Lectures notes on orthogonal matrices (with exercises)
92.222 - Linear Algebra II - Spring 2004
by D. Klain

1. Orthogonal matrices and orthonormal sets

An $n \times n$ real-valued matrix A is said to be an *orthogonal matrix* if

$$A^T A = I, \tag{1}$$

or, equivalently, if $A^T = A^{-1}$.

If we view the matrix A as a family of column vectors:

$$A = \left[\begin{array}{c|c|c|c} A_1 & A_2 & \cdots & A_n \end{array} \right]$$

then

$$A^T A = \left[\begin{array}{c} A_1^T \\ \hline A_2^T \\ \hline \vdots \\ \hline A_n^T \end{array} \right] \left[\begin{array}{c|c|c|c} A_1 & A_2 & \cdots & A_3 \end{array} \right] = \begin{bmatrix} A_1^T A_1 & A_1^T A_2 & \cdots & A_1^T A_n \\ A_2^T A_1 & A_2^T A_2 & \cdots & A_2^T A_n \\ \vdots & \vdots & \ddots & \vdots \\ A_n^T A_1 & A_n^T A_2 & \cdots & A_n^T A_n \end{bmatrix}$$

So the condition (1) asserts that A is an orthogonal matrix iff

$$A_i^T A_j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

that is, iff the columns of A form an orthonormal set of vectors.

Orthogonal matrices are also characterized by the following theorem.

Theorem 1 *Suppose that A is an $n \times n$ matrix. The following statements are equivalent:*

1. *A is an orthogonal matrix.*
2. *$|AX| = |X|$ for all $X \in \mathbb{R}^n$.*
3. *$AX \cdot AY = X \cdot Y$ for all $X, Y \in \mathbb{R}^n$.*

In other words, a matrix A is orthogonal iff A preserves distances and iff A preserves dot products.

Proof: We will prove that *1.* $\Rightarrow$ *2.* $\Rightarrow$ *3.* $\Rightarrow$ *1.*

1. $\Rightarrow$ *2.* : Suppose that A is orthogonal, so that $A^TA = I$. For all column vectors $X \in \mathbb{R}^n$, we have

$$|AX|^2 = (AX)^TAX = X^TA^TAX = X^TIX = X^TX = |X|^2,$$

so that $|AX| = |X|$.

2. $\Rightarrow$ *3.* : Suppose that A is a square matrix such that $|AX| = |X|$ for all $X \in \mathbb{R}^n$. Then, for all $X, Y \in \mathbb{R}^n$, we have

$$|X+Y|^2 = (X+Y)^T(X+Y) = X^TX + 2X^TY + Y^TY = |X|^2 + 2X^TY + |Y|^2$$

and similarly,

$$|A(X+Y)|^2 = |AX+AY|^2 = (AX+AY)^T(AX+AY) = |AX|^2 + 2(AX)^TAY + |AY|^2.$$

Since $|AX| = |X|$ and $|AY| = |Y|$ and $|A(X+Y)| = |X+Y|$, it follows that $(AX)^TAY = X^TY$. In other words, $AX \cdot AY = X \cdot Y$.

3. $\Rightarrow$ *1.* : Suppose that A is a square matrix such that $AX \cdot AY = X \cdot Y$ for all $X, Y \in \mathbb{R}^n$ Let e_i denote the i-th standard basis vector for $\mathbb{R}^n$, and let A_i denote the i-th column of A, as above. Then

$$A_i^TA_j = (Ae_i)^TAe_j = Ae_i \cdot Ae_j = e_i \cdot e_j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

so that the columns of A are an orthonormal set, and A is an orthogonal matrix. ∎

We conclude this section by observing two useful properties of orthogonal matrices.

Proposition 2 *Suppose that A and B are orthogonal matrices.*

1. *AB is an orthogonal matrix.*
2. *Either* $\det(A) = 1$ *or* $\det(A) = -1$.

The proof is left to the exercises.

Note: The converse is false. There exist matrices with determinant ± 1 that are *not* orthogonal.

2. The n-Reflections Theorem

Recall that if $u \in \mathbb{R}^n$ is a unit vector and $W = u^\perp$ then

$$H = I - 2uu^T$$

is the reflection matrix for the subspace W. Since reflections preserve distances, it follows from Theorem 1 that H must be an orthogonal matrix. (You can also verify condition (1) directly.) We also showed earlier in the course that $H = H^{-1} = H^T$ and $H^2 = I$.

It turns out that every orthogonal matrix can be expressed as a product of reflection matrices.

Theorem 3 (**n-Reflections Theorem**) *Let A be an $n \times n$ orthogonal matrix. There exist $n \times n$ reflection matrices $H_1, H_2, \ldots, H_k$ such that $A = H_1 H_2 \cdots H_k$, where $0 \le k \le n$.*

In other words, every $n \times n$ orthogonal matrix can be expressed as a product of at most n reflections.

Proof: The theorem is trivial in dimension 1. Assume it holds in dimension $n-1$.

For the n-dimensional case, let $z = Ae_n$, and let H be a reflection of $\mathbb{R}^n$ that exchanges z and e_n. Then $HAe_n = Hz = e_n$, so HA fixes e_n. Moreover, HA is also an orthogonal matrix by Proposition 2, so HA preserves distances and angles. In particular, if we view $\mathbb{R}^{n-1}$ as the hyperplane $e_n^\perp$, then HA must map $\mathbb{R}^{n-1}$ to itself. By the induction assumption, HA must be expressible as a product of at most $n-1$ reflections on $\mathbb{R}^{n-1}$, which extend (along the e_n direction) to reflections of $\mathbb{R}^n$ as well. In other words, either $HA = I$ or

$$HA = H_2 \cdots H_k,$$

where $k \le n$. Setting $H_1 = H$, and keeping in mind that $HH = I$ (since H is a reflection!), we have

$$A = HHA = H_1 H_2 \cdots H_k.$$

∎

Proposition 4 *If H is a reflection matrix, then* $\det H = -1$.

Proof: See Exercises.

Corollary 5 *If A is an orthogonal matrix and $A = H_1 H_2 \cdots H_k$, then* $\det A = (-1)^k$.

So an orthogonal matrix A has determinant equal to $+1$ iff A is a product of an *even* number of reflections.

3. Classifying 2×2 Orthogonal Matrices

Suppose that A is a 2×2 orthogonal matrix. We know from the first section that the columns of A are unit vectors and that the two columns are perpendicular (orthonormal!). Any unit vector $\mathbf{u}$ in the plane $\mathbb{R}^2$ lies on the unit circle centered at the origin, and so can be expressed in the form $\mathbf{u} = (\cos\theta, \sin\theta)$ for some angle θ. So we can describe the first column of A as follows:

$$A = \begin{bmatrix} \cos\theta & ?? \\ \sin\theta & ?? \end{bmatrix}$$

What are the possibilities for the second column? Since the second column must be a unit vector perpendicular to the first column, there remain only two choices:

$$A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \qquad \text{or} \qquad A = \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}$$

The first case is the *rotation matrix*, which rotates $\mathbb{R}^2$ counterclockwise around the origin by the angle θ. The second case is a *reflection* across the line that makes an angle $\theta/2$ from the x-axis (counterclockwise).

But which is which? You can check that the rotation matrix (on the left) has determinant 1, while the reflection matrix (on the right) has determinant -1. This is consistent with Proposition 4 and Corollary 5, since a rotation of $\mathbb{R}^2$ can always be expressed as a product of two reflections (how?).

4. Classifying 3×3 Orthogonal Matrices

The n-Reflection Theorem 3 leads to a complete description of the $3\times$ orthogonal matrices. In particular, a 3×3 orthogonal matrix must a product of 0, 1, 2, or 3 reflections.

Theorem 6 *Let A be a 3×3 orthogonal matrix.*

1. *If $\det A = 1$ then A is a rotation matrix.*
2. *If $\det A = -1$ and $A^T = A$, then either $A = -I$ or A is a reflection matrix.*
3. *If $\det A = -1$ and $A^T \neq A$, then A is a product of 3 reflections (that is, A is a non-trivial rotation followed by a reflection).*

Proof: By Theorem 3, A is a product of 0, 1, 2, or 3 reflections. Note that if $A = H_1 \cdots H_k$, then $\det A = (-1)^k$.

If $\det A = 1$ then A must be a product of an even number of reflections, either 0 reflections (so that $A = I$, the trivial rotation), or 2 reflections, so that A is a rotation.

If $\det A = -1$ then A must be a product of an odd number of reflections, either 1 or 3.

If A is a single reflection then $A = H$ for some Householder matrix H. In this case we observed earlier that $H^T = H$ so $A^T = A$.

Conversely, if $\det A = -1$ and $A^T = A$ then $\det(-A) = 1$ (since A is a 3×3 matrix) and $-A^T = -A = -A^{-1}$ as well. It follows that $-A$ is a rotation that squares to the identity. If $A \neq I$, then the only time this happens is when we rotate by the angle π (that is, 180^o) around some axis. But if $-A$ is a 180^o rotation around some axis, then $A = -(-A)$ must be the reflection across the equatorial plane for that axis (draw a picture!). So A is a single reflection.

Finally if $\det A = -1$ and $A^T \neq A$, then A cannot be a rotation or a pure reflection, so A must be a product of at least 3 reflections. ∎

Corollary 7 *Let A be a 3×3 orthogonal matrix.*

1. *If $\det A = 1$ then A is a rotation matrix.*
2. *If $\det A = -1$ then $-A$ is a rotation matrix.*

Proof: If $\det A = 1$ then A is a rotation matrix, by Theorem 6. If $\det A = -1$ then $\det(-A) = (-1)^3 \det A = 1$. Since $-A$ is also orthogonal, $-A$ must be a rotation. ∎

Corollary 8 *Suppose that A and B are 3×3 rotation matrices. Then AB is also a rotation matrix.*

Proof: If A and B are 3×3 rotation matrices, then A and B are both orthogonal with determinant $+1$. It follows that AB is orthogonal, and $\det AB = \det A \det B = 1 \cdot 1 = 1$. Theorem 6 then implies that AB is also a rotation matrix. ∎

Note that the rotations represented by A, B, and AB may each have completely different angles and axes of rotation! Given two rotations A and B around two *different* axes of rotation, it is far from obvious that AB will also be a rotation (around some mysterious third axis). But this is true, by Corollary 8. Later on we will see how to compute precisely the angle and axis of rotation of a rotation matrix.

Exercises:

1. (a) Suppose that A is an orthogonal matrix.
Prove that either $\det A = 1$ or $\det A = -1$.

(b) Find a 2×2 matrix A such that $\det A = 1$, but also such that A is *not* an orthogonal matrix.

2. Suppose that A and B are orthogonal matrices. Prove that AB is an orthogonal matrix.

3. Suppose that $H = I - 2uu^T$ is a reflection matrix. Let $v_1, \ldots, v_{n-1}$ be an orthonormal basis for the subspace $u^\perp$. (Here $u^\perp$ denotes the orthogonal complement to the line spanned by u.)

(a) What is $Hu = ?$

(b) What is $Hv_i = ?$

(c) Let M be the matrix with columns as follows

$$M = \left[\begin{array}{c|c|c|c} v_1 & \cdots & v_{n-1} & u \end{array} \right]$$

What is $HM = ?$ What are the columns of HM?

(d) By comparing $\det HM$ to $\det M$ prove that $\det H = -1$.

4. (a) Give a geometric description (and sketch) of the reflection performed by the matrix

$$H_1 = \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}$$

(b) Give a geometric description (and sketch) of the reflection performed by the matrix

$$H_2 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

(c) Show that a rotation in $\mathbb{R}^2$ is a product of two reflections by showing that $H_1 H_2$ is a rotation matrix. Give a geometric description (and sketch) of the rotation performed by the matrix $H_1 H_2$.

5. Let u, v, w be an orthonormal basis (of column vectors) for $\mathbb{R}^3$, and let A be the matrix given by

$$A = uu^T + vv^T + ww^T.$$

(a) What is $Au =$? What is $Av =$? What is $Aw =$?

(b) Let P be the matrix

$$P = \left[\begin{array}{c|c|c} u & v & w \end{array}\right]$$

What is $AP =$?

(c) Prove that $A = I$ (the identity!).

6. Give geometric descriptions of the what happens when you multiply a vector in $\mathbb{R}^3$ by each of the following orthogonal matrices:

$$\begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \begin{bmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix}$$

Sketch what happens in each case.