April 10 Math 3260 sec. 56 Spring 2018

Section 6.2: Orthogonal Sets

Definition: An **orthogonal basis** for a subspace W of $\mathbb{R}^n$ is a basis that is also an orthogonal set.

Theorem: Let $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ be an orthogonal basis for a subspace W of $\mathbb{R}^n$. Then each vector $\mathbf{y}$ in W can be written as the linear combination

$$\mathbf{y} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_p\mathbf{u}_p, \quad \text{where the weights}$$

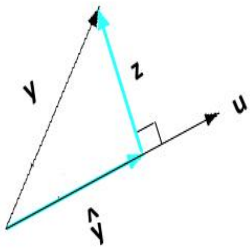
$$c_j = \frac{\mathbf{y} \cdot \mathbf{u}_j}{\mathbf{u}_j \cdot \mathbf{u}_j}.$$

Projection

Given a nonzero vector $\mathbf{u}$, suppose we wish to decompose another nonzero vector $\mathbf{y}$ into a sum of the form

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

in such a way that $\hat{\mathbf{y}}$ is parallel to $\mathbf{u}$ and $\mathbf{z}$ is perpendicular to $\mathbf{u}$.



Projection

Since $\hat{\mathbf{y}}$ is parallel to $\mathbf{u}$, there is a scalar α such that

$$\hat{\mathbf{y}} = \alpha\mathbf{u}.$$

$\vec{y} = \hat{y} + \vec{z} = \alpha\vec{u} + \vec{z}$ we want $\vec{z} \perp \vec{u}$ i.e. $\vec{z}\cdot\vec{u} = 0$

$$\vec{y}\cdot\vec{u} = (\alpha\vec{u} + \vec{z})\cdot\vec{u}$$
$$= \alpha\vec{u}\cdot\vec{u} + \vec{z}\cdot\vec{u}$$
$$= \alpha\vec{u}\cdot\vec{u} + 0 = \alpha\vec{u}\cdot\vec{u}$$

$$\Rightarrow \quad \alpha = \frac{\vec{y}\cdot\vec{u}}{\vec{u}\cdot\vec{u}}$$

Projection onto the subspace $L = \text{Span}\{\mathbf{u}\}$

Notation: $\hat{\mathbf{y}} = \text{proj}_L \mathbf{y} = \left(\dfrac{\mathbf{y} \cdot \mathbf{u}}{\mathbf{u} \cdot \mathbf{u}}\right) \mathbf{u}$

The distance between $\mathbf{y}$ and L is the norm $\|\mathbf{y} - \text{proj}_L\|$.

Example:

Let $\mathbf{y} = \begin{bmatrix} 7 \\ 6 \end{bmatrix}$ and $\mathbf{u} = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$. Write $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$ where $\hat{\mathbf{y}}$ is in Span$\{\mathbf{u}\}$ and $\mathbf{z}$ is orthogonal to $\mathbf{u}$.

$$\hat{y} = \frac{\vec{y}\cdot\vec{u}}{\vec{u}\cdot\vec{u}}\,\vec{u}$$

$$\vec{y}\cdot\vec{u} = 7(4)+6(2) = 40$$

$$\vec{u}\cdot\vec{u} = 4^2+2^2 = 20$$

$$= \frac{40}{20}\begin{bmatrix} 4 \\ 2 \end{bmatrix}$$

$$= 2\begin{bmatrix} 4 \\ 2 \end{bmatrix} = \begin{bmatrix} 8 \\ 4 \end{bmatrix}$$

$$\vec{z} = \vec{y} - \hat{y} = \begin{bmatrix} 7 \\ 6 \end{bmatrix} - \begin{bmatrix} 8 \\ 4 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$$

$$\vec{y} = \underbrace{\begin{bmatrix} 8 \\ 4 \end{bmatrix}}_{\hat{y}} + \underbrace{\begin{bmatrix} -1 \\ 2 \end{bmatrix}}_{\vec{z}}$$

Example Continued...

Determine the distance between the point $(7, 6)$ and the line Span$\{\mathbf{u}\}$.

The distance is $\|\vec{z}\| = \sqrt{(-1)^2 + 2^2} = \sqrt{5}$

Orthonormal Sets

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 **orthonormal basis** of a subspace W of $\mathbb{R}^n$ is a basis that is also an 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 $\mathbb{R}^2$.

To show it's a basis consider $\det \begin{bmatrix} \frac{3}{5} & \frac{-4}{5} \\ \frac{4}{5} & \frac{3}{5} \end{bmatrix} = \frac{9}{25} + \frac{16}{25} = 1 \neq 0$

The columns are lin. independent and span $\mathbb{R}^2$ (Invertible matrix thm).

We have a basis for $\mathbb{R}^2$.

Let $\vec{u}_1 = \begin{bmatrix} 3/5 \\ 4/5 \end{bmatrix}$ and $\vec{u}_2 = \begin{bmatrix} -\frac{4}{5} \\ \frac{3}{5} \end{bmatrix}$

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

$$\vec{u}_2 \cdot \vec{u}_2 = \left(\frac{-4}{5}\right)^2 + \left(\frac{3}{5}\right)^2 = \frac{25}{25} = 1 \Rightarrow \|\vec{u}_2\| = \sqrt{1} = 1$$

$$\vec{u}_1 \cdot \vec{u}_2 = \frac{3}{5}\left(\frac{-4}{5}\right) + \frac{4}{5}\left(\frac{3}{5}\right) = \frac{-12}{25} + \frac{12}{25} = 0$$

They are orthogonal unit vectors. Hence 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

$$U^TU = \begin{bmatrix} \frac{3}{5} & \frac{4}{5} \\ \frac{-4}{5} & \frac{3}{5} \end{bmatrix}\begin{bmatrix} \frac{3}{5} & \frac{-4}{5} \\ \frac{4}{5} & \frac{3}{5} \end{bmatrix} = \begin{bmatrix} \left(\frac{3}{5}\right)^2+\left(\frac{4}{5}\right)^2 & \frac{-12}{25}+\frac{12}{25} \\ \frac{-12}{25}+\frac{12}{25} & \left(\frac{-4}{5}\right)^2+\left(\frac{3}{5}\right)^2 \end{bmatrix}$$

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

What does this say about U^{-1}?

$$U^{-1} = U^T$$

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 $\mathbb{R}^n$.

The linear transformation associated to an orthogonal matrix preserves *lenghts* and *angles* in the following sense:

Theorem: Orthogonal Matrices

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 $\|\vec{x}\|^2 = \vec{x} \cdot \vec{x}$, $\vec{x} \cdot \vec{x} = \vec{x}^T \vec{x}$

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

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

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

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

$$= \vec{x}^T I \vec{x}$$

$$= \vec{x}^T \vec{x}$$

$$= \|\vec{x}\|^2$$

Taking the square root of both sides,

since the norms 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 $\mathbf{y}$.

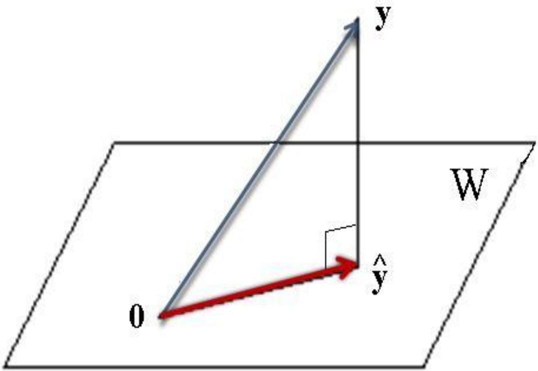


Figure: Illustration of an orthogonal projection. Note that $\text{dist}(\mathbf{y}, \hat{\mathbf{y}})$ is the shortest distance between $\mathbf{y}$ and the points on W.

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

Remark: The vector $\hat{\mathbf{y}}$ is called the **orthogonal projection of y onto** W. We can denote it

$$\mathrm{proj}_W\, \mathbf{y}.$$

Example

Let $\mathbf{y} = \begin{bmatrix} 4 \\ 8 \\ 1 \end{bmatrix}$ and

$$W = \text{Span}\left\{ \overset{\vec{v}_1}{\begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}}, \overset{\vec{v}_2}{\begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix}} \right\}.$$

(a) Verify that the spanning vectors for W given are an orthogonal basis for W.

Noting that $\vec{v}_1 \neq k\vec{v}_2$ for any scalar k, they are linearly independent.

$$\vec{v}_1 \cdot \vec{v}_2 = 2(-2)+1(2)+2(1) = -4+2+2 = 0$$

They are orthogonal.

Example Continued...

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

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

$$\hat{y} = \frac{\vec{y}\cdot\vec{v}_1}{\vec{v}_1\cdot\vec{v}_1}\vec{v}_1 + \frac{\vec{y}\cdot\vec{v}_2}{\vec{v}_2\cdot\vec{v}_2}\vec{v}_2$$

$$\vec{y}\cdot\vec{v}_1 = 2\cdot 4 + 1\cdot 8 + 2\cdot 1 = 18 \qquad \vec{v}_1\cdot\vec{v}_1 = 2^2 + 1^2 + 2^2 = 9$$

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

$$\hat{y} = \frac{18}{9}\vec{v}_1 + \frac{9}{9}\vec{v}_2$$

$$= 2\begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix} + 1\begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \\ 5 \end{bmatrix}$$

(c) Find the shortest distance between $\mathbf{y}$ and the subspace W.

If $\vec{y} = \hat{y} + \vec{z}$ then the distance is

$$\|\vec{z}\| = \|\vec{y} - \hat{y}\|.$$

$$\vec{z} = \vec{y} - \hat{y} = \begin{bmatrix} 4 \\ 8 \\ 1 \end{bmatrix} - \begin{bmatrix} 2 \\ 4 \\ 5 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \\ -4 \end{bmatrix}$$

$$\|\vec{z}\| = \sqrt{2^2 + 4^2 + (-4)^2} = \sqrt{36} = 6$$

Computing Orthogonal Projections

Theorem: If $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is an orthonormal basis of a subspace W of $\mathbb{R}^n$, and $\mathbf{y}$ is any vector in $\mathbb{R}^n$ then

$$\hat{y} = \quad \text{proj}_W\, \mathbf{y} = \sum_{j=1}^{p} (\mathbf{y} \cdot \mathbf{u}_j)\, \mathbf{u}_j.$$

And, if U is the matrix $U = [\mathbf{u}_1 \quad \cdots \quad \mathbf{u}_p]$, then the above is equivalent to

$$\text{proj}_W\, \mathbf{y} = UU^T\mathbf{y}.$$

Remark: In general, U is not square; it's $n \times p$. So even though UU^T will be a square matrix, it is not the same matrix as U^TU and it is not the identity matrix.

Example

$$W = \text{Span}\left\{ \overset{\vec{v}_1}{\begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}}, \overset{\vec{v}_2}{\begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix}} \right\}$$

Find an orthonormal basis $\{\mathbf{u}_1, \mathbf{u}_2\}$ for W. Then compute the matrices $U^T U$ and UU^T where $U = [\mathbf{u}_1 \;\; \mathbf{u}_2]$.

$$\vec{u}_1 = \frac{1}{\|\vec{v}_1\|}\vec{v}_1 = \frac{1}{3}\begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 2/3 \\ 1/3 \\ 2/3 \end{bmatrix}$$

$$\vec{u}_2 = \frac{1}{\|\vec{v}_2\|}\vec{v}_2 = \frac{1}{3}\begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} -2/3 \\ 2/3 \\ 1/3 \end{bmatrix}$$

$$U = \begin{bmatrix} 2/3 & -2/3 \\ 1/3 & 2/3 \\ 2/3 & 1/3 \end{bmatrix}$$

$$U^TU = \frac{1}{3}\begin{bmatrix} 2 & 1 & 2 \\ -2 & 2 & 1 \end{bmatrix} \frac{1}{3}\begin{bmatrix} 2 & -2 \\ 1 & 2 \\ 2 & 1 \end{bmatrix} = \frac{1}{9}\begin{bmatrix} 9 & 0 \\ 0 & 9 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$UU^T = \frac{1}{3}\begin{bmatrix} 2 & -2 \\ 1 & 2 \\ 2 & 1 \end{bmatrix} \frac{1}{3}\begin{bmatrix} 2 & 1 & 2 \\ -2 & 2 & 1 \end{bmatrix} = \frac{1}{9}\begin{bmatrix} 8 & -2 & 2 \\ -2 & 5 & 4 \\ 2 & 4 & 5 \end{bmatrix}$$

Example

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

Use the matrix formulation to find $\text{proj}_W\, \mathbf{y}$.

$$\hat{y} = UU^T\vec{y} = \frac{1}{9}\begin{bmatrix} 8 & -2 & 2 \\ -2 & 5 & 4 \\ 2 & 4 & 5 \end{bmatrix}\begin{bmatrix} 4 \\ 8 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \\ 5 \end{bmatrix}$$