MATH 311 - Orthogonality

Jasraj Sandhu

June 2023

Orthogonal Complements and Projections

Review

Recall: Two vectors $\vec{v}, \vec{w}$ are orthogonal if $\vec{v} \cdot \vec{w} = 0$.

Recall: A set of vectors $\{\vec{v_1}, \ldots, \vec{v_k}\}$ is orthogonal if each $\vec{v}_i \neq 0$ and $\vec{v_i} \cdot \vec{v_j} = 0$ for all $i \neq j$.

Theorem 13

If $\{\vec{v_1}, \ldots, \vec{v_k}\}$ is an orthogonal set in $\mathbb{R}^n$, then $\{\vec{v_1}, \ldots, \vec{v_k}\}$ is linearly independent.

Proof. Suppose $\{\vec{v_1}, \ldots, \vec{v_k}\}$ is orthogonal. We want to show that $c_1\vec{v_1} + \ldots + c_k\vec{v_k} = \vec{0}$, where $c_1, \ldots, c_k \in \mathbb{R}$ and $c_1 = \ldots = c_k = 0$. $\square$

Fourier Expansion Theorem

Let $\{\vec{f}_1, \ldots, \vec{f}_k\}$ be an **orthogonal** basis for a subspace U of $\mathbb{R}^n$. For any $\vec{u} \in U$, we have

$$\vec{u} = \left(\frac{\vec{u} \cdot \vec{f}_1}{||\vec{f}_1||^2} \right) \vec{f}_1 + \ldots + \left(\frac{\vec{u} \cdot \vec{f}_k}{||\vec{f}_k||^2} \right) \vec{f}_k \ .$$

Note that

$$\frac{\vec{u} \cdot \vec{f}_1}{||\vec{f}_1||^2}, \ldots, \frac{\vec{u} \cdot \vec{f}_k}{||\vec{f}_k||^2}$$

are Fourier coefficients.

Note: Notice that each term being summed is the projection of $\vec{u}$ onto $\vec{f}_i$ for all $1 \leq i \leq k$.

Gram Schmidt Algorithm

We use this algorithm to orthogonalize a basis.

Gram-Schmidt Algorithm

Let $\{\vec{u}_1, \ldots, \vec{u}_k\}$ be a basis for a subspace U of $\mathbb{R}^n$. We construct an orthogonal basis $\{\vec{f}_1, \ldots, \vec{f}_k\}$ as follows:

$$\begin{aligned}
\vec{f}_1 &= \vec{u}_1 \\
\vec{f}_2 &= \vec{u}_2 - \left(\frac{\vec{u}_2 \cdot \vec{f}_1}{||\vec{f}_1||^2}\right)\vec{f}_1 \\
\vec{f}_3 &= \vec{u}_3 - \left(\frac{\vec{u}_3 \cdot \vec{f}_1}{||\vec{f}_1||^2}\right)\vec{f}_1 - \left(\frac{\vec{u}_3 \cdot \vec{f}_2}{||\vec{f}_2||^2}\right)\vec{f}_2 \\
\vec{f}_k &= \vec{u}_k - \left(\frac{\vec{u}_k \cdot \vec{f}_1}{||\vec{f}_1||^2}\right)\vec{f}_1 - \ldots - \left(\frac{\vec{u}_k \cdot \vec{f}_{k-1}}{||\vec{f}_{k-1}||^2}\right)\vec{f}_{k-1}
\end{aligned}$$

1. Consider the basis $\left\{ \begin{bmatrix}1\\1\\1\end{bmatrix}, \begin{bmatrix}0\\1\\1\end{bmatrix}, \begin{bmatrix}0\\0\\1\end{bmatrix} \right\}$ of $\mathbb{R}^3$.

   (a) Construct an orthogonal basis for $\mathbb{R}^3$.

   We use the Gram-Schmidt Algorithm to construct an orthogonal basis using the given basis for $\mathbb{R}^3$. Let

$$\vec{v_1} = \begin{bmatrix}1\\1\\1\end{bmatrix}, \vec{v_2} = \begin{bmatrix}0\\1\\1\end{bmatrix}, \vec{v_3} = \begin{bmatrix}0\\0\\1\end{bmatrix}.$$

   Then

$$\begin{aligned}
\vec{f_1} = \vec{v}_1 &= \begin{bmatrix}1\\1\\1\end{bmatrix}, \\
\vec{f_2} = \vec{v}_2 &- \left(\frac{\vec{v_2} \cdot \vec{f_1}}{||\vec{f_1}||^2}\right)\vec{f_1} \\
&= \begin{bmatrix}0\\1\\1\end{bmatrix} - \frac{2}{3}\begin{bmatrix}1\\1\\1\end{bmatrix} \\
&= \begin{bmatrix}-2/3\\1/3\\1/3\end{bmatrix},
\end{aligned}$$

$$
\begin{aligned}
\vec{f}_3 &= \vec{v}_3 - \left(\frac{\vec{v}_3 \cdot \vec{f}_1}{||\vec{f}_1||^2}\right)\vec{f}_1 - \left(\frac{\vec{v}_3 \cdot \vec{f}_2}{||\vec{f}_2||^2}\right)\vec{f}_2 \\
&= \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} - \frac{1}{3}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \frac{1/3}{6/9}\begin{bmatrix} -2/3 \\ 1/3 \\ 1/3 \end{bmatrix} \\
&= \begin{bmatrix} 0 \\ -1/2 \\ 1/2 \end{bmatrix} .
\end{aligned}
$$

So, $\{\vec{f}_1, \vec{f}_2, \vec{f}_3\} = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -2/3 \\ 1/3 \\ 1/3 \end{bmatrix}, \begin{bmatrix} 0 \\ -1/2 \\ 1/2 \end{bmatrix} \right\}$ is an orthogonal basis for $\mathbb{R}^3$.

(b) To produce an orthonormal basis, we normalize the vectors in that form the orthogonal basis. So, we normalize each of $\vec{f}_1$, $\vec{f}_2$, and $\vec{f}_3$. Then we get that

$$
||\vec{f}_1|| = \sqrt{1^2 + 1^2 + 1^2} = \sqrt{3} \,,
$$

$$
\begin{aligned}
||\vec{f}_2|| &= \sqrt{\left(-\frac{2}{3}\right)^2 + \left(\frac{1}{3}\right)^2 + \left(\frac{1}{3}\right)^2} \\
&= \sqrt{\frac{4}{9} + \frac{1}{9} + \frac{1}{9}} \\
&= \sqrt{\frac{6}{9}} \\
&= \sqrt{\frac{2}{3}} \\
&= \frac{\sqrt{2}}{\sqrt{3}} \,,
\end{aligned}
$$

$$
\begin{aligned}
||\vec{f}_3|| &= \sqrt{0^2 + \left(-\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2} \\
&= \sqrt{0 + \frac{1}{4} + \frac{1}{4}} \\
&= \sqrt{\frac{1}{2}} \\
&= \frac{1}{\sqrt{2}} \,.
\end{aligned}
$$

Hence,

$$\left\{ \frac{1}{||\vec{f}_1||}\vec{f}_1, \frac{1}{||\vec{f}_2||}\vec{f}_2, \frac{1}{||\vec{f}_3||}\vec{f}_3, \right\} = \left\{ \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \frac{\sqrt{3}}{\sqrt{2}} \begin{bmatrix} -2/3 \\ 1/3 \\ 1/3 \end{bmatrix}, \sqrt{2} \begin{bmatrix} 0 \\ -1/2 \\ 1/2 \end{bmatrix} \right\}$$
$$= \left\{ \begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}, \begin{bmatrix} \frac{-2\sqrt{3}}{3\sqrt{2}} \\ \frac{\sqrt{3}}{3\sqrt{2}} \\ \frac{\sqrt{3}}{3\sqrt{2}} \end{bmatrix}, \begin{bmatrix} 0 \\ \frac{-\sqrt{2}}{2} \\ \frac{\sqrt{3}}{3\sqrt{2}} \end{bmatrix} \right\}$$

is an orthonormal basis for $\mathbb{R}^3$.

Theorem (Orthogonal Lemma)

Let $\{\vec{f}_1, \ldots, \vec{f}_m\}$ be an **orthonormal** subset of $\mathbb{R}^n$. Given $\vec{x} \in \mathbb{R}^n$, define:

$$\vec{f}_{m+1} := \vec{x} - \left(\frac{\vec{x} \cdot \vec{f}_1}{||\vec{f}_1||^2} \right) \vec{f}_1 - \ldots - \left(\frac{\vec{x} \cdot \vec{f}_m}{||\vec{f}_m||^2} \right) \vec{f}_m \ .$$

If $\vec{x} \notin \operatorname{span}\{\vec{f}_1, \ldots, \vec{f}_m\}$, then $\{\vec{f}_1, \ldots, \vec{f}_m, \vec{f}_{m+1}\}$ is orthogonal.

Proof. Suppose $\{\vec{f}_1, \ldots, \vec{f}_m\}$ is an orthonormal subset of $\mathbb{R}^n$. For each $1 \leq k \leq m$, we have that

$$\begin{aligned}
\vec{f}_{m+1} \cdot \vec{f}_k &= \left(\vec{x} - \left(\frac{\vec{x} \cdot \vec{f}_1}{||\vec{f}_1||^2} \right) \vec{f}_1 - \ldots - \left(\frac{\vec{x} \cdot \vec{f}_k}{||\vec{f}_k||^2} \right) \vec{f}_k \right) \cdot \vec{f}_k \\
&= \vec{x} \cdot \vec{f}_k - \left(\frac{\vec{x} \cdot \vec{f}_1}{||\vec{f}_1||^2} \right) \vec{f}_1 \cdot \vec{f}_k - \ldots - \left(\frac{\vec{x} \cdot \vec{f}_k}{||\vec{f}_k||^2} \right) \vec{f}_k \cdot \vec{f}_k \\
&= \vec{x} \cdot \vec{f}_k - \left(\frac{\vec{x} \cdot \vec{f}_k}{||\vec{f}_k||^2} \right) \vec{f}_k \cdot \vec{f}_k \ ,
\end{aligned}$$

since $\vec{f}_i \cdot \vec{f}_j = 0$ for all $i \neq j$.. So,

$$\begin{aligned}
\vec{f}_{m+1} \cdot \vec{f}_k &= \vec{x} \cdot \vec{f}_k - \left(\frac{\vec{x} \cdot \vec{f}_k}{||\vec{f}_k||^2} \right) \vec{f}_k \cdot \vec{f}_k \\
&= \vec{x} \cdot \vec{f}_k - \frac{\vec{x} \cdot \vec{f}_k}{||\vec{f}_k||^2} (\vec{f}_k \cdot \vec{f}_k) \\
&= \vec{x} \cdot \vec{f}_k - \frac{\vec{x} \cdot \vec{f}_k}{||\vec{f}_k||^2} \ ||\vec{f}_k||^2 \\
&= \vec{x} \cdot \vec{f}_k - \vec{x} \cdot \vec{f}_k \\
&= 0 \ .
\end{aligned}$$

Moreover, $\vec{f}_{m+1} \neq \vec{0}$, since $\vec{x} \notin \operatorname{span}\{\vec{f}_1, \ldots, \vec{f}_m\}$. $\square$

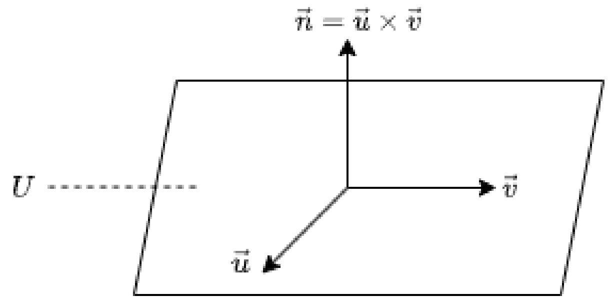
Orthogonal Complements

Definition (Orthogonal Complement)

Let U be a subspace of $\mathbb{R}^n$. The orthogonal complement of U is

$$U^{\perp} = \{\vec{x} \in \mathbb{R}^n : \vec{x} \cdot \vec{u} = 0 \ \ \forall \vec{u} \in U\} \ .$$

1. Let $U \subseteq \mathbb{R}^3$ be a plane through $\vec{0}$ (the origin). Then $U^{\perp}$ is a line through $\vec{0}$ in the direction of the **normal vector** of the plane.



Here we have the following:

- U is the plane. So, $U = \text{span}\{\vec{u}, \vec{v}\}$.
- $\vec{n} = \vec{u} \times \vec{v}$ is the normal vector (the vector perpendicular to the plane).
- $U^{\perp} = \text{span}\{\vec{n}\} = \{c\vec{n} : c \in \mathbb{R}\}$, where $c \in \mathbb{R}$. That is, $U^{\perp}$ is the set of all linear combinations (in this case, scalar multiples) of the normal vector $\vec{n}$ perpendicular/orthogonal to the plane U.

Indeed, we can prove that $U^{\perp}$ is a subspace of $\mathbb{R}^n$ using the subspace test.

(1) $\vec{0} \in U^{\perp}$ because $0 \cdot \vec{n} = \vec{0}$ for all $\vec{n} \in U^{\perp}$.

(2) Suppose $\vec{x}, \vec{y} \in U^{\perp}$. We show that $\vec{x} + \vec{y} \in U^{\perp}$. Since $\vec{x}, \vec{y} \in U^{\perp}$, this means that $\vec{x} = a\vec{n}$ and $\vec{y} = b\vec{n}$ for $a, b \in \mathbb{R}$. So,

$$\begin{aligned} \vec{x} + \vec{y} &= a\vec{n} + b\vec{n} \\ &= (a+b)\vec{n} \ , \end{aligned}$$

where $a + b \in \mathbb{R}$. Hence, $\vec{x} + \vec{y} \in U^{\perp}$, and so $U^{\perp}$ is closed under addition.

(3) Suppose $\vec{x} \in U^{\perp}$ and $k \in \mathbb{R}$. We show that $k\vec{x} \in U^{\perp}$. Since $\vec{x} \in U^{\perp}$, this means that $\vec{x} = a\vec{n}$, where $a \in \mathbb{R}$. So,

$$\begin{aligned} k\vec{x} &= k(a\vec{n}) \\ &= ka\vec{n} \\ &= (ka)\vec{n} \ , \end{aligned}$$

where $ka \in \mathbb{R}$. Hence, $k\vec{x} \in U^{\perp}$, and so $U^{\perp}$ is closed under scalar multiplication.

Thus, because $U^{\perp}$ contains the zero vectors, is closed under addition, and is closed under scalar multiplication, $U^{\perp}$ is a subspace of $\mathbb{R}^n$. $\square$

Lemma

Let $U = \text{span}\{\vec{u_1}, \ldots, \vec{u_k}\} \subseteq \mathbb{R}^n$. Then

$$U^\perp = \{\vec{x} \in \mathbb{R}^n : \vec{x} \cdot \vec{u}_i = 0 \text{ for all } 1 \leq i \leq k\} \ .$$

Proof. Suppose $U = \text{span}\{\vec{u_1}, \ldots, \vec{u_k}\} \subseteq \mathbb{R}^n$. We show that $U^\perp = \{\vec{x} \in \mathbb{R}^n : \vec{x} \cdot \vec{u}_i = 0 \text{ for all } 1 \leq i \leq k\}$.

($\subseteq$) Suppose $\vec{v} \in U^\perp$. Then this means that $\vec{v} \cdot \vec{u} = 0$ for all $\vec{u} \in U$. Then it holds that $\vec{v} \cdot \vec{u}_i = 0$ for all $1 \leq i \leq k$, since each $\vec{u}_i \in U$. Hence, $\vec{v} \in \{\vec{x} \in \mathbb{R}^n : \vec{x} \cdot \vec{u}_i = 0 \text{ for all } 1 \leq i \leq k\}$.

($\supseteq$) Suppose $\vec{v} \in \{\vec{x} \in \mathbb{R}^n : \vec{x} \cdot \vec{u}_i = 0 \text{ for all } 1 \leq i \leq k\}$. Then this means that $\vec{v} \cdot \vec{u}_i = 0$ for all $1 \leq i \leq k$. Let $\vec{u} \in U$. Since $U = \text{span}\{\vec{u}_1, \ldots, \vec{u}_k\} = \{c_1\vec{u}_1 + \ldots + c_k\vec{u}_k : c_1, \ldots, c_k \in \mathbb{R}\}$, it follows that $\vec{u} = c_1\vec{u}_1 + \ldots + c_k\vec{u}_k$ for some $c_1, \ldots, c_k \in \mathbb{R}$. From this we get that

$$\begin{aligned}
\vec{v} \cdot \vec{u} &= \vec{v} \cdot (c_1\vec{u}_1 + \ldots + c_k\vec{u}_k) \\
&= \vec{v} \cdot c_1\vec{u}_1 + \ldots + \vec{v} \cdot c_k\vec{u}_k \\
&= c_1(\vec{v} \cdot \vec{u}_1) + \ldots + c_k(\vec{v} \cdot \vec{u}_k) \\
&= c_1 \cdot 0 + \ldots c_k \cdot 0 \\
&= 0 \ ,
\end{aligned}$$

since $\vec{v} \cdot \vec{u}_i = 0$ for all $1 \leq i \leq k$. Hence, $\vec{v} \in U^\perp$.

So, we have shown that $U^\perp \subseteq \{\vec{x} \in \mathbb{R}^n : \vec{x} \cdot \vec{u}_i = 0 \text{ for all } 1 \leq i \leq k\}$ and $U^\perp \supseteq \{\vec{x} \in \mathbb{R}^n : \vec{x} \cdot \vec{u}_i = 0 \text{ for all } 1 \leq i \leq k\}$. Thus, $U^\perp = \{\vec{x} \in \mathbb{R}^n : \vec{x} \cdot \vec{u}_i = 0 \text{ for all } 1 \leq i \leq k\}$. $\square$

1. If $U = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$, find $U^\perp$.

Let $\vec{u}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ and $\vec{u}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$. We know that

$$U^\perp = \{\vec{x} \in \mathbb{R}^3 : \vec{x} \cdot \vec{u}_1 = 0 \text{ and } \vec{x} \cdot \vec{u}_2 = 0\} \ .$$

Let $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$. Then

$$\vec{x} \cdot \vec{u}_1 = x_1 + 0x_2 + 0x_3 = 0$$

and

$$\vec{x} \cdot \vec{u}_2 = 0x_1 + x_2 + 0x_3 = 0 \ .$$

So, we can solve the following lienar system of equations:

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{array}\right]$$

which gives us $x_1 = 0$, $x_2 = 0$, and $x_3 = t \in \mathbb{R}$. Then we get that

$$\vec{x} = \begin{bmatrix} 0 \\ 0 \\ t \end{bmatrix} = t \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \ .$$

Hence,

$$\begin{aligned} U^\perp &= \{\vec{x} \in \mathbb{R}^3 : \vec{x} \cdot \vec{u}_1 = 0 \text{ and } \vec{x} \cdot \vec{u}_2 = 0\} \\ &= \left\{ t \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} : t \in \mathbb{R} \right\} \\ &= \operatorname{span} \left\{ \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\} \ . \end{aligned}$$

2. If $W = \text{span}\left\{ \begin{bmatrix} -2 \\ 3 \\ 1 \end{bmatrix}, \begin{bmatrix} 5 \\ -1 \\ 2 \end{bmatrix} \right\}$, find $W^\perp$.

Let $\vec{u_1} = \begin{bmatrix} -2 \\ 3 \\ 1 \end{bmatrix}$ and $\vec{u_2} = \begin{bmatrix} 5 \\ -1 \\ 2 \end{bmatrix}$. We know that

$$\begin{aligned} W^\perp &= \{\vec{x} \in \mathbb{R}^3 : \vec{x} \cdot \vec{u}_i = 0 \ \ \textbf{for all } 1 \leq i \leq 3\} \\ &= \{\vec{x} \in \mathbb{R}^3 : \vec{x} \cdot \vec{u}_1 = 0,\ \vec{x} \cdot \vec{u}_2 = 0\} \ . \end{aligned}$$

(Note that $\vec{u}_i \in W$.) Let $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3$. So, solving the system of equations

$$\begin{aligned} -2x_1 + 3x_2 + x_3 &= 0 \\ 5x_1 - x_2 + 2x_3 &= 0 \end{aligned}$$

gets us

$$\left[\begin{array}{rrr|r} -2 & 3 & 1 & 0 \\ 5 & -1 & 2 & 0 \end{array}\right] \to \left[\begin{array}{rrr|r} 1 & 0 & 7/13 & 0 \\ 0 & 1 & 9/13 & 0 \end{array}\right] \ .$$

From this we get that

$$\begin{aligned} x_1 &= -\frac{7}{13}t \\ x_2 &= -\frac{9}{13}t \\ x_3 &= t \end{aligned}$$

for $t \in \mathbb{R}$. Then

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -7/13t \\ -9/13t \\ t \end{bmatrix} = t \begin{bmatrix} -7/13 \\ -9/13 \\ 1 \end{bmatrix} \ .$$

Hence,

$$\begin{aligned} W^\perp &= \left\{ \begin{bmatrix} -7/13t \\ -9/13t \\ t \end{bmatrix} : t \in \mathbb{R} \right\} \\ &= \text{span}\left\{ \begin{bmatrix} -7 \\ -9 \\ 13 \end{bmatrix} \right\} \ , \end{aligned}$$

if we take $t = 13$.

Orthogonal Projections

Definition (Projection)

The projection of $\vec{v}$ onto $\vec{w}$ is

$$\text{proj}_{\vec{w}}\vec{v} = \frac{\vec{v} \cdot \vec{w}}{||\vec{w}||^2}\ \vec{w}\ .$$

Definition (Orthogonal Projection)

Let $U \subseteq \mathbb{R}^n$ be a subspace with orthogonal basis $\{\vec{f}_1, \ldots, \vec{f}_k\}$. For $\vec{x} \in \mathbb{R}^n$, we define the orthogonal projection of $\vec{x}$ onto U by

$$\text{proj}_U\vec{x} = \left(\frac{\vec{x} \cdot \vec{f}_1}{||\vec{f}_1||^2}\right)\vec{f}_1 + \ldots + \left(\frac{\vec{x} \cdot \vec{f}_k}{||\vec{f}_k||^2}\right)\vec{f}_k\ .$$

Observations:

(1) For all $\vec{x} \in \mathbb{R}^n$, $(\vec{x} - \text{proj}_U\vec{x})$ is orthogonal to every $\vec{u} \in U$. That is, $(\vec{x} - \text{proj}_U\vec{x}) \in U^\perp$.

Proof. □

(2) If $\vec{x} \in U$, then $\text{proj}_U\vec{x}$ is the Fourier Expansion of $\vec{x}$.

(3) $\text{proj}_U : \mathbb{R}^n \to U$ is a <u>linear transformation</u> with $\ker(\text{proj}_U) = U^\perp$ and $\text{Im}(\text{proj}_U) = U$.

1. Let $\mathrm{proj}_U\big|_U : U \to U$ be the restriction of proj_U to U.

   (a) Prove that $\mathrm{proj}_U\big|_U$ is the identity map on U.

   (b) $(\mathrm{proj}_U\big|_U) \circ \mathrm{proj}_U = \mathrm{proj}_U$.

2. Consider $U = \text{span}\left\{ \begin{bmatrix} 4/5 \\ 0 \\ -3/5 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$. Express $\vec{x} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ in the form $\vec{x} = \vec{u} + \vec{w}$ with $\vec{u} \in U$ and $\vec{w} \in U^{\perp}$.

We know that $\vec{x} = \vec{u} + \vec{w}$. Then

$$\vec{x} = \vec{u} + \vec{w} = \text{proj}_U \vec{x} + (\vec{x} - \text{proj}_U \vec{x}) \ ,$$

where $\text{proj}_U \vec{x} \in U$ and $(\vec{x} - \text{proj}_U \vec{x}) \in U^{\perp}$. To compute $\text{proj}_U \vec{x}$, we need an orthogonal basis for U. Is $\left\{ \begin{bmatrix} 4/5 \\ 0 \\ -3/5 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$ orthogonal and a basis? This set is a basis as it is linearly independent. It is also orthogonal because

$$\begin{aligned} \begin{bmatrix} 4/5 \\ 0 \\ -3/5 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} &= \left(\frac{4}{5} \cdot 0 \right) + (0 \cdot 1) + \left(-\frac{3}{5} \cdot 0 \right) \\ &= 0 + 0 + 0 \\ &= 0 \ . \end{aligned}$$

Then

$$\begin{aligned} \vec{u} &:= \text{proj}_U \vec{x} \\ &= \frac{\vec{x} \cdot \vec{v}_1}{||\vec{v}_1||^2} \vec{v}_1 + \frac{\vec{x} \cdot \vec{v}_2}{||\vec{v}_2||^2} \vec{v}_2 \\ &= \begin{bmatrix} -4/5 \\ 2 \\ 3/5 \end{bmatrix} \\ &\in U \end{aligned}$$

and

$$\begin{aligned} \vec{w} &:= \vec{x} - \vec{u} \\ &= \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} - \begin{bmatrix} -4/5 \\ 2 \\ 3/5 \end{bmatrix} \\ &= \begin{bmatrix} 9/5 \\ 0 \\ 12/5 \end{bmatrix} \\ &\in U^{\perp} \ . \end{aligned}$$

So, $\vec{x} = \vec{u} + \vec{w}$, where $\vec{u} \in U$ and $\vec{w} \in U^{\perp}$.

Lemma

The orthogonal projection of $\vec{x}$ onto U is the vector in U that is closest to $\vec{x}$. That is,

$$||\vec{x} - \mathrm{proj}_U \vec{x}|| < ||\vec{x} - \vec{y}||$$

for all $\vec{y} \in U$, $\vec{y} \neq \mathrm{proj}_U \vec{x}$.

1. Find the point on the plane $3x + y - 2z = 0$ that is closest to the point $(1, 1, 1)$. (Note that this point isn't on the plane. Try plugging in these values into the equation of plane and you will find that the equation will not hold.)

   First, solve $3x + y - 2z = 0$. We can set it up as an augmented matrix.

$$\begin{aligned} &\left[\begin{array}{ccc|c} 3 & 1 & -2 & 0 \end{array}\right] \\ \xrightarrow{\frac{1}{3}R1} &\left[\begin{array}{ccc|c} 1 & 1/3 & -2/3 & 0 \end{array}\right] \end{aligned}$$

   From this we get that

$$\begin{aligned} x &= -\frac{1}{3}s + \frac{2}{3}t \\ y &= s \\ z &= t \end{aligned}$$

   for $s, t \in \mathbb{R}$. So,

$$\begin{aligned} \begin{bmatrix} x \\ y \\ z \end{bmatrix} &= \begin{bmatrix} -\frac{1}{3}s + \frac{2}{3}t \\ s \\ t \end{bmatrix} \\ &= \begin{bmatrix} -\frac{1}{3}s \\ s \\ 0 \end{bmatrix} + \begin{bmatrix} \frac{2}{3}t \\ 0 \\ t \end{bmatrix} \\ &= s \begin{bmatrix} -\frac{1}{3} \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} \frac{2}{3} \\ 0 \\ 1 \end{bmatrix} . \end{aligned}$$

   What does this mean? It means that the plane, which we'll denote as U, can be expressed as

$$\begin{aligned} U &= \left\{ s \begin{bmatrix} -\frac{1}{3} \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} \frac{2}{3} \\ 0 \\ 1 \end{bmatrix} : s, t \in \mathbb{R} \right\} \\ &= \operatorname{span} \left\{ \begin{bmatrix} -\frac{1}{3} \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} \frac{2}{3} \\ 0 \\ 1 \end{bmatrix} \right\} \\ &= \operatorname{span} \left\{ \begin{bmatrix} -1 \\ 3 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 3 \end{bmatrix} \right\} . \end{aligned}$$

Now, this spanning set is not orthogonal because

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

However, this set is indeed a basis, which means we can use the Gram-Schmidt process to construct an orthogonal basis from this basis. So, let $\vec{u}_1 = \begin{bmatrix}-1\\3\\0\end{bmatrix}$ and $\vec{u}_2 = \begin{bmatrix}2\\0\\3\end{bmatrix}$. We will construct an orthogonal basis $\{\vec{f}_1, \vec{f}_2\}$ from the basis $\{\vec{u}_1, \vec{u}_2\}$.

$$\vec{f}_1 = \vec{u}_1 = \begin{bmatrix}-1\\3\\0\end{bmatrix} \ ,$$

$$\begin{aligned}
\vec{f}_2 &= \vec{u}_2 - \left(\frac{\vec{u}_2 \cdot \vec{f}_1}{||\vec{f}_1||^2}\right) \vec{f}_1 \\
&= \begin{bmatrix}2\\0\\3\end{bmatrix} - \frac{-2}{10} \begin{bmatrix}-1\\3\\0\end{bmatrix} \\
&= \begin{bmatrix}2\\0\\3\end{bmatrix} - \frac{-1}{5} \begin{bmatrix}-1\\3\\0\end{bmatrix} \\
&= \begin{bmatrix}10/5\\0\\3\end{bmatrix} - \begin{bmatrix}1/5\\-3/5\\0\end{bmatrix} \\
&= \begin{bmatrix}9/5\\3/5\\3\end{bmatrix} \ .
\end{aligned}$$

So, we get an orthogonal basis

$$\{\vec{f}_1, \vec{f}_2\} = \left\{ \begin{bmatrix}-1\\3\\0\end{bmatrix}, \begin{bmatrix}9/5\\3/5\\3\end{bmatrix} \right\}$$

Hence, the closest point in U to $\vec{x} = \begin{bmatrix}1\\1\\1\end{bmatrix}$ is

$$\text{proj}_U \vec{x} = \left(\frac{\vec{x} \cdot \vec{f}_1}{||\vec{f}_1||^2}\right) \vec{f}_1 + \left(\frac{\vec{x} \cdot \vec{f}_2}{||\vec{f}_2||^2}\right) \vec{f}_2$$

$$= \frac{1}{5}\begin{bmatrix}-1\\3\\0\end{bmatrix} + \frac{27/5}{}$$

Orthogonal Diagonalization

Recall one of the earlier Theorems:

Theorem 15

An $n \times n$ matrix is diagonalizable if and only if A has n linearly independent eigenvectors.

Theorem 28

Let $A \in M_{n\times n}$ be a square matrix. If A is symmetric, then all eigenvalues of A are real numbers.

Proof. $\square$

Definition (Orthogonal Matrix)

A square real matrix B is called orthogonal if $B^{-1} = B^{\mathsf{T}}$

1. Is $A = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$ orthogonal?

   Well, let's check if $A^{-1} = A^{\mathsf{T}}$. So,

$$\begin{aligned}
A^{-1} &= \left(\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \right)^{-1} \\
&= \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}^{-1} \\
&= \frac{1}{1} \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \\
&= \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \\
&= \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}^{\mathsf{T}} \\
&= \left(\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \right)^{\mathsf{T}} \\
&= A^{\mathsf{T}} .
\end{aligned}$$

   Indeed, we can also check that $AA^{-1} = AA^{\mathsf{T}} = I_2 = A^{\mathsf{T}}A = A^{-1}A$.

Theorem 29

Let A be an $n \times n$ matrix. The following are equivalent:

(1) A is orthogonal.

(2) The rows of A are orthonormal.

(3) The columns of A are orthonormal.

Proof. $\square$

Definition (Orthogonally Diagonalizable)

A matrix A is called orthogonally diagonalizable if there exists an orthogonal matrix P such that $D = P^{-1}AP = P^{\mathsf{T}}AP$ is a diagonal matrix.

Theorem 30 (Principal Axis Theorem)

Let A be an $n \times n$ matrix. The following are equivalent:

(1) A is orthogonally diagonalizable.

(2) A is symmetric.

(3) A has an orthonormal set of n eigenvectors.

Proof. (Theorem 8.2.2 in Textbook) $\square$

Question: How do you orthogonally diagonalize a symmetric matrix?

Answer: First, find the eigenvalues of the matrix. Then find the corresponding eigenvectors of each eigenvalue. Then apply Gram-Schmidt (if necessary) to get an orthogonal basis for each eigenspace E_λ. Finally, normalize each vector in the orthogonal basis to get an orthonormal basis for the eigenspace E_λ.

Question: What about eigenvectors corresponding to different eigenvalues?

If A is a symmetric matrix, then the eigenvectors of A corresponding to distinct eigenvalues are orthogonal.

Proof. $\square$

Procedure for Orth. Diagonalizing a Symmetric Matrix

(1) Find the eigenvalues of A.

(2) For each eigenvalue λ, find an orthonormal basis for the eigenspace of that eigenvalue, $E_\lambda(A) = \{\vec{v} \in \mathbb{R}^n : A\vec{v} = \lambda\vec{v}\}$ (may require Gram-Schmidt).

(3) Diagonalize A using the orthonormal basis for $\mathbb{R}^n$ consisting of the eigenvectors of A.

1. Orthogonally diagonalize $A = \begin{bmatrix} 1 & -2 & -2 \\ -2 & 1 & -2 \\ -2 & -2 & 1 \end{bmatrix}$.

(1) First, we find the eigenvalues via the characteristic polynomial $C_A(\lambda) = \det(\lambda I_3 - A) = 0$. We have that

$$\begin{aligned}
\det(\lambda I_3 - A) &= \det\left(\begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix} - \begin{bmatrix} 1 & -2 & -2 \\ -2 & 1 & -2 \\ -2 & -2 & 1 \end{bmatrix}\right) \\
&= \det\left(\begin{bmatrix} \lambda-1 & 2 & 2 \\ 2 & \lambda-1 & 2 \\ 2 & 2 & \lambda-1 \end{bmatrix}\right) \\
&= \det\left(\begin{bmatrix} \lambda-1 & 2 & 2 \\ 2 & \lambda-1 & 2 \\ 0 & -\lambda+3 & \lambda-3 \end{bmatrix}\right) \\
&= (\lambda-1)[(\lambda-1)(\lambda-3) - (2)(-\lambda+3)] - (2)[(2)(\lambda-3) - (2)(-\lambda+3)] \\
&= (\lambda-1)[\lambda^2 - 4\lambda + 3 - (-2\lambda+6)] - (2)[2\lambda - 6 - (-2\lambda+6)] \\
&= (\lambda-1)(\lambda^2 - 2\lambda - 3) - (2)(4\lambda - 12) \\
&= \lambda^3 - 2\lambda^2 - 3\lambda - \lambda^2 + 2\lambda + 3 - (8\lambda - 24) \\
&= \lambda^3 - 3\lambda^2 - 9\lambda + 27 \\
&= \lambda^2(\lambda-3) - 9(\lambda-3) \\
&= (\lambda^2 - 9)(\lambda-3) \\
&= (\lambda+3)(\lambda-3)(\lambda-3)\ .
\end{aligned}$$

So, $\lambda_1 = \lambda_2 = 3$ with multiplicity 2, and $\lambda_3 = -3$ with multiplicity 1, are the eigenvalues of A. Now, we find the eigenvectors associated with each eigenvalue.

(Continued on next page.)

For $\lambda_1 = \lambda_2 = 3$, solve the system $\left[\lambda_1 I_3 - A \mid 0\right]$. So,

$$\left[\begin{array}{ccc|c} 2 & 2 & 2 & 0 \\ 2 & 2 & 2 & 0 \\ 2 & 2 & 2 & 0 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

Here we get that

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -s-t \\ s \\ t \end{bmatrix} = \begin{bmatrix} -s \\ s \\ 0 \end{bmatrix} + \begin{bmatrix} -t \\ 0 \\ t \end{bmatrix} = s\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + t\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}.$$

So, we get that the eigenspace of $\lambda_1 = \lambda_2 = 3$ is

$$\begin{aligned} E_3(A) &= \{\vec{v} \in \mathbb{R}^3 : A\vec{v} = 3\vec{v}\} \\ &= \left\{ s\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + t\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} : s, t \in \mathbb{R} \right\} \\ &= \operatorname{span}\left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\} \end{aligned}$$

which has a basis $\left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\}$. So, $\dim(E_3(A)) = 2$.

For $\lambda_3 = -3$, solve the system $\left[\lambda_3 I_3 - A \mid 0\right]$. So,

$$\left[\begin{array}{ccc|c} -4 & 2 & 2 & 0 \\ 2 & -4 & 2 & 0 \\ 2 & 2 & -4 & 0 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

Here we get that

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} t \\ t \\ t \end{bmatrix} = t\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

So, we get that the eigenspace of $\lambda_3 = -3$ is

$$\begin{aligned} E_{-3}(A) &= \left\{ t\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} : t \in \mathbb{R} \right\} \\ &= \operatorname{span}\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}. \end{aligned}$$

(2) Now, we check if the spanning sets for $E_3(A)$ and $E_{-3}(A)$ are orthogonal. Our basis for $E_3(A)$ which is $\left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\}$ is not orthogonal. So, we use the Gram-Schmidt process to construct an orthogonal basis $\{\vec{f_1}, \vec{f_2}\}$. After going through the process, we get that

$$\{\vec{f_1}, \vec{f_2}\} = \left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1/2 \\ -1/2 \\ 1 \end{bmatrix} \right\} .$$

From this we can get an orthonormal basis

$$\left\{ \frac{1}{||\vec{f_1}||}\vec{f_1}, \frac{1}{||\vec{f_2}||}\vec{f_2} \right\} = \left\{ \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix}, \begin{bmatrix} -1/\sqrt{6} \\ -1/\sqrt{6} \\ \sqrt{2}/\sqrt{3} \end{bmatrix} \right\}$$

Next, our basis for $E_{-3}(A)$ which is $\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$ is trivially orthogonal. So, an orthonormal basis for $E_{-3}(A)$ is

$$\left\{ \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\} = \left\{ \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix} \right\}$$

(3) Hence, our orthonormal matrix is

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

2. Orthogonally diagonalize $A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$.

(1) Find the eigenvalues via the characteristic polynomial $C_A(\lambda) = \det(\lambda I_3 - A) = 0$. So,

$$\begin{aligned}
\det(\lambda I_3 - A) &= \det\left(\begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix} - \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}\right) \\
&= \det\left(\begin{bmatrix} \lambda - 1 & -1 & -1 \\ -1 & \lambda - 1 & -1 \\ -1 & -1 & \lambda - 1 \end{bmatrix}\right) \\
&= \det\left(\begin{bmatrix} \lambda - 1 & -1 & -1 \\ -1 & \lambda - 1 & -1 \\ 0 & -\lambda & \lambda \end{bmatrix}\right) \\
&= (\lambda - 1)[(\lambda - 1)(\lambda) - (-1)(-\lambda)] - (-1)(-\lambda - (\lambda)) \\
&= (\lambda - 1)(\lambda^2 - \lambda - (\lambda)) - (-1)(-2\lambda) \\
&= (\lambda - 1)(\lambda^2 - 2\lambda) - (2\lambda) \\
&= \lambda^3 - 2\lambda^2 - \lambda^2 + 2\lambda - (2\lambda) \\
&= \lambda^3 - 3\lambda^2 \\
&= \lambda^2(\lambda - 3) \ .
\end{aligned}$$

So, $\lambda_1 = \lambda_2 = 0$ with multiplicity 2, and $\lambda_3 = 3$ with multiplicity 1, are the eigenvalues of A. Now, we find the eigenvectors of each eigenvalue to determine the eigenspaces of each eigenvalue.

For $\lambda_1 = \lambda_2 = 0$, solve the system $\left[\lambda_1 I_3 - A \mid 0\right]$. So,

$$\left[\begin{array}{ccc|c} -1 & -1 & -1 & 0 \\ -1 & -1 & -1 & 0 \\ -1 & -1 & -1 & 0 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \ .$$

So, we get the solution

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -s - t \\ s \\ t \end{bmatrix} = \begin{bmatrix} -s \\ s \\ 0 \end{bmatrix} + \begin{bmatrix} -t \\ 0 \\ t \end{bmatrix} = s\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + t\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$