an element of O_2 is equivalent to writing down an orthonormal basis $\{v_1, v_2\}$ of $\mathbb{R}^2$. Evidently, v_1 must be a unit vector, which can always be described as $v_1 = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}$ for some angle θ. Then v_2 must also have length 1 and be perpendicular to v_1. There are two choices, $v_2 = \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix}$ or $\begin{pmatrix} \sin\theta \\ -\cos\theta \end{pmatrix}$. This characterizes all 2×2 orthogonal matrices:

$$O_2 = \left\{ \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix} \right\}.$$



In particular, the first type of matrix has determinant 1, and forms the subgroup SO_n, and the second has determinant -1 and forms its the non-trivial coset. Geometrically, the first type of matrix in O_2 are rotations by θ around the origin. The matrices of the second type, $A = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}$, have characteristic polynomial $p_A(t) = t^2 - 1 = (t+1)(t-1)$. Thus, they have distinct eigenvalues ± 1, in contrast to rotation matrices, which do not have any real eigenvalues. Because the eigenvalues are distinct, there is an eigenbasis $\{\vec{v_+}, \vec{v_-}\}$.

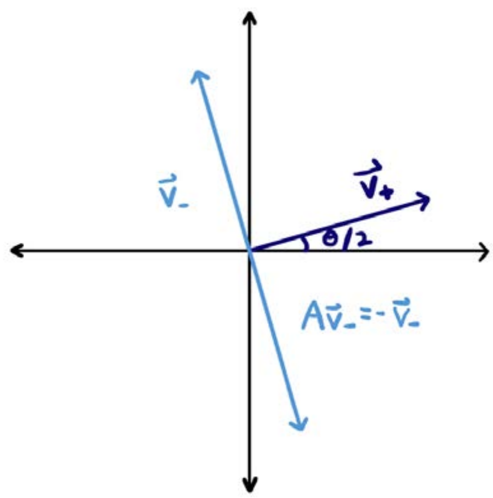
Theorem 12.8
The matrices of the second type are reflections across a line through the origin at an angle of $\theta/2$.



Proof. Consider the line $L = \mathrm{Span}(\vec{v_+})$; since $\vec{v_+}$ is an eigenvector with eigenvalue 1, A fixes this line. Notice that

$$\vec{v_+} \cdot \vec{v_-} = A\vec{v_+} \cdot A\vec{v_-} = \vec{v_+} \cdot (-\vec{v_-}),$$

where the first equality comes from the fact that A is orthogonal, and the second comes from the eigenvalues 1 and -1 of v_+ and v_-. The only possibility is $\vec{v_+} \cdot \vec{v_-} = 0$, so the two eigenvectors are orthogonal. Writing out any other vector in terms of the eigenvectors, Av is precisely the reflection across L.



□

As expected, rotations and reflections preserve distance, and in fact they make up all the 2×2 orthogonal matrices. A fun fact that comes from this analysis is that the composition of two reflections over different lines will be a rotation, since the product of determinants will be $(-1)\cdot(-1) = 1$. Orthogonal matrices can be thought of either geometrically or algebraically!

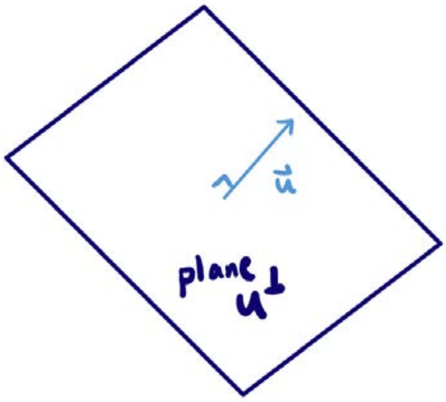
12.4 Orthogonal Matrices in Three Dimensions

In two dimensions, SO_2 consists of rotation matrices. It turns out that in three dimensions, SO_3 also consists of rotation matrices.

In particular, a rotation in $\mathbb{R}^3$ is characterized by the axis of the rotation, which is a unit vector $\vec{u} \in \mathbb{R}^3$, and the angle of the rotation, which is some $\theta \in \mathbb{R}$. The plane

$$u^{\perp} = \{v \in \mathbb{R}^3 : u \cdot v = 0\}$$

consists of all the vectors in $\mathbb{R}^3$ that are perpendicular to $\mathbb{R}^3$.



Definition 12.9
The **rotation operator** with **spin labels** u and θ is $\rho_{(u,\theta)}$, the linear operator $\rho : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ such that $\rho(u) = u$ and $\rho|_{u^\perp}$ is the rotation by θ counterclockwise with respect to the direction that u points in.[a]

[a]Since every vector in $\mathbb{R}^3$ is a linear combination of u and some vector in $u^\perp$, the rotation operator is described completely by these conditions.

There is some redundancy in this description; for example, $\rho_{(u,\theta)} = \rho_{(-u,-\theta)}$.

Theorem 12.10
The rotation operators are exactly SO_3.

From geometric intuition, this result is not very surprising, since rotations preserve distance.[42]

Proof. First, we show that all the rotation matrices are in SO_3, and then we show that all matrices in SO_3 are rotation matrices.

- We first show that all of these rotation matrices belong to SO_3. Let $\{v, w\}$ be an orthonormal basis for the plane $u^\perp$, and let P be a 3×3 matrix with columns (u, v, w). Since v and w are orthogonal to each other, and u is orthogonal to both v and w, $P \in O_3$. Conjugating a rotation matrix by P demonstrates the action of $\rho_{(u,\theta)}$ with respect to the basis (u, v, w). Since u is fixed by the rotation matrix, the first column is $(1, 0, 0)^t$, and since the plane $u^\perp$ is being rotated by θ, the rest of the matrix M is given by the form of a 2×2 rotation matrix. That is,

$$P^{-1}\rho_{(u,\theta)}P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix} = M,$$

  which is in SO_3. Since $\rho_{(u,\theta)} = PMP^{-1}$, and since $P, P^{-1} \in O_3$ and $M \in SO_3$, the rotation matrix $\rho_{(u,\theta)}$ is also in O_3. Taking the determinant of both sides[43] demonstrates that $\rho_{(u,\theta)} \in SO_3$.

- To show the other direction, an element $A \in SO_3$ must be shown to be rotation around some axis u, which has to be some eigenvector with eigenvalue $\lambda = 1$. There exists such an eigenvector if and only if 1 is a root of the characteristic polynomial of A, which is precisely when $\det(I - A) = 0$.

[42]And orientation
[43]$\det(PMP^{-1}) = \det(P)\det(M)\det(P)^{-1} = \det(M) = 1$

Since $\det(A^T) = 1$, $\det(A - I) = \det(A^T(A - I))$. Using the fact that A is orthogonal, this is $\det(I - A^T)$. Taking the transpose, this is $\det(I - A)$. Since the matrices are 3×3, $\det(I - A) = (-1)^3 \det(A - I)$. Combining these,

$$\begin{aligned} \det(A - I) &= \det(A^T(A - I)) \\ &= \det(I - A^T) \\ &= \det(I - A) \\ &= (-1)^3 \det(A - I), \end{aligned}$$

implying that $\det(A - I) = 0$. Therefore, there does exist an eigenvector of eigenvalue 1 for A, which can be scaled to be a unit vector u.

We extend u to an orthonormal basis $P = (u, v, w)$ by picking an orthonormal basis for $u^\perp$. Consider taking A in this basis. The first column is $(1, 0, 0)^t$, since u is an eigenvector, and the first row is $(1, 0, 0)$ because the columns are orthogonal. Then, the bottom right submatrix is an element of SO_2 by taking the determinant. So

$$P^{-1}AP = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos & -\sin \\ 0 & \sin & \cos \end{pmatrix},$$

and we are done.

$\square$

Resource: Algebra I Student Notes
Fall 2021
Instructor: Davesh Maulik
Notes taken by Jakin Ng, Sanjana Das, and Ethan Yang