MTHSC 3110 Section 6.2 – Orthogonal Sets

Kevin James

Definition

A set $\{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_k\}$ of non-zero vectors is said to be orthogonal if for every $i \neq j$,

$$\vec{u}_i \cdot \vec{u}_j = 0,$$

that is, if every pair of vectors is orthogonal.

Example

Show that $\left\{ \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 4 \\ -7 \end{pmatrix} \right\}$. is orthogonal.

Solution

$$\begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix} = (3)(-1) + (1)(2) + (1)(1) = 0.$$

$$\begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 4 \\ -7 \end{pmatrix} = (3)(1) + (1)(4) + (1)(-7) = 0$$

$$\begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 4 \\ -7 \end{pmatrix} = (-1)(1) + (2)(4) + (1)(-7) = 0.$$

Theorem

If $\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k\}$ is an orthogonal set of nonzero vectors, then it is linearly independent. Hence it is a basis for the space that it spans.

Proof.

Suppose that $\sum_{i=1}^{k} c_i \vec{u}_i = \vec{0}$.
Then for $1 \leq j \leq k$, we have

$$\begin{aligned} 0 &= \vec{u}_j \cdot \vec{0} == \vec{u}_j \cdot \sum_{i=1}^{k} c_i \vec{u}_i \\ &= \sum_{i=1}^{k} c_i \left(\vec{u}_j \cdot \vec{u}_i\right) = c_j \left(\vec{u}_j \cdot \vec{u}_j\right). \\ \Rightarrow & \quad c_j = 0 \text{ (because, } \vec{u}_j \neq \vec{0}\text{).} \end{aligned}$$

Thus $c_1 = c_2 = \cdots = c_k = 0$, and our set is indeed independent. $\square$

Definition

An orthogonal basis for a subspace $W < \mathbb{R}^n$ is a basis for W which is an orthogonal set.

Theorem

Let $W < \mathbb{R}^n$, and let $S = \{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k\}$ be an orthogonal basis for W. Then if the vector $\vec{y}$ in W is given in terms of the basis S by

$$\vec{y} = c_1\vec{u}_1 + c_2\vec{u}_2 + \cdots + c_k\vec{u}_k$$

then

$$c_j = \frac{\vec{y} \cdot \vec{u}_j}{\vec{u}_j \cdot \vec{u}_j}$$

Proof.

Since, S is a basis, we know that every vector $\vec{y} \in W$ has a unique representation as a linear combination of vectors in S.
That is, we can write $y = \sum_{i=1}^{k} c_i \vec{u_i}$.
Thus, for $1 \leq j \leq k$, we have

$$\begin{aligned} \vec{u_j} \cdot \vec{y} &= \vec{u_j} \cdot \sum_{i=1}^{k} c_i \vec{u_i} = \sum_{i=1}^{k} c_i \left(\vec{u_j} \cdot \vec{u_i} \right) \\ &= c_j \left(\vec{u_j} \cdot \vec{u_j} \right). \end{aligned}$$

$$\Rightarrow c_j = \frac{\vec{u_j} \cdot \vec{y}}{\vec{u_j} \cdot \vec{u_j}}.$$

$\square$

Example

We saw above that

$$S = \left\{ \vec{v_1} = \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix}, \vec{v_2} = \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix}, \vec{v_3} = \begin{pmatrix} 1 \\ 4 \\ -7 \end{pmatrix} \right\}$$ is an orthogonal set in $\mathbb{R}^3$. Since it is linearly independent and has three vectors in it, it must be a basis for $\mathbb{R}^3$. Express the vector

$$\vec{y} = \begin{pmatrix} 2 \\ 4 \\ 6 \end{pmatrix}$$ in terms of the vectors in S.

Solution

$\vec{y} = c_1\vec{v_1} + c_2\vec{v_2} + c_3\vec{v_3}$, where

$$\begin{aligned}
c_1 &= \frac{\vec{v_1}\cdot\vec{y}}{\vec{v_1}\cdot\vec{v_1}} = \frac{(3)(2)+(1)(4)+(1)(6)}{3^2+1^2+1^2} = \frac{16}{11}\\
c_2 &= \frac{\vec{v_2}\cdot\vec{y}}{\vec{v_2}\cdot\vec{v_2}} = \frac{(-1)(2)+(2)(4)+(1)(6)}{(-1)^2+2^2+1^2} = \frac{12}{6} = 2\\
c_3 &= \frac{\vec{v_3}\cdot\vec{y}}{\vec{v_3}\cdot\vec{v_3}} = \frac{(1)(2)+(4)(4)+(-7)(6)}{1^2+4^2+(-7)^2} = \frac{-24}{66} = \frac{-4}{11}
\end{aligned}$$

Orthogonal Projections

Earlier we saw how to find the component of $\vec{v}$ in the direction of $\vec{u}$ and the component orthogonal to $\vec{u}$. We revisit this idea to introduce some notation, and to extend it to projecting onto a subspace.

Given $\vec{y} \in \mathbb{R}^n$ and $\vec{u} \in \mathbb{R}^n$, find $\hat{y}, \vec{z} \in \mathbb{R}^n$ so that

1. $\vec{y} = \hat{y} + \vec{z}$.
2. $\hat{y} = \alpha\vec{u}$ $(\alpha \in \mathbb{R})$.
3. $\vec{u} \cdot \vec{z} = 0$.

As before, we see that

$$\vec{y} \cdot \vec{u} = \hat{y} \cdot \vec{u} = (\alpha \vec{u}) \cdot \vec{u} = \alpha \|\vec{u}\|^2$$

so

$$\alpha = \frac{\vec{y} \cdot \vec{u}}{\|\vec{u}\|^2}$$

and

$$\vec{z} = \vec{y} - \frac{\vec{y} \cdot \vec{u}}{\|\vec{u}\|^2} \vec{u}$$

Notation

$$\hat{y} = \mathrm{Proj}_L(\vec{y}) = \frac{\vec{y} \cdot \vec{u}}{\|\vec{u}\|^2} \vec{u}$$

where $L = \mathrm{Span}(\vec{u})$.

Terminology

$\vec{y} = \hat{y} + \vec{z}$:
$\hat{y}$ is the orthogonal projection of $\vec{y}$ onto L
$\vec{z}$ is the component of $\vec{y}$ orthogonal to L.

Example

Let $\vec{y} = \begin{pmatrix} 7 \\ 6 \end{pmatrix}$, $\vec{u} = \begin{pmatrix} 4 \\ 2 \end{pmatrix}$ and $L = \text{Span}(\vec{u})$. Compute the orthogonal projection of $\vec{y}$ onto L and the component of $\vec{y}$ orthogonal to L. Plot $\vec{y}, \vec{u}, \hat{y}$ and $\vec{z}$. Compute the distance from $\vec{y}$ to L. (Note: the subspace L here is the line through $\vec{0}$ and $\vec{u}$.)

Let $\{\vec{u}_1, \vec{u}_2\}$ be an orthogonal basis for $\mathbb{R}^2$. Put

$$\hat{y}_1 = \frac{\vec{y} \cdot \vec{u}_1}{\|\vec{u}_1\|^2}\vec{u}_1 = \mathrm{Proj}_{\vec{u}_1}(\vec{y})$$

$$\hat{y}_2 = \frac{\vec{y} \cdot \vec{u}_2}{\|\vec{u}_2\|^2}\vec{u}_2 = \mathrm{Proj}_{\vec{u}_2}(\vec{y})$$

Then

$$\vec{y} = \hat{y}_1 + \hat{y}_2.$$

Definition

Orthonormal Sets A set $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_p$ is called *orthonormal* if it is orthogonal and $\|\vec{u}_i\| = 1$ for $1 \leq i \leq p$. In this case, if $W = \text{Span}(\vec{u}_1, \ldots, \vec{u}_p)$, then the set is called an *orthonormal basis* for W.

Example

The set $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$ is an orthonormal basis for $\mathbb{R}^n$.

Example

Show that the set $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ is an orthornormal basis for $\mathbb{R}^3$, where

$$\vec{v}_1 = \begin{pmatrix} 3/\sqrt{11} \\ 1/\sqrt{11} \\ 1/\sqrt{11} \end{pmatrix}, \quad \vec{v}_2 = \begin{pmatrix} -1/\sqrt{6} \\ 2/\sqrt{6} \\ 1/\sqrt{6} \end{pmatrix}, \quad \vec{v}_3 = \begin{pmatrix} 1/\sqrt{66} \\ 4/\sqrt{66} \\ -7/\sqrt{66} \end{pmatrix}.$$

Theorem

An $m \times n$ matrix U has orthonormal columns if and only if $U^T U = I$.

Proof: proof strategy: interpret the entries of $U^T U$ in terms of inner products of the columns of U.

Theorem

Let U be an $m \times n$ matrix with orthonormal columns, and $\vec{x}, \vec{y} \in \mathbb{R}^n$. Then

1. $\|U\vec{x}\| = \|\vec{x}\|$.
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* $\vec{x} \cdot \vec{y} = 0$.

Note

1. U preserves length
2. U preserves orthonormality.

Proof.

$$(U\vec{x}) \cdot (U\vec{y}) = (U\vec{x})^T(U\vec{y}) = (\vec{x}^T U^T)(U\vec{y}) = \vec{x}^T(U^T U)\vec{y} = \vec{x} \cdot \vec{y}$$

This proves part 2. 1 & 3 follow from 2. □

Example

Let

$$U = \begin{pmatrix} 1/\sqrt{2} & 2/3 \\ 1/\sqrt{2} & -2/3 \\ 0 & 1/3 \end{pmatrix} \qquad \text{and} \qquad \vec{x} = \begin{pmatrix} \sqrt{2} \\ 3 \end{pmatrix}.$$

Check that U has orthonormal columns and that $\|U\vec{x}\| = \|\vec{x}\|$.

Note

In the case where U is a an $n \times n$ matrix with orthonormal columns, we see that $U^T U = I$ so $U^T = U^{-1}$, so $UU^T = I$ as well: that is, U has orthonormal rows too!