6.3 & 6.4: Orthogonal Projections & Gram-Schmidt
Math 220: Linear Algebra

Given a vector $\mathbf{y}$ and a subspace W in $\mathbb{R}^n$ there is a vector $\hat{\mathbf{y}} \in W$ such that

1) $\hat{\mathbf{y}}$ is the unique vector in *W* for which $\mathbf{y}-\hat{\mathbf{y}}$ is orthogonal to *W*

2) $\hat{\mathbf{y}}$ is the unique vector in *W* closest to $\mathbf{y}$

Theorem 8 The Orthogonal Decomposition Theorem

Let *W* be a subspace of $\mathbb{R}^n$. Then each **y** in $\mathbb{R}^n$ can be written uniquely in the form

$$\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z} \qquad (1)$$

where $\hat{\mathbf{y}}$ is in *W* and **z** is in $W^\perp$. In fact, if $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is any orthogonal basis of *W*, then

$$\hat{\mathbf{y}} = \frac{\mathbf{y}\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}\mathbf{u}_1 + \cdots + \frac{\mathbf{y}\cdot\mathbf{u}_p}{\mathbf{u}_p\cdot\mathbf{u}_p}\mathbf{u}_p \qquad (2)$$

and $\mathbf{z} = \mathbf{y} - \hat{\mathbf{y}}$.

Ex 1: Let $W = \text{Span}\{\mathbf{u}_1, \mathbf{u}_2\}$. Write $\mathbf{y}$ as the sum of a vector in W and a vector orthogonal to W.

$$\mathbf{y} = \begin{bmatrix}1\\3\\5\end{bmatrix}, \mathbf{u}_1 = \begin{bmatrix}1\\3\\-2\end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix}5\\1\\4\end{bmatrix}$$

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

$$= \frac{0}{14}\begin{bmatrix}1\\3\\-2\end{bmatrix} + \frac{28}{42}\begin{bmatrix}5\\1\\4\end{bmatrix}$$

$$= \begin{bmatrix}10/3\\2/3\\8/3\end{bmatrix}$$

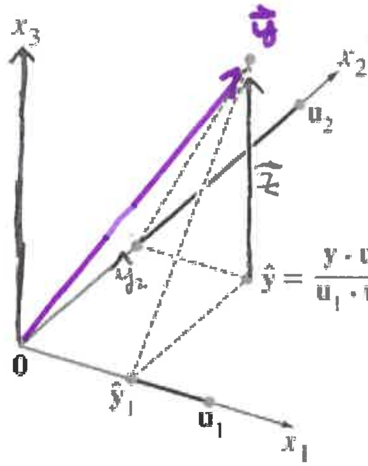
$$\vec{z} = \vec{y} - \hat{y}$$

$$= \begin{bmatrix}1\\3\\5\end{bmatrix} - \begin{bmatrix}10/3\\2/3\\8/3\end{bmatrix}$$

$$= \begin{bmatrix}-7/3\\7/3\\7/3\end{bmatrix}$$

$$\therefore \vec{y} = \underbrace{\overset{\hat{y}}{\begin{bmatrix}10/3\\2/3\\8/3\end{bmatrix}}}_{\text{vector in } W} + \underbrace{\overset{\vec{z}}{\begin{bmatrix}-7/3\\7/3\\7/3\end{bmatrix}}}_{\text{vector in } W^\perp}$$

6.3 & 6.4: Orthogonal Projections & Gram-Schmidt



$$\hat{\mathbf{y}} = \frac{\mathbf{y}\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}\mathbf{u}_1 + \frac{\mathbf{y}\cdot\mathbf{u}_2}{\mathbf{u}_2\cdot\mathbf{u}_2}\mathbf{u}_2 = \hat{\mathbf{y}}_1 + \hat{\mathbf{y}}_2$$

Spoiler: $\vec{y} \approx \hat{y}$ and the error in the approximation is $\vec{z}$.

Theorem 9 The Best Approximation Theorem

Let W be a subspace of $\mathbb{R}^n$, let $\mathbf{y}$ be any vector in $\mathbb{R}^n$, and let $\hat{\mathbf{y}}$ be the orthogonal projection of $\mathbf{y}$ onto W. Then $\hat{\mathbf{y}}$ is the closest point in W to $\mathbf{y}$, in the sense that

$$\|\mathbf{y} - \hat{\mathbf{y}}\| < \|\mathbf{y} - \mathbf{v}\| \qquad (3)$$

for all $\mathbf{v}$ in W distinct from $\hat{\mathbf{y}}$.

Ex 2: As in Ex 1, $\begin{bmatrix} 10/3 \\ 2/3 \\ 8/3 \end{bmatrix}$ is the closest point in $W = \text{Span}\left\{\mathbf{u}_1 = \begin{bmatrix} 1 \\ 3 \\ -2 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} 5 \\ 1 \\ 4 \end{bmatrix}\right\}$ to $\mathbf{y} = \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix}$.

Find the distance from $\mathbf{y}$ to W

The distance from $\vec{y}$ to $W = \|\vec{z}\|$

$$= \sqrt{\frac{49}{9} + \frac{49}{9} + \frac{49}{9}}$$

$$= \frac{7\sqrt{3}}{3}$$

6.3 & 6.4: Orthogonal Projections & Gram-Schmidt

Practice Problems

1. Let $\mathbf{u}_1 = \begin{bmatrix} -7 \\ 1 \\ 4 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} -1 \\ 1 \\ -2 \end{bmatrix}$, $\mathbf{y} = \begin{bmatrix} -9 \\ 1 \\ 6 \end{bmatrix}$, and $W = \text{Span}\{\mathbf{u}_1, \mathbf{u}_2\}$.
Use the fact that $\mathbf{u}_1$ and $\mathbf{u}_2$ are orthogonal to compute $\text{proj}_W\, \mathbf{y}$.

$$\text{proj}_W \vec{y} = \hat{y} = \frac{\vec{y} \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1} \vec{u}_1 + \frac{\vec{y} \cdot \vec{u}_2}{\vec{u}_2 \cdot \vec{u}_2} \vec{u}_2 = \frac{38}{66} \begin{bmatrix} -7 \\ 1 \\ 4 \end{bmatrix} + \frac{-2}{6} \begin{bmatrix} -1 \\ 1 \\ -2 \end{bmatrix} = \begin{bmatrix} -9 \\ 1 \\ 6 \end{bmatrix}$$

notice that $\hat{y} = \vec{y}$
which means $\vec{z} = 0$,

Phrased another way
$\vec{y} \in W$

2. Let W be the subspace spanned by the **u**'s, and write **y** as the sum of a vector in W and a vector orthogonal to W.

$$\mathbf{y} = \begin{bmatrix} 4 \\ 3 \\ 3 \\ -1 \end{bmatrix}, \mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} -1 \\ 3 \\ 1 \\ -2 \end{bmatrix}, \mathbf{u}_3 = \begin{bmatrix} -1 \\ 0 \\ 1 \\ 1 \end{bmatrix}$$

$$\hat{y} = \frac{\vec{y} \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1} \vec{u}_1 + \frac{\vec{y} \cdot \vec{u}_2}{\vec{u}_2 \cdot \vec{u}_2} \vec{u}_2 + \frac{\vec{y} \cdot \vec{u}_3}{\vec{u}_3 \cdot \vec{u}_3} \vec{u}_3$$

$$= \frac{6}{3} \vec{u}_1 + \frac{10}{15} \vec{u}_2 + \frac{-2}{3} \vec{u}_3$$

$$= \begin{bmatrix} 2 - 2/3 + 2/3 \\ 2 + 2 + 0 \\ 0 + 2/3 - 2/3 \\ 2 - 4/3 - 2/3 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \\ 0 \\ 0 \end{bmatrix}$$

$$\rightarrow \quad z = \vec{y} - \hat{y} = \begin{bmatrix} 4 \\ 3 \\ 3 \\ -1 \end{bmatrix} - \begin{bmatrix} 2 \\ 4 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \\ 3 \\ -1 \end{bmatrix}$$

AND $\vec{y} = \underbrace{\begin{bmatrix} 2 \\ 4 \\ 0 \\ 0 \end{bmatrix}}_{\hat{y}} + \underbrace{\begin{bmatrix} 2 \\ -1 \\ 3 \\ -1 \end{bmatrix}}_{\vec{z}}$

vector in W vector in $W^{\perp}$

6.3 & 6.4: Orthogonal Projections & Gram-Schmidt

The Gram-Schmidt Process

Basis you are given

Basis you want to find/construct

Ex 3: Let $W = \text{Span}\left\{ \mathbf{x}_1 = \begin{bmatrix} 3 \\ 0 \\ -1 \end{bmatrix}, \mathbf{x}_2 = \begin{bmatrix} 8 \\ 5 \\ -6 \end{bmatrix} \right\}$, construct an orthogonal basis $\{\mathbf{v}_1, \mathbf{v}_2\}$.

$$\vec{v}_1 = \vec{x}_1$$

$\vec{v}_2 = \vec{x}_2 - \hat{x}_2$ where $\hat{x}_2$ is the orthogonal projection of $\vec{x}_2$ onto $\vec{v}_1$

$$= \vec{x}_2 - \frac{\vec{x}_2 \cdot \vec{v}_1}{\vec{v}_1 \cdot \vec{v}_1}\vec{v}_1 = \begin{bmatrix} 8 \\ 5 \\ -6 \end{bmatrix} - \frac{30}{10}\begin{bmatrix} 3 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} -1 \\ 5 \\ -3 \end{bmatrix}$$

$\therefore$ An orthogonal basis for W is:

$$\left\{ \overset{\vec{v}_1}{\begin{bmatrix} 3 \\ 0 \\ -1 \end{bmatrix}}, \overset{\vec{v}_2}{\begin{bmatrix} -1 \\ 5 \\ -3 \end{bmatrix}} \right\}$$

Ex 4:

Let $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{x}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}$, and $\mathbf{x}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}$. Then $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ is clearly linearly independent and thus is a basis for a subspace W of $\mathbb{R}^4$.
Construct an orthogonal basis for W. $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$

$$\vec{v}_1 = \vec{x}_1$$

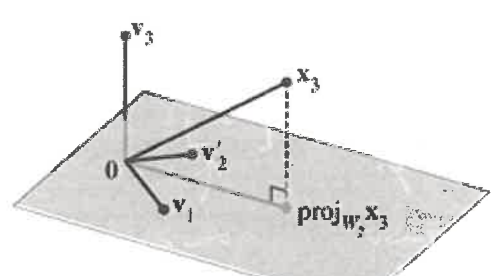
key: use the given basis to iteratively construct an orthogonal basis.

$$\vec{v}_2 = \vec{x}_2 - \frac{\vec{x}_2 \cdot \vec{v}_1}{\vec{v}_1 \cdot \vec{v}_1}\vec{v}_1 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} - \frac{3}{4}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -3/4 \\ 1/4 \\ 1/4 \\ 1/4 \end{bmatrix}$$

$$\vec{v}_3 = \vec{x}_3 - \frac{\vec{x}_3 \cdot \vec{v}_1}{\vec{v}_1 \cdot \vec{v}_1}\vec{v}_1 - \frac{\vec{x}_3 \cdot \vec{v}_2}{\vec{v}_2 \cdot \vec{v}_2}\vec{v}_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} - \frac{2}{4}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} - \overbrace{\frac{2/4}{12/16}}^{1/6} \cdot \frac{1}{4}\begin{bmatrix} -3 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ -2/3 \\ 1/3 \\ 1/3 \end{bmatrix}$$

$\therefore$ An orthogonal basis for W is:

$$\left\{ \overset{\vec{v}_1}{\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}}, \overset{\vec{v}_2}{\frac{1}{4}\begin{bmatrix} -3 \\ 1 \\ 1 \\ 1 \end{bmatrix}}, \overset{\vec{v}_3}{\frac{1}{3}\begin{bmatrix} 0 \\ -2 \\ 1 \\ 1 \end{bmatrix}} \right\}$$



$W_2 = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2'\}$

Note: This orthogonal basis is not unique (the only orthogonal basis for W).

6.3 & 6.4: Orthogonal Projections & Gram-Schmidt

Theorem 11 The Gram–Schmidt Process
Given a basis $\{\mathbf{x}_1, \ldots, \mathbf{x}_p\}$ for a nonzero subspace *W* of $\mathbb{R}^n$, define

$$\begin{aligned}
\mathbf{v}_1 &= \mathbf{x}_1 \\
\mathbf{v}_2 &= \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 \\
\mathbf{v}_3 &= \mathbf{x}_3 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\mathbf{v}_2 \\
&\vdots \\
\mathbf{v}_p &= \mathbf{x}_p - \frac{\mathbf{x}_p \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 - \frac{\mathbf{x}_p \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\mathbf{v}_2 - \cdots - \frac{\mathbf{x}_p \cdot \mathbf{v}_{p-1}}{\mathbf{v}_{p-1} \cdot \mathbf{v}_{p-1}}\mathbf{v}_{p-1}
\end{aligned}$$

Then $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ is an orthogonal basis for *W*. In addition

$$\text{Span}\,\{\mathbf{v}_1, \ldots, \mathbf{v}_k\} = \text{Span}\,\{\mathbf{x}_1, \ldots, \mathbf{x}_k\} \quad \text{for } 1 \le k \le p$$

The result of this is that every nonzero subspace W in $\mathbb{R}^n$ has an orthogonal basis.

An orthonormal basis is constructed easily by normalizing all the $\mathbf{v}_k$'s to unit vectors.

$\perp$ length 1

Ex 5: Re-write the orthogonal basis found in Ex 3 as an orthonormal basis.

The orthogonal basis we found was: $\left\{ \overset{\vec{v}_1}{\begin{bmatrix} 3 \\ 0 \\ -1 \end{bmatrix}}, \overset{\vec{v}_2}{\begin{bmatrix} -1 \\ 5 \\ -3 \end{bmatrix}} \right\}$

$\vec{v}_1 \perp \vec{v}_2$ by construction. We just need to scale them to unit vectors.

$\|\vec{v}_1\| = \sqrt{10}$ and $\|\vec{v}_2\| = \sqrt{35}$, so an orthonormal basis is: $\left\{ \begin{bmatrix} 3/\sqrt{10} \\ 0 \\ -1/\sqrt{10} \end{bmatrix}, \begin{bmatrix} -1/\sqrt{35} \\ 5/\sqrt{35} \\ -3/\sqrt{35} \end{bmatrix} \right\}$

using the same method, an orthonormal basis for the space in ex 4 is: $\left\{ \frac{1}{2}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \frac{1}{2\sqrt{3}}\begin{bmatrix} -3 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \frac{1}{\sqrt{6}}\begin{bmatrix} 0 \\ -2 \\ 1 \\ 1 \end{bmatrix} \right\}$

6.3 & 6.4: Orthogonal Projections & Gram-Schmidt

Practice Problems

orthogonal basis since $\frac{\vec{x}_1 \cdot \vec{x}_2}{\vec{x}_1 \cdot \vec{x}_1} = 0$

1. Let $W = \text{Span}\{\mathbf{x}_1, \mathbf{x}_2\}$, where $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ and $\mathbf{x}_2 = \begin{bmatrix} 1/3 \\ 1/3 \\ -2/3 \end{bmatrix}$

Construct an orthonormal basis for W.

$\|\vec{x}_1\| = \sqrt{3}$ $\qquad \|\vec{x}_2\| = \frac{\sqrt{6}}{3}$

orthonormal basis

$$\left\{ \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \frac{3}{\sqrt{6}} \begin{bmatrix} 1/3 \\ 1/3 \\ -2/3 \end{bmatrix} \right\}$$

we can organize all this with

$$\begin{bmatrix} 1 & 1/3 \\ 1 & 1/3 \\ 1 & -2/3 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{6} \\ 1/\sqrt{3} & 1/\sqrt{6} \\ 1/\sqrt{3} & -2/\sqrt{6} \end{bmatrix} \begin{bmatrix} \sqrt{3} & 0 \\ 0 & \frac{\sqrt{6}}{3} \end{bmatrix}$$

$$A = Q \quad R$$

A: columns of A are linearly independent

Q: The columns of Q are an orthonormal basis for col A

R: R is an upper triangular matrix

2. Use the Gram–Schmidt process to produce an orthogonal basis for W.

$$W = \text{Span}\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\} \text{ where } \mathbf{x}_1 = \begin{bmatrix} -1 \\ 3 \\ 1 \\ 1 \end{bmatrix}, \mathbf{x}_2 = \begin{bmatrix} 6 \\ -8 \\ -2 \\ -4 \end{bmatrix}, \mathbf{x}_3 = \begin{bmatrix} 6 \\ 3 \\ 6 \\ -3 \end{bmatrix}$$

$$\vec{v}_1 = \vec{x}_1 = \begin{bmatrix} -1 \\ 3 \\ 1 \\ 1 \end{bmatrix} \quad \text{and } \|\vec{v}_1\| = \sqrt{12}$$

$$\vec{v}_2 = \vec{x}_2 - \frac{\vec{x}_2 \cdot \vec{v}_1}{\vec{v}_1 \cdot \vec{v}_1} \vec{v}_1 = \begin{bmatrix} 6 \\ -8 \\ -2 \\ -4 \end{bmatrix} - \frac{-36}{12} \begin{bmatrix} -1 \\ 3 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \\ 1 \\ -1 \end{bmatrix} \quad \text{and } \|\vec{v}_2\| = \sqrt{12}$$

$$\vec{v}_3 = \vec{x}_3 - \frac{\vec{x}_3 \cdot \vec{v}_1}{\vec{v}_1 \cdot \vec{v}_1} \vec{v}_1 - \frac{\vec{x}_3 \cdot \vec{v}_2}{\vec{v}_2 \cdot \vec{v}_2} \vec{v}_2 = \begin{bmatrix} 6 \\ 3 \\ 6 \\ -3 \end{bmatrix} - \frac{6}{12} \begin{bmatrix} -1 \\ 3 \\ 1 \\ 1 \end{bmatrix} - \frac{30}{12} \begin{bmatrix} 3 \\ 1 \\ 1 \\ -1 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \\ 3 \\ -1 \end{bmatrix}$$

and $\|\vec{v}_3\|$ is $\sqrt{12}$

This can be organized w/ a QR Factorization.

$$\begin{bmatrix} -1 & 6 & 6 \\ 3 & -8 & 3 \\ 1 & -2 & 6 \\ 1 & -4 & -3 \end{bmatrix} = \begin{bmatrix} -1/\sqrt{12} & 3/\sqrt{12} & -1/\sqrt{12} \\ 3/\sqrt{12} & 1/\sqrt{12} & -1/\sqrt{12} \\ 1/\sqrt{12} & 1/\sqrt{12} & 3/\sqrt{12} \\ 1/\sqrt{12} & -1/\sqrt{12} & -1/\sqrt{12} \end{bmatrix} \begin{bmatrix} \sqrt{12} & -3\sqrt{12} & \frac{1}{2}\sqrt{12} \\ 0 & \sqrt{12} & \frac{5}{2}\sqrt{12} \\ 0 & 0 & \sqrt{12} \end{bmatrix}$$