Math 254: Introduction to Linear Algebra
Lecture Notes #5.3 — Orthogonal Transformations and Orthogonal Matrices

Peter Blomgren,
⟨blomgren.peter@gmail.com⟩
Department of Mathematics and Statistics
Dynamical Systems Group
Computational Sciences Research Center
San Diego State University
San Diego, CA 92182-7720
http://terminus.sdsu.edu/

Spring 2017

Outline

1 Student Learning Objectives
   ● SLOs: Orthogonal Transformations and Orthogonal Matrices

2 Orthogonal Transformations and Orthogonal Matrices
   ● Examples, and Fundamental Theorems
   ● Products, Inverses, and Transposes of Orthogonal Matrices
   ● Matrix of Orthogonal Projection, using Orthonormal Basis

3 Suggested Problems
   ● Suggested Problems 5.3
   ● Lecture–Book Roadmap

Orthogonal Transformations

For many reasons, we tend to “like” linear transformations that preserve the length of vectors; and angles between vectors:

Definition (Orthogonal Transformations)

A linear transformation \( T : \mathbb{R}^n \rightarrow \mathbb{R}^n \) is called orthogonal if it preserves the length of vectors:

\[
\| T(\vec{x}) \| = \| \vec{x} \|, \ \forall \vec{x} \in \mathbb{R}^n.
\]

If \( T(\vec{x}) = A\vec{x} \) is an orthogonal transformation, we say that \( A \) is an orthogonal (or unitary, when it has complex entries) matrix.
**Example: Rotations**

The rotation

\[ T(x) = \begin{bmatrix} \cos(\phi) & -\sin(\phi) \\ \sin(\phi) & \cos(\phi) \end{bmatrix} x \]

is an orthogonal transformation from \( \mathbb{R}^2 \) to \( \mathbb{R}^2 \), and \( \forall \theta \).

**Example (Rotations)**

![Rotation 1](image1.png)  ![Rotation 2](image2.png)  ![Rotation 3](image3.png)

**Example: Reflections**

Consider a subspace \( V \) of \( \mathbb{R}^n \). For a vector \( \vec{x} \in \mathbb{R}^n \), the vector \( \text{ref}_V(\vec{x}) = \vec{x}^\perp - \vec{x}^\perp \equiv 2\text{proj}_V(\vec{x}) - \vec{x} \) is the reflection of \( \vec{x} \) in \( V \). We show that reflections are orthogonal transformations:

By the Pythagorean theorem, we have

\[ ||\text{ref}_V(\vec{x})||^2 = ||\vec{x}^\perp - \vec{x}^\perp||^2 = ||\vec{x}^\perp||^2 + ||-\vec{x}^\perp||^2 = ||\vec{x}^\perp||^2 + ||\vec{x}^\perp||^2 = ||\vec{x}||^2 \]

**Theorem (Preservation of Orthogonality)**

Consider an orthogonal transformation \( T : \mathbb{R}^n \rightarrow \mathbb{R}^n \). If the vectors \( \vec{v}, \vec{w} \in \mathbb{R}^n \) are orthogonal, then so are \( T(\vec{v}) \) and \( T(\vec{w}) \).

**Proof.**

By the theorem of Pythagoras, we have to show that

\[ ||T(\vec{v}) + T(\vec{w})||^2 = ||T(\vec{v})||^2 + ||T(\vec{w})||^2 : \]

\[ ||T(\vec{v}) + T(\vec{w})||^2 = ||\vec{v} + \vec{w}||^2 \quad \text{[Linearity of } T\text{]} \]

\[ = ||\vec{v} + \vec{w}||^2 \quad \text{[Orthogonality of } T\text{]} \]

\[ = ||\vec{v}||^2 + ||\vec{w}||^2 \quad \text{[Orthogonality of } T\text{]} \]

\[ = ||T(\vec{v})||^2 + ||T(\vec{w})||^2 \quad \text{[Orthogonality of } T\text{]} \]

**Theorem (Orthogonal Transformations and Orthonormal Bases)**

a. A linear transformation \( T : \mathbb{R}^n \rightarrow \mathbb{R}^n \) is orthogonal if and only if the vectors \( T(\vec{e}_1), \ldots, T(\vec{e}_n) \) form an orthonormal basis of \( \mathbb{R}^n \).

b. An \( n \times n \) matrix \( A \) is orthogonal if and only if its columns form an orthonormal basis of \( \mathbb{R}^n \).

**Part (a).**

\[ \Rightarrow \text{ If } T \text{ is orthogonal, then, by definition, the } T(\vec{e}_i) \text{ are unit vectors, and orthogonal by the previous theorem; hence a basis for } \mathbb{R}^n. \]

\[ \Leftarrow \text{ Conversely, suppose } T(\vec{e}_1), \ldots, T(\vec{e}_n) \text{ form an orthonormal basis. Consider a vector } \vec{x} = x_1\vec{e}_1 + \cdots + x_n\vec{e}_n \in \mathbb{R}^n. \text{ Then} \]

\[ ||T(\vec{x})||^2 = ||x_1T(\vec{e}_1) + \cdots + x_nT(\vec{e}_n)||^2 \]

\[ = ||x_1(\vec{e}_1)||^2 + \cdots + ||x_n(\vec{e}_n)||^2 \quad \text{[by Pythagoras]} \]

\[ = x_1^2 + \cdots + x_n^2 \]

\[ = ||\vec{x}||^2. \]
Orthogonal Transformations and Orthonormal Bases

Part (b).
This follows from the result from Notes 2.1 restated below...

Theorem (The Columns of the Matrix of a Linear Transformation)
Consider a linear transformation \( T : \mathbb{R}^m \to \mathbb{R}^n \). Then, the matrix of \( T \) is

\[
A = \begin{bmatrix}
T(e_1) & T(e_2) & \ldots & T(e_m)
\end{bmatrix},
\]

where \( e_i \in \mathbb{R}^m \) is the vector of all zeros, except entry \( \#i \) which is 1.

A Warning

A matrix with orthogonal columns need not be an orthogonal matrix, e.g.

\[
A = \begin{bmatrix} 4 & -3 \\ 3 & 4 \end{bmatrix}.
\]

Example (\( A = \begin{bmatrix} 4 & -3 \\ 3 & 4 \end{bmatrix} \))

\[
\vec{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \|\vec{x}\| = \sqrt{2}, \quad A\vec{x} = \begin{bmatrix} 1 \\ 7 \end{bmatrix}, \quad \|A\vec{x}\| = \sqrt{50}
\]

Products and Inverses of Orthogonal Matrices

Theorem (Products and Inverses of Orthogonal Matrices)

\( a \). The product \( AB \) of two orthogonal \( n \times n \) matrices \( A \) and \( B \) is orthogonal.

\( b \). The inverse \( A^{-1} \) of an orthogonal \( n \times n \) matrix \( A \) is orthogonal.

Proof.
In part (a), the linear transformation \( T(\vec{x}) = AB\vec{x} \) preserves length, because \( \|T(\vec{x})\| = \|A(B\vec{x})\| = \|B\vec{x}\| = \|\vec{x}\| \). In part (b), the linear transformation \( T(\vec{x}) = A^{-1}\vec{x} \) preserves length, since \( \|A^{-1}\vec{x}\| = \|AA^{-1}\vec{x}\| = \|\vec{x}\| \).
Matrix Transpose, Symmetric and Skew-symmetric Matrices

Definition (Matrix Transpose, Symmetric and Skew-symmetric Matrices)
Consider an \( m \times n \) matrix \( A \).
- The transpose \( A^T \) of \( A \) is the \( n \times m \) matrix whose \( ij \)th entry is the \( ji \)th entry of \( A \): The roles of rows and columns are reversed.
- We say that a square matrix \( A \) is symmetric if \( A^T = A \), and
- \( A \) is called skew-symmetric if \( A^T = -A \).

Example (Transpose)
\[
A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}, \quad A^T = \begin{bmatrix} 1 & 2 \\ 4 & 5 \\ 3 & 6 \end{bmatrix}
\]

 transpose of a vector

Example (Transpose of a Vector)
\[
\vec{v} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \Rightarrow \vec{v}^T = [1 \ 2 \ 3]
\]

We use this all the time:

Theorem
If \( \vec{v} \) and \( \vec{w} \) are two (column) vectors \( \in \mathbb{R}^n \), then
\[
\vec{v} \cdot \vec{w} = \vec{v}^T \vec{w}
\]

Dot Product "Matrix" Product

Orthogonal Matrices: \( A^T \) and \( A^{-1} \)

Theorem
Consider an \( n \times n \) matrix \( A \). The matrix \( A \) is orthogonal if and only if \( A^T A = I_n \) or, equivalently, if \( A^{-1} = A^T \).

Proof.
Write \( A \) in terms of its columns:
\[
A = [\vec{v}_1 \ldots \vec{v}_n]
\]
then
\[
A^T A = \begin{bmatrix} \vec{v}_1^T \\ \vdots \\ \vec{v}_n^T \end{bmatrix} \begin{bmatrix} \vec{v}_1 & \ldots & \vec{v}_n \end{bmatrix} = \begin{bmatrix} \vec{v}_1^T \vec{v}_1 & \ldots & \vec{v}_1^T \vec{v}_n \\ \vdots & \ddots & \vdots \\ \vec{v}_n^T \vec{v}_1 & \ldots & \vec{v}_n^T \vec{v}_n \end{bmatrix}
\]
this is \( I_n \) if and only if \( A \) is orthogonal.

Summary
Consider an \( n \times n \) matrix \( A \). The following statements are equivalent:

i. \( A \) is an orthogonal matrix.
ii. The transformation \( T(\vec{x}) = A\vec{x} \) preserves length, that is, \( \|A\vec{x}\| = \|\vec{x}\| \forall \vec{x} \in \mathbb{R}^n \).
iii. The columns of \( A \) form an orthonormal basis of \( \mathbb{R}^n \).
iv. \( A^T A = I_n \).
v. \( A^{-1} = A^T \).
vi. \( A \) preserves the dot product, meaning that \( (A\vec{x}) \cdot (A\vec{y}) = \vec{x} \cdot \vec{y} \forall \vec{x}, \vec{y} \in \mathbb{R}^n \).
**Properties of the Matrix Transpose**

**Theorem (Properties of the Transpose)**

- (a) \((A + B)^T = A^T + B^T\) \(\forall A, B \in \mathbb{R}^{m \times n}\)
- (b) \((kA)^T = kA^T\) \(\forall A \in \mathbb{R}^{m \times n}, \forall k \in \mathbb{R}\)
- (c) \((AB)^T = (B^T A^T)\) \(\forall A \in \mathbb{R}^{m \times p}, \forall B \in \mathbb{R}^{p \times n}\)
- (d) \(\text{rank}(A) = \text{rank}(A^T)\) \(\forall \text{matrices } A\)
- (e) \((A^T)^{-1} = (A^{-1})^T\) \(\forall \text{invertible matrices } A\)

**Note:** Outer Product

\[ A = \begin{bmatrix} u_1 & \cdots & u_n \end{bmatrix} \begin{bmatrix} u_1^T & \cdots & u_n^T \end{bmatrix} \text{ is known as the outer product.} \]

**The Matrix of an Orthogonal Projection**

We can use our expanded matrix-notation-language to express orthogonal projections. First consider

\[ \text{proj}_L(\vec{x}) = (\vec{u}_1 \cdot \vec{x})\vec{u}_1 \]

onto a line \(L\) in \(\mathbb{R}^n\); where \(\vec{u}_1\) is a unit vector in \(L\). Think of this vector as an \(n \times 1\) matrix, and the scalar \((\vec{u}_1 \cdot \vec{x})\) as an \(1 \times 1\) matrix; we can rearrange

\[ \text{proj}_L(\vec{x}) = \vec{u}_1(\vec{u}_1 \cdot \vec{x}) = \vec{u}_1 \vec{u}_1^T \vec{x} = (\vec{u}_1 \vec{u}_1^T)\vec{x} = A\vec{x} \]

where \(A = \vec{u}_1 \vec{u}_1^T\).

**The Matrix of an Orthogonal Projection: Summary**

We apply the same idea to the general projection formula

\[ \text{proj}_V(\vec{x}) = (\vec{u}_1 \cdot \vec{x})\vec{u}_1 + \cdots + (\vec{u}_n \cdot \vec{x})\vec{u}_n = \vec{u}_1\vec{u}_1^T\vec{x} + \cdots + \vec{u}_n\vec{u}_n^T\vec{x} = \underbrace{\begin{bmatrix} \vec{u}_1^T \\ \vdots \\ \vec{u}_n^T \end{bmatrix} \vec{x}}_A \]

and we can also write

\[ A = \begin{bmatrix} \vec{u}_1 & \cdots & \vec{u}_n \end{bmatrix} \begin{bmatrix} \vec{u}_1^T \\ \vdots \\ \vec{u}_n^T \end{bmatrix} \]

We summarize on the next slide...

\[ \text{Note: } P = QQ^T, \text{ where } Q = \begin{bmatrix} \vec{q}_1 & \cdots & \vec{q}_m \end{bmatrix}. \]

\[ \text{Note that it is } QQ^T \text{ not } Q^T Q \]

\(P\) is symmetric — \(P^T = (QQ^T)^T = (Q^T)^T QT = QQ^T = P\).
Example (Following 5.2.7)

In (5.2.7) [see Learning Glass] we orthogonalized the vectors
\[
\begin{pmatrix}
2 \\
1
\end{pmatrix}, \quad \begin{pmatrix}
2 \\
1
\end{pmatrix}, \quad \begin{pmatrix}
18 \\
0
\end{pmatrix},
\]

using the Gram-Schmidt method, and got
\[
\begin{pmatrix}
\frac{2}{3} \\
\frac{1}{3}
\end{pmatrix}, \quad \begin{pmatrix}
\frac{-2}{3} \\
\frac{2}{3}
\end{pmatrix}, \quad \begin{pmatrix}
\frac{1}{3} \\
\frac{-2}{3}
\end{pmatrix},
\]

Let’s define
\[
Q_1 = \begin{bmatrix} \bar{q}_1 \end{bmatrix}, \quad Q_2 = \begin{bmatrix} \bar{q}_1 & \bar{q}_2 \end{bmatrix}, \quad Q_3 = \begin{bmatrix} \bar{q}_1 & \bar{q}_2 & \bar{q}_3 \end{bmatrix}
\]

Now,
\[
Q_1 Q_1^T = \frac{1}{9} \begin{bmatrix} 4 & 4 & 2 \\ 4 & 4 & 2 \\ 2 & 2 & 1 \end{bmatrix}, \quad Q_1^T Q_1 = [1]
\]
\[
Q_2 Q_2^T = \frac{1}{9} \begin{bmatrix} 8 & 2 & -2 \\ 2 & 5 & 4 \\ -2 & 4 & 5 \end{bmatrix}, \quad Q_2^T Q_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}
\]
\[
Q_3 Q_3^T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad Q_3^T Q_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}
\]

Note, $Q_1 Q_1^T$, $Q_2 Q_2^T$, and $Q_3 Q_3^T$ are the matrices of orthogonal projections onto a line span($\bar{q}_1$), a plane span($\bar{q}_1, \bar{q}_2$), and $\mathbb{R}^3$ span($\bar{q}_1, \bar{q}_2, \bar{q}_3$).

Available on Learning Glass videos:

5.3 — 1, 2, 5, 6, 13, 15, 17, 19, 28, 32, 33, 36, 41