• Instructor: He Wang Email: he.wang@northeastern.edu

Definition.

A linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ is called **orthogonal** iff

$$||T(\vec{x})|| = ||\vec{x}|| \text{ for all } \vec{x} \in \mathbb{R}^n$$

that is, T preserves the length of vectors. The matrix of an orthogonal transformation is called an **orthogonal matrix**.

Example 1. Whether or not the following transformations are orthogonal.

(1.) Rotations $S : \mathbb{R}^2 \to \mathbb{R}^2$ are orthogonal transformations.

The matrix of rotation $S = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ is orthogonal.

(2.) Reflections $R : \mathbb{R}^2 \to \mathbb{R}^2$ are orthogonal transformations.

The matrix of reflection matrix $R = \begin{bmatrix} a & b \\ b & -a \end{bmatrix}$ with $a^2 + b^2 = 1$ is orthogonal.

(3.) Orthogonal projections $P : \mathbb{R}^2 \to \mathbb{R}^2$ are NOT orthogonal transformations.

Theorem.

Let U be an $n \times n$ orthogonal matrix and let $\vec{x}$ and $\vec{y}$ be any vectors in $\mathbb{R}^n$. Then

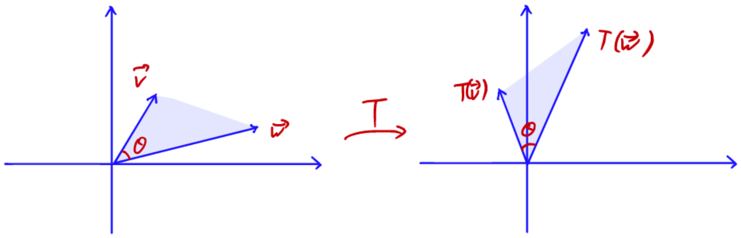
1. $||U \cdot \vec{x}|| = ||\vec{x}||$.
2. $(U\vec{x}) \cdot (U\vec{y}) = \vec{x} \cdot \vec{y}$.
3. $(U\vec{x}) \cdot (U\vec{y}) = 0$ if and only if $\vec{x} \cdot \vec{y} = 0$.

The transformation $T(\vec{x}) = U\vec{x}$ is orthogonal. So, we have 1.
For 2. $||U(\vec{x} + \vec{y})||^2 = ||\vec{x} + \vec{y}||^2 = (\vec{x} + \vec{y}) \cdot (\vec{x} + \vec{y}) = ||\vec{x}||^2 + 2\vec{x} \cdot \vec{y} + ||\vec{y}||^2$
$||U(\vec{x} + \vec{y})||^2 = ||U(\vec{x}) + U(\vec{y})||^2 = ||U(\vec{x})||^2 + ||U(\vec{y})||^2 + 2(U\vec{x}) \cdot (U\vec{y})$.
Compare two formulas, we have $(U\vec{x}) \cdot (U\vec{y}) = \vec{x} \cdot \vec{y}$.

Proposition.

U is an orthogonal matrix if and only if $(U\vec{x}) \cdot (U\vec{y}) = \vec{x} \cdot \vec{y}$ for any $\vec{x}$ and $\vec{y}$ in $\mathbb{R}^n$.

The above theorem says that orthogonal transformations **preserve dot products**, hence also **preserve angles** and orthogonality.



Using the geometric meaning of the orthogonal transformation, we have

Theorem.

1. If A is orthogonal, then A is invertible and A^{-1} is orthogonal.
2. If A and B are orthogonal, then AB is orthogonal.

Theorem.

The $n \times n$ matrix U is orthogonal if and only if $\{\vec{u}_1, \dots, \vec{u}_n\}$ is an orthonormal set.

Proof. "$\Rightarrow$" Suppose U is an orthogonal matrix. We prove that $\{\vec{u}_1, \dots, \vec{u}_n\}$ is an orthonormal set.
Use the property that U is orthogonal if and only if $(U\vec{x}) \cdot (U\vec{y}) = \vec{x} \cdot \vec{y}$. Apply the formula to standard vectors $\vec{x} = \vec{e}_i$ and $\vec{y} = \vec{e}_j$.

$$U\vec{e}_i = [u_1 \; u_2 \; \cdots \; u_n] \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} = \vec{u}_i$$

Hence $\vec{u}_i \cdot \vec{u}_j = (U\vec{e}_i) \cdot (U\vec{e}_j) = \vec{e}_i \cdot \vec{e}_j = \begin{cases} 0 & \text{when } i \neq j \\ 1 & \text{when } i = j \end{cases}$

So, $\{\vec{u}_1, \dots, \vec{u}_n\}$ is orthonormal.
"$\Leftarrow$" Suppose $\{\vec{u}_1, \dots, \vec{u}_n\}$ is an orthonormal set. We show that U is an orthogonal matrix.

For any $\vec{x} \in \mathbb{R}^n$, $U\vec{x} = [u_1 \; u_2 \; \cdots \; u_n] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1\vec{u}_1 + x_2\vec{u}_2 + \cdots + x_n\vec{u}_n$

$||U\vec{x}||^2 = (x_1\vec{u}_1 + x_2\vec{u}_2 + \cdots + x_n\vec{u}_n) \cdot (x_1\vec{u}_1 + x_2\vec{u}_2 + \cdots + x_n\vec{u}_n) = x_1^2 + \cdots + x_n^2 = ||\vec{x}||^2$
So, $||U\vec{x}|| = ||\vec{x}||$ and hence U is an orthogonal matrix.

Example 2. Verify that the matrix $A = \begin{bmatrix} 1/\sqrt{6} & -1/\sqrt{2} & -1/\sqrt{3} \\ 1/\sqrt{6} & 1/\sqrt{2} & -1/\sqrt{3} \\ 2/\sqrt{6} & 0 & 1/\sqrt{3} \end{bmatrix}$ is orthogonal.

Verify that $\{\vec{a}_1, \vec{a}_2, \vec{a}_3\}$ is orthonormal by verify that $\vec{a}_1 \cdot \vec{a}_2 = 0$, $\vec{a}_1 \cdot \vec{a}_3 = 0$, $\vec{a}_2 \cdot \vec{a}_3 = 0$, and $||\vec{a}_1|| = ||\vec{a}_2|| = ||\vec{a}_3|| = 1$.

Example 3. Verify that the matrix $B = \frac{1}{7}\begin{bmatrix} 2 & 3 & 6 \\ 6 & 2 & -3 \\ 3 & -6 & 2 \end{bmatrix}$ is orthogonal.

Verify that $\{\vec{a}_1, \vec{a}_2, \vec{a}_3\}$ is orthonormal by verify that $\vec{a}_1 \cdot \vec{a}_2 = 0$, $\vec{a}_1 \cdot \vec{a}_3 = 0$, $\vec{a}_2 \cdot \vec{a}_3 = 0$, and $||\vec{a}_1|| = ||\vec{a}_2|| = ||\vec{a}_3|| = 1$.

Recall the transpose of a matrix: Given an $m \times n$ matrix A, we define the ***transpose matrix*** A^T as the $n \times m$ matrix whose (i, j)-th entry is the (j, i)-th entry of A. The dot product can be written as matrix product

$$\vec{v} \cdot \vec{w} = \vec{v}^T \vec{w}$$

Theorem.

The $n \times n$ matrix A is orthogonal if and only if $A^T A = I_n$; if and only if $A^{-1} = A^T$.

Proof. A is orthogonal if and only if $\{\vec{a}_1, \dots, \vec{a}_n\}$ is orthonormal, i.e., $\vec{a}_i \cdot \vec{a}_j = 1$ if $i \neq j$ and $||\vec{a}_i|| = 1$.
On the other side, (write for the case $n = 3$)

$$A^T A = \begin{bmatrix} \vec{a}_1^T \\ \vec{a}_2^T \\ \vec{a}_3^T \end{bmatrix} \begin{bmatrix} \vec{a}_1 & \vec{a}_2 & \vec{a}_3 \end{bmatrix} = \begin{bmatrix} \vec{a}_1^T\vec{a}_1 & \vec{a}_1^T\vec{a}_2 & \vec{a}_1^T\vec{a}_3 \\ \vec{a}_2^T\vec{a}_1 & \vec{a}_2^T\vec{a}_2 & \vec{a}_2^T\vec{a}_3 \\ \vec{a}_3^T\vec{a}_1 & \vec{a}_3^T\vec{a}_2 & \vec{a}_3^T\vec{a}_3 \end{bmatrix} = \begin{bmatrix} \vec{a}_1 \cdot \vec{a}_1 & \vec{a}_1 \cdot \vec{a}_2 & \vec{a}_1 \cdot \vec{a}_3 \\ \vec{a}_2 \cdot \vec{a}_1 & \vec{a}_2 \cdot \vec{a}_2 & \vec{a}_2 \cdot \vec{a}_3 \\ \vec{a}_3 \cdot \vec{a}_1 & \vec{a}_3 \cdot \vec{a}_2 & \vec{a}_3 \cdot \vec{a}_3 \end{bmatrix} = I_3$$

Example 4. Find the inverse of matrices in Examples 2 and 3.

Since they are orthogonal, $A^{-1} = A^T = \begin{bmatrix} 1/\sqrt{6} & 1/\sqrt{6} & 2/\sqrt{6} \\ -1/\sqrt{2} & 1/\sqrt{2} & 0 \\ -1/\sqrt{3} & -1/\sqrt{3} & 1/\sqrt{3} \end{bmatrix}$ and

$B^{-1} = B^T = \frac{1}{7}\begin{bmatrix} 2 & 6 & 3 \\ 3 & 2 & -6 \\ 6 & -3 & 2 \end{bmatrix}$.

Theorem.

Let W be any subspace of $\mathbb{R}^n$ with an orthonormal basis $\{\vec{u}_1, \ldots, \vec{u}_p\}$. Let $U = [\vec{u}_1\ \vec{u}_2 \cdots \vec{u}_p]$. For any $\vec{y} \in \mathbb{R}^n$,

$$\text{proj}_W(\vec{y}) = UU^T\vec{y}.$$

That is, the **matrix of the projection** onto W is $P = UU^T$.

Remark: 1. $p < n$ since W is a subspace of $\mathbb{R}^n$. When $p = n$, then $P = I_n$.

2. We always have $U^TU = I$ for orthonormal basis $\{\vec{u}_1, \ldots, \vec{u}_p\}$.

The theorem comes from the following formula from §5.1. The idea is to translate dot product to matrix product.

$$\begin{aligned}
\text{proj}_W(\vec{y}) &= (\vec{y}\cdot\vec{u}_1)\vec{u}_1 + (\vec{y}\cdot\vec{u}_2)\vec{u}_2 + \cdots + (\vec{y}\cdot\vec{u}_p)\vec{u}_p \\
&= [\vec{u}_1\ \vec{u}_2\ \cdots\ \vec{u}_p]\begin{bmatrix}\vec{y}\cdot\vec{u}_1\\ \vec{y}\cdot\vec{u}_2\\ \vdots\\ \vec{y}\cdot\vec{u}_p\end{bmatrix} \\
&= [\vec{u}_1\ \vec{u}_2\ \cdots\ \vec{u}_p]\begin{bmatrix}\vec{u}_1\cdot\vec{y}\\ \vec{u}_2\cdot\vec{y}\\ \vdots\\ \vec{u}_p\cdot\vec{y}\end{bmatrix} \\
&= [\vec{u}_1\ \vec{u}_2\ \cdots\ \vec{u}_p]\begin{bmatrix}\vec{u}_1^T\vec{y}\\ \vec{u}_2^T\vec{y}\\ \vdots\\ \vec{u}_p^T\vec{y}\end{bmatrix} \\
&= [\vec{u}_1\ \vec{u}_2\ \cdots\ \vec{u}_p]\begin{bmatrix}\vec{u}_1^T\\ \vec{u}_2^T\\ \vdots\\ \vec{u}_p^T\end{bmatrix}\vec{y} \\
&= UU^T\vec{y}
\end{aligned}$$

Example 5.