Orthogonal Transforms and Orthogonal Matrices

SLOs 5.3

After this lecture you should:

- Know what Orthogonal Transformations are; and their relation to Orthonormal Bases.
- Know the Properties of Orthogonal Matrices.
- Be able to perform an Orthogonal Projection using Orthonormal Basis you have constructed.

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
\[
T(\vec{x}) = \begin{bmatrix} \cos(\phi) & -\sin(\phi) \\ \sin(\phi) & \cos(\phi) \end{bmatrix} \vec{x}
\]
is an orthogonal transformation from \(\mathbb{R}^2\) to \(\mathbb{R}^2\), and \(\forall \theta\).

Example (Rotations)

Example (Reflections)

Consider a subspace \(V\) of \(\mathbb{R}^n\). For a vector \(\vec{x} \in \mathbb{R}^n\), the vector \(\text{ref}_V(\vec{x}) = \vec{x}^\perp - \vec{x}^\perp = 2\text{proj}_V(\vec{x}) - \vec{x}\) is the reflection of \(\vec{x}\) in \(V\).

We show that reflections are orthogonal transformations:

By the Pythagorean theorem, we have
\[
||\text{ref}_V(\vec{x})||^2 = ||\vec{x}^\perp||^2 + ||\vec{x}^\perp||^2 = ||\vec{x}||^2 + ||\vec{x}^\perp||^2 = ||\vec{x}||^2
\]

Orthogonal Transformations and Orthogonal Matrices — (5/24)

Preservation of Orthogonality

Theorem (Preservation of Orthogonality)

Consider an orthogonal transformation \(T : \mathbb{R}^n \to \mathbb{R}^n\). 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
\[
||T(\vec{v}) + T(\vec{w})||^2 = ||T(\vec{v})||^2 + ||T(\vec{w})||^2 :
\]

\[
\begin{align*}
||T(\vec{v}) + T(\vec{w})||^2 &= ||T(\vec{v} + \vec{w})||^2 & [\text{Linearity of } T] \\
&= ||\vec{v} + \vec{w}||^2 & [\text{Orthogonality of } T] \\
&= ||\vec{v}||^2 + ||\vec{w}||^2 & [\vec{v} \perp \vec{w}] \\
&= ||T(\vec{v})||^2 + ||T(\vec{w})||^2 & [\text{Orthogonality of } T]
\end{align*}
\]

Orthogonal Transformations and Orthogonal Matrices — (6/24)

Orthogonal Transformations and Orthonormal Bases

Theorem (Orthogonal Transformations and Orthonormal Bases)

a. A linear transformation \(T : \mathbb{R}^n \to \mathbb{R}^n\) 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).

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

\(\Leftarrow\) Conversely, suppose \(T(\vec{e}_1), \ldots, T(\vec{e}_n)\) form an orthonormal basis. Consider a vector \(\vec{x} = x_1\vec{e}_1 + \cdots + x_n\vec{e}_n \in \mathbb{R}^n\). Then
\[
||T(\vec{x})||^2 = ||x_1 T(\vec{e}_1) + \cdots + x_n T(\vec{e}_n)||^2
\]

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

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

\[
= ||\vec{x}||^2.
\]
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

$$A = \begin{bmatrix} T(\vec{e}_1) & T(\vec{e}_2) & \cdots & T(\vec{e}_m) \end{bmatrix},$$

where $\vec{e}_i \in \mathbb{R}^m$ is the vector of all zeros, except entry #i which is 1.

Products and Inverses of Orthogonal Matrices

Example

Consider the orthogonal matrix

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

We compute $\|A\vec{x}\| = \|\vec{x}\|$, $\|B\vec{x}\| = \|\vec{x}\|$, $\|AB\vec{x}\| = \|\vec{x}\|$, $\|BA\vec{x}\| = \|\vec{x}\|$.

Example: Properties of the Transpose of an Orthonormal Matrix

Consider the orthogonal matrix $A$, and the matrix where the $ij$ entry has been shifted to the $ji$ position ($B$):

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 7 & 2 & 6 \\ 3 & 2 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 2 & 3 \\ 6 & 2 & -3 \\ 3 & -6 & 2 \end{bmatrix}.$$

We compute

$$BA = \begin{bmatrix} 1 & 9 \\ 0 & 49 \\ 0 & 49 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

The $k, \ell$ entry in $BA$ is the dot product of the $k$th row of $B$, and the $\ell$th column of $A$; by construction this is the dot product of the $k$th and $\ell$th columns of $A$; since $A$ is orthogonal this gives 1 when $k = \ell$, and 0 otherwise.
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)
\[
A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}, \quad A^T = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}
\]

Transpose of a Vector

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

We use this all the time:

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

Orthogonal Matrices: Summary

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. \((A + B)^T = A^T + B^T\) \quad \forall A, B \in \mathbb{R}^{m \times n}

b. \((kA)^T = kA^T\) \quad \forall A \in \mathbb{R}^{m \times n}, \forall k \in \mathbb{R}

c. \((AB)^T = (B^T A^T)\) \quad \forall A \in \mathbb{R}^{m \times p}, \forall B \in \mathbb{R}^{p \times n}

d. \text{rank}(A) = \text{rank}(A^T) \quad \forall \text{matrices } A

e. \((A^T)^{-1} = (A^{-1})^T\) \quad \forall \text{invertible matrices } A

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

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

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

Note: Outer Product

\[ \begin{bmatrix} A \end{bmatrix} = \begin{bmatrix} \vec{u}_1 \vec{u}_1^T \end{bmatrix} \]

is known as the **outer product**.

The Matrix of an Orthogonal Projection: Summary

Theorem (The Matrix of an Orthogonal Projection: Summary)

Consider a subspace \(V\) of \(\mathbb{R}^n\) with orthonormal basis \(\vec{q}_1, \ldots, \vec{q}_m\). The matrix \(P\) of the orthogonal projection onto \(V\) is

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

- Note that it is \(QQ^T\) **not** \(Q^TQ\)
- \(P\) is symmetric — \(P^T = (QQ^T)^T = (Q^T)^TQ^T = QQ^T = P\)
Example (Following 5.2.7)

In (5.2.7) [see learning glass] we orthogonalized the vectors

\[ \vec{v}_1 = \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} -2 \\ 1 \\ 2 \end{bmatrix}, \quad \vec{v}_3 = \begin{bmatrix} 18 \\ 0 \\ 0 \end{bmatrix}, \]

using the Gram-Schmidt method, and got

\[ \vec{q}_1 = \begin{bmatrix} 2/3 \\ 2/3 \\ 1/3 \end{bmatrix}, \quad \vec{q}_2 = \begin{bmatrix} -2/3 \\ 1/3 \\ 2/3 \end{bmatrix}, \quad \vec{q}_3 = \begin{bmatrix} 1/3 \\ -2/3 \\ 2/3 \end{bmatrix}, \]

Let’s define

\[ Q_1 = [\vec{q}_1], \quad Q_2 = [\vec{q}_1 \quad \vec{q}_2], \quad Q_3 = [\vec{q}_1 \quad \vec{q}_2 \quad \vec{q}_3] \]

Now,

\[ Q_1 Q_1^T = \frac{1}{9} \begin{bmatrix} 4 & 4 & 2 \\ 4 & 4 & 2 \\ 2 & 2 & 1 \end{bmatrix}, \quad Q_1^T Q_1 = \begin{bmatrix} 1 \end{bmatrix} \]

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

\[ Q_3 Q_3^T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad Q_3^T Q_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \]

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 \( \text{span}(\vec{q}_1) \), a plane \( \text{span}(\vec{q}_1, \vec{q}_2) \), and \( \mathbb{R}^3 \) span(\( \vec{q}_1, \vec{q}_2, \vec{q}_3 \)).

Suggested Problems 5.3

Available on Learning Glass videos:

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