April 12 Math 3260 sec. 51 Spring 2024

Section 6.2: Orthogonal Sets

Projection onto a Line

The project of a vector $\mathbf{y}$ onto the line $L = \text{Span}\{\mathbf{u}\}$ for nonzero vector $\mathbf{u}$ is denoted $\text{proj}_L\,\mathbf{y}$ or $\text{proj}_{\mathbf{u}}\,\mathbf{y}$. It is given by

$$\text{proj}_L\,\mathbf{y} = \left(\frac{\mathbf{y}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\right)\mathbf{u}.$$

Given any nonzero vector $\mathbf{u}$ in $\mathbb{R}^n$, each vector $\mathbf{y}$ in $\mathbb{R}^n$ can be written as

$$\mathbf{y} = \text{proj}_{\mathbf{u}}\,\mathbf{y} + \mathbf{z}$$

where $\mathbf{z}$ is orthogonal to $\mathbf{u}$. (Note: if $\mathbf{y}$ is already in $\text{Span}\{\mathbf{u}\}$, then $\mathbf{z}$ will be the zero vector.)

Orthogonal and Orthonormal Sets and Bases

Definition:

An indexed set $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ in $\mathbb{R}^n$ is said to be an **orthogonal set** provided each pair of distinct vectors in the set is orthogonal. That is, provided $\mathbf{u}_i \cdot \mathbf{u}_j = 0 \quad \text{whenever} \quad i \neq j$.

Definition:

A set $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is called an **orthonormal set** if it is an orthogonal set of **unit vectors**.

Definition:

An **orthogonal** (**orthonormal**) basis for a subspace W of $\mathbb{R}^n$ is a basis that is also an orthogonal (respectively orthonormal) set.

Example:

Show that $\left\{ \begin{bmatrix} \frac{3}{5} \\ \frac{4}{5} \end{bmatrix}, \begin{bmatrix} -\frac{4}{5} \\ \frac{3}{5} \end{bmatrix} \right\}$ is an orthonormal basis for $\mathbb{R}^2$.

Call these $\vec{u}_1$ and $\vec{u}_2$ in the order given.

Note that $[\vec{u}_1, \vec{u}_2] \xrightarrow{\text{rref}} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. So $\{\vec{u}_1, \vec{u}_2\}$ is linearly independent. Since there are two vectors and $\dim(\mathbb{R}^2) = 2$, the set is a basis.

$$\vec{u}_1 = \begin{bmatrix} 3/5 \\ 4/5 \end{bmatrix}, \quad \vec{u}_2 = \begin{bmatrix} -4/5 \\ 3/5 \end{bmatrix}.$$

$$\vec{u}_1 \cdot \vec{u}_2 = \frac{3}{5}\left(\frac{-4}{5}\right) + \left(\frac{4}{5}\right)\left(\frac{3}{5}\right) = 0 \Rightarrow \vec{u}_1 \perp \vec{u}_2$$

$$\|\vec{u}_1\|^2 = \left(\frac{3}{5}\right)^2 + \left(\frac{4}{5}\right)^2 = \frac{9}{25} + \frac{16}{25} = \frac{25}{25} = 1$$

$$\|\vec{u}_2\| = \left(\frac{-4}{5}\right)^2 + \left(\frac{3}{5}\right)^2 = 1$$

They're orthogonal unit vectors, so we have an orthonormal basis for $\mathbb{R}^2$.

Orthogonal Matrix

Consider the matrix $U = \begin{bmatrix} \frac{3}{5} & -\frac{4}{5} \\ \frac{4}{5} & \frac{3}{5} \end{bmatrix}$ whose columns are the vectors in the last example. Compute the product

$$U^T U = \begin{bmatrix} \frac{3}{5} & \frac{4}{5} \\ \frac{-4}{5} & \frac{3}{5} \end{bmatrix} \begin{bmatrix} \frac{3}{5} & \frac{-4}{5} \\ \frac{4}{5} & \frac{3}{5} \end{bmatrix} = \begin{bmatrix} \frac{9+16}{25} & \frac{-12+12}{25} \\ \frac{-12+12}{25} & \frac{16+9}{25} \end{bmatrix}$$

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

What does this say about U^{-1}? Recall $U^{-1}U = UU^{-1} = I$

Orthogonal Matrix

Definition:

A square matrix U is called an **orthogonal matrix** if $U^T = U^{-1}$.

Theorem:

An $n \times n$ matrix U is orthogonal if and only if it's columns form an orthonormal basis of $\mathbb{R}^n$.

Remark: The linear transformation associated to an orthogonal matrix preserves *lengths* and *angles* in the sense of the following theorem.

Theorem: Orthogonal Matrices

Theorem

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

(a) $\|U\mathbf{x}\| = \|\mathbf{x}\|$

(b) $(U\mathbf{x}) \cdot (U\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$, in particular

(c) $(U\mathbf{x}) \cdot (U\mathbf{y}) = 0$ if and only if $\mathbf{x} \cdot \mathbf{y} = 0$.

Proof of (a):

Recall: $\|\vec{x}\|^2 = \vec{x}^T \vec{x}$

$(AB)^T = B^T A^T$

Assume U is orthogonal, so $U^T U = I$.

Then
$$\|U\vec{x}\|^2 = (U\vec{x})^T (U\vec{x})$$
$$= \vec{x}^T U^T U \vec{x}$$
$$= \vec{x}^T I \vec{x}$$
$$= \vec{x}^T \vec{x} = \|\vec{x}\|^2.$$

Since $\|U\vec{x}\|$ and $\|\vec{x}\|$ are nonnegative,

$$\|U\vec{x}\| = \|\vec{x}\|.$$

Section 6.3: Orthogonal Projections

Equating points with position vectors, we may wish to find the point $\hat{\mathbf{y}}$ in a subspace W of $\mathbb{R}^n$ that is *closest* to a given point **y**.

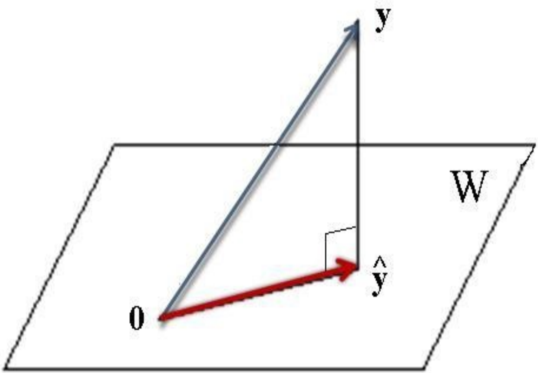

Figure: Illustration of an orthogonal projection. Note that $\text{dist}(\mathbf{y}, \hat{\mathbf{y}})$ is the shortest distance between **y** and the points on W.

Orthogonal Decomposition Theorem

Let W be a subspace of $\mathbb{R}^n$. Each vector $\mathbf{y}$ in $\mathbb{R}^n$ can be written uniquely as a sum

$$\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$$

where $\hat{\mathbf{y}}$ is in W and $\mathbf{z}$ is in $W^\perp$.

If $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is **any orthogonal basis** for W, then

$$\hat{\mathbf{y}} = \sum_{j=1}^{p} \left(\frac{\mathbf{y} \cdot \mathbf{u}_j}{\mathbf{u}_j \cdot \mathbf{u}_j} \right) \mathbf{u}_j, \quad \text{and} \quad \mathbf{z} = \mathbf{y} - \hat{\mathbf{y}}.$$

Remark: The formula for $\hat{\mathbf{y}}$ looks just like the projection onto a line, but with more terms. That is,

$$\hat{\mathbf{y}} = \left(\frac{\mathbf{y} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1} \right) \mathbf{u}_1 + \left(\frac{\mathbf{y} \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2} \right) \mathbf{u}_2 + \cdots + \left(\frac{\mathbf{y} \cdot \mathbf{u}_p}{\mathbf{u}_p \cdot \mathbf{u}_p} \right) \mathbf{u}_p$$

Remarks on the Orthogonal Decomposition Thm.

- Note that the basis must be orthogonal, but otherwise the vector $\hat{\mathbf{y}}$ is **independent** of the particular basis used!

- The vector $\hat{\mathbf{y}}$ is called the **orthogonal projection of y onto** W. We can denote it
$$\text{proj}_W\,\mathbf{y}.$$

- All you really have to do is remember how to project onto a line. Notice that
$$\text{proj}_{\mathbf{u}_1}\,\mathbf{y} = \left(\frac{\mathbf{y}\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}\right)\mathbf{u}_1.$$

  If $W = \text{Span}\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ with the $\mathbf{u}$'s orthogonal, then
$$\text{proj}_W\,\mathbf{y} = \text{proj}_{\mathbf{u}_1}\,\mathbf{y} + \text{proj}_{\mathbf{u}_2}\,\mathbf{y} + \cdots + \text{proj}_{\mathbf{u}_p}\,\mathbf{y}.$$

Example

Let $\mathbf{y} = \begin{bmatrix} 4 \\ 8 \\ 1 \end{bmatrix}$, $\mathbf{u}_1 = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix}$ and $W = \text{Span}\{\mathbf{u}_1, \mathbf{u}_2\}$. (a) Verify that $\{\mathbf{u}_1, \mathbf{u}_2\}$ is an orthogonal basis for W.

We can verify that $\{\vec{u}_1, \vec{u}_2\}$ is linearly independent

We need to show that they are orthogonal.

$$\vec{u}_1 \cdot \vec{u}_2 = 2(-2) + 1(2) + 2(1) = 0.$$

So $\{\vec{u}_1, \vec{u}_2\}$ is an orthogonal basis for W.

Example Continued...

$$W = \text{Span}\left\{ \overset{\vec{u}_1}{\begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}}, \overset{\vec{u}_2}{\begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix}} \right\} \quad \text{and} \quad \mathbf{y} = \begin{bmatrix} 4 \\ 8 \\ 1 \end{bmatrix}$$

(b) Find the orthogonal projection of **y** onto W.

$$\text{proj}_W \vec{y} = \frac{\vec{u}_1 \cdot \vec{y}}{\vec{u}_1 \cdot \vec{u}_1} \vec{u}_1 + \frac{\vec{u}_2 \cdot \vec{y}}{\vec{u}_2 \cdot \vec{u}_2} \vec{u}_2$$

$$\vec{u}_1 \cdot \vec{y} = 2(4) + 1(8) + 2(1) = 18$$

$$\vec{u}_2 \cdot \vec{y} = -2(4) + 2(8) + 1 \cdot 1 = 9$$

$$\vec{u}_1 \cdot \vec{u}_1 = 2^2 + 1^2 + 2^2 = 9$$

$$\vec{u}_2 \cdot \vec{u}_2 = (-2)^2 + 2^2 + 1^2 = 9$$

$$\text{proj}_W \vec{y} = \frac{18}{9}\vec{u}_1 + \frac{9}{9}\vec{u}_2$$

$$= 2\begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix} + 1\begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} 2 \\ 4 \\ 5 \end{bmatrix}$$

(c) Find the shortest distance between **y** and the subspace W.

$$\text{dist}(\vec{y}, W) = \|\vec{y} - \text{proj}_W \vec{y}\|$$

$$= \left\| \begin{bmatrix} 4 \\ 8 \\ 1 \end{bmatrix} - \begin{bmatrix} 2 \\ 4 \\ 5 \end{bmatrix} \right\|$$

$$= \left\| \begin{bmatrix} 2 \\ 4 \\ -4 \end{bmatrix} \right\| = \sqrt{2^2 + 4^2 + (-4)^2}$$

$$= \sqrt{4+16+16} = \sqrt{36} = 6.$$

Note: $\text{proj}_W \vec{y}$ is the point in W closest to $\vec{y}$. So the distance from $\vec{y}$ to W is this $\text{dist}(\vec{y}, \text{proj}_W \vec{y})$.