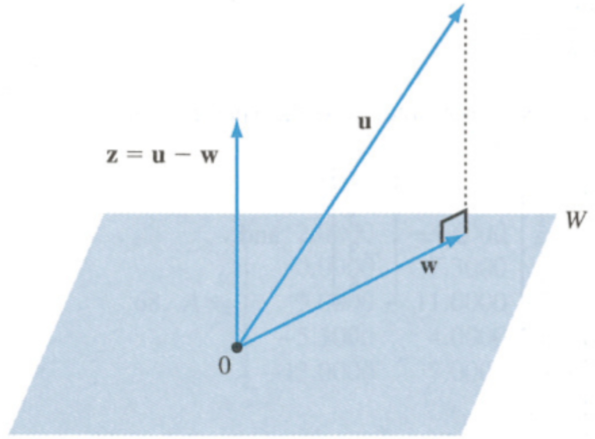
Section 6.3 Orthogonal Projections

Orthogonal Projection:

The unique vector **w** in subspace W that is "closest" to vector **u**.



We have $\mathbf{u} = \mathbf{w} + \mathbf{z}$, where **z** is always orthogonal to all vectors in W. It is therefore worthy of studying "orthogonal complement" of W, the set of vectors that possess this property.

Definition

The **orthogonal complement** of a nonempty subset $\mathcal{S}$ of $\mathcal{R}^n$, denoted by $\mathcal{S}^\perp$ (read "$\mathcal{S}$ perp), is the set of all vectors in $\mathcal{R}^n$ that are **orthogonal** to every vector in $\mathcal{S}$. That is,

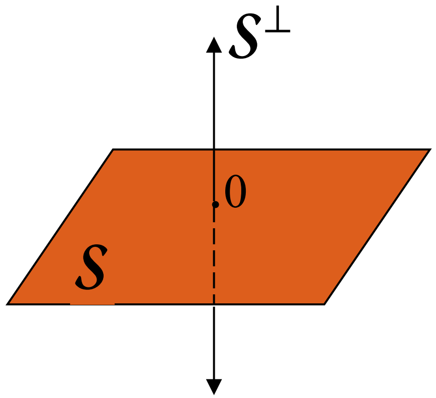
$$\mathcal{S}^\perp = \{\mathbf{v} \in \mathcal{R}^n : \mathbf{v} \cdot \mathbf{u} = 0, \forall \mathbf{u} \in \mathcal{S}\}.$$

Definition

The **orthogonal complement** of a nonempty subset $\mathcal{S}$ of $\mathcal{R}^n$, denoted by $\mathcal{S}^\perp$ (read "$\mathcal{S}$ perp"), is the set of all vectors in $\mathcal{R}^n$ that are **orthogonal** to every vector in $\mathcal{S}$. That is,

$$\mathcal{S}^\perp = \{\mathbf{v} \in \mathcal{R}^n : \mathbf{v} \cdot \mathbf{u} = 0, \forall \mathbf{u} \in \mathcal{S}\}.$$

Example:



Example: $\mathcal{S} = \mathcal{R}^n \Rightarrow \mathcal{S}^\perp = \{0\}$,

$\mathcal{S} = \{0\} \Rightarrow \mathcal{S}^\perp = \mathcal{R}^n$.

Definition

The **orthogonal complement** of a nonempty subset $\mathcal{S}$ of $\mathcal{R}^n$, denoted by $\mathcal{S}^\perp$ (read "$\mathcal{S}$ perp"), is the set of all vectors in $\mathcal{R}^n$ that are **orthogonal** to every vector in $\mathcal{S}$. That is,

$$\mathcal{S}^\perp = \{\mathbf{v} \in \mathcal{R}^n : \mathbf{v} \cdot \mathbf{u} = 0, \forall \mathbf{u} \in \mathcal{S}\}.$$

Example: $W = \{ [\, w_1 \;\; w_2 \;\; 0 \,]^T \mid w_1, \;\; w_2 \in \mathcal{R}.\}$
$V = \{ [\, 0 \;\; 0 \;\; v_3 \,]^T \mid v_3 \in \mathcal{R}.\}$

Then, $V = W^\perp$.

Proof:

(1) for all $\mathbf{v} \in V$ and $\mathbf{w} \in W$, $\mathbf{v} \cdot \mathbf{w} = 0 \Rightarrow V \subseteq W^\perp$;

(2) since $\mathbf{e}_1$, $\mathbf{e}_2 \in W$, all $\mathbf{z} = [\, z_1 \;\; z_2 \;\; z_3 \,]^T \in W^\perp$ must have $z_1 = z_2 = 0$ by $\mathbf{z} \cdot \mathbf{e}_1 = \mathbf{z} \cdot \mathbf{e}_2 = 0 \Rightarrow W^\perp \subseteq V$.

Property

1. $\mathbf{0} \in \mathcal{S}^{\perp}$ for every nonempty subset $\mathcal{S}$ of $\mathcal{R}^n$.
2. If $\mathbf{v}, \mathbf{w} \in \mathcal{S}^{\perp}$ for a nonempty subset $\mathcal{S}$ of $\mathcal{R}^n$, then $\mathbf{v} + \mathbf{w} \in \mathcal{S}^{\perp}$ and $c\mathbf{v} \in \mathcal{S}^{\perp}$ for any scalar $c \in \mathcal{R}$.

Proof You show it.

Property

The orthogonal complement of any nonempty subset of $\mathcal{R}^n$ is a subspace of $\mathcal{R}^n$.

Proposition

For any nonempty subset $\mathcal{S}$ of $\mathcal{R}^n$, $(\text{Span } \mathcal{S})^{\perp} = \mathcal{S}^{\perp}$.

Proposition

Let W be a subspace of $\mathcal{R}^n$, and $\mathcal{B}$ be a basis of W. Then $\mathcal{B}^{\perp} = W^{\perp}$.

Example: For $W = \text{Span}\{\mathbf{u}_1, \mathbf{u}_2\}$, where $\mathbf{u}_1 = [\ 1\ \ 1\ \ \text{-}1\ \ 4\ ]^T$ and $\mathbf{u}_2 = [\ 1\ \text{-}1\ \ 1\ \ 2\ ]^T$, $\mathbf{v} \in W^{\perp}$ if and only if $\mathbf{u}_1 \cdot \mathbf{v} = \mathbf{u}_2 \cdot \mathbf{v} = \mathbf{0}$

i.e., $\mathbf{v} = [\ x_1\ \ x_2\ \ x_3\ \ x_4\ ]^T$ satisfies

$$\begin{aligned} x_1 + x_2 - x_3 + 4x_4 &= 0 \\ x_1 - x_2 + x_3 + 2x_4 &= 0. \end{aligned} \Leftrightarrow \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -3x_4 \\ x_3 - x_4 \\ x_3 \\ x_4 \end{bmatrix} = x_3 \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -3 \\ -1 \\ 0 \\ 1 \end{bmatrix}$$

$$\Leftrightarrow \mathcal{B} = \left\{ \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ -1 \\ 0 \\ 1 \end{bmatrix} \right\} \text{is a basis for } W^{\perp}.$$

Let $A = \begin{bmatrix} 1 & 1 & -1 & 4 \\ 1 & -1 & 1 & 2 \end{bmatrix}$. Then the above system of linear equation is equivalent to

$$A\mathbf{x} = \mathbf{0}.$$

Question 1: Which is correct? 1) $W = \text{Col } A$? 2) $W = \text{Row } A$?
Question 2: Which is correct? 1) $W^{\perp} = \text{Null } A$? 2) $W^{\perp} = \text{Null } A^T$?

Property

For any real matrix A, $(\text{Row } A)^{\perp} = \text{Null } A$, or $(\text{Col } A)^{\perp} = \text{Null } A^T$.

Proof $\mathbf{v} \in (\text{Row } A)^{\perp} \Leftrightarrow \mathbf{w} \cdot \mathbf{v} = 0$ for all $\mathbf{w} \in \text{Span}\{\text{rows of } A\}$

$\Leftrightarrow A\mathbf{v} = \mathbf{0}$.

Also, $(\text{Col } A)^{\perp} = (\text{Row } A^T) = \text{Null } A^T$.

Theorem 6.7 (Orthogonal Decomposition Theorem)

Let $W \subset \mathcal{R}^n$ be a subspace of $\mathcal{R}^n$. Then, for any vector $\mathbf{u}$ in $\mathcal{R}^n$, there exist unique vectors $\mathbf{w} \in W$ and $\mathbf{z} \in W^\perp$ such that $\mathbf{u} = \mathbf{w} + \mathbf{z}$.

In addition, if $\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k\}$ is an orthonormal basis for W, then

$$\mathbf{w} = (\mathbf{u} \cdot \mathbf{v}_1)\mathbf{v}_1 + (\mathbf{u} \cdot \mathbf{v}_2)\mathbf{v}_2 + \cdots + (\mathbf{u} \cdot \mathbf{v}_k)\mathbf{v}_k.$$

Proof

Theorem 6.7 (Orthogonal Decomposition Theorem)

Let $W \subset \mathcal{R}^n$ be a subspace of $\mathcal{R}^n$. Then, for any vector $\mathbf{u}$ in $\mathcal{R}^n$, there exist unique vectors $\mathbf{w} \in W$ and $\mathbf{z} \in W^\perp$ such that $\mathbf{u} = \mathbf{w} + \mathbf{z}$.

In addition, if $\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k\}$ is an orthonormal basis for W, then

$$\mathbf{w} = (\mathbf{u} \cdot \mathbf{v}_1)\mathbf{v}_1 + (\mathbf{u} \cdot \mathbf{v}_2)\mathbf{v}_2 + \cdots + (\mathbf{u} \cdot \mathbf{v}_k)\mathbf{v}_k.$$

Proof For all $\mathbf{u} \in \mathcal{R}^n$ and $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$, let $\mathbf{z} = \mathbf{u} - \mathbf{w}$.

$\Rightarrow \mathbf{w} \in W$ and $\mathbf{u} = \mathbf{w} + \mathbf{z}$.

Now, $\mathbf{z} \in W^\perp$, since by the orthogonality of $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$,

$$\mathbf{z} \cdot \mathbf{v}_i = (\mathbf{u} - \mathbf{w}) \cdot \mathbf{v}_i = \mathbf{u} \cdot \mathbf{v}_i - \mathbf{w} \cdot \mathbf{v}_i = \mathbf{u} \cdot \mathbf{v}_i - \mathbf{u} \cdot \mathbf{v}_i = 0.$$

Suppose there is another decomposition $\mathbf{u} = \mathbf{w}' + \mathbf{z}'$, where $\mathbf{w}' \in W$ and $\mathbf{z}' \in W^\perp$. Then $\mathbf{w} + \mathbf{z} = \mathbf{w}' + \mathbf{z}' \Rightarrow \mathbf{w} - \mathbf{w}' = \mathbf{z}' - \mathbf{z}$

$\Rightarrow \mathbf{w} - \mathbf{w}' \in W \cap W^\perp \Rightarrow (\mathbf{w} - \mathbf{w}') \cdot (\mathbf{w} - \mathbf{w}') = 0 \Rightarrow \mathbf{w} - \mathbf{w}' = \mathbf{0}$

$\Rightarrow \mathbf{w} = \mathbf{w}'$ and $\mathbf{z}' = \mathbf{z}$.

Example: Let W be the solution space of $x_1 - x_2 + 2x_3 = 0$.
Then $\mathbf{u} = [\ 1\ \ 3\ \ 4\ ]^T \in \mathcal{R}^3$ can be uniquely decomposed as
$\mathbf{u} = \mathbf{w} + \mathbf{z}$ where $\mathbf{w} \in W$ and $\mathbf{v} \in W^{\perp}$. How do you find $\mathbf{w}$ and $\mathbf{z}$?

First of all, W has an orthonormal basis

$$\mathcal{B} = \{\mathbf{w}_1, \mathbf{w}_2\} = \left\{ \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \frac{1}{\sqrt{3}} \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} \right\}$$

Then,

$$\mathbf{w} = (\mathbf{u} \cdot \mathbf{w}_1)\mathbf{w}_1 + (\mathbf{u} \cdot \mathbf{w}_2)\mathbf{w}_2 = \frac{4}{\sqrt{2}}\mathbf{w}_1 + \frac{6}{\sqrt{3}}\mathbf{w}_2$$

$$= 2 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + 2 \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 4 \\ 2 \end{bmatrix} \in W, \text{ and}$$

$$\mathbf{z} = \mathbf{u} - \mathbf{w} = \begin{bmatrix} 1 \\ 3 \\ 4 \end{bmatrix} - \begin{bmatrix} 0 \\ 4 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} \in W^{\perp}.$$

Property

For any subspace W of $\mathcal{R}^n$,

$$\dim W + \dim W^{\perp} = n.$$

Proof You show that (a basis of W)$\cup$(a basis of $W^{\perp}$) = a basis of $\mathcal{R}^n$.

Homework Set for Section 6.3

- Section 6.3: Problems 2, 4, 5, 9, 11, 14, 21, 22, 24, 58, 62, 65-67, 71, 72, 75, 79, 82.

Definition

The **orthogonal complement** of a nonempty subset $\mathcal{S}$ of $\mathcal{R}^n$, denoted by $\mathcal{S}^\perp$ (read "$\mathcal{S}$ perp"), is the set of all vectors in $\mathcal{R}^n$ that are **orthogonal** to every vector in $\mathcal{S}$. That is,

$$\mathcal{S}^\perp = \{\mathbf{v} \in \mathcal{R}^n : \mathbf{v} \cdot \mathbf{u} = 0, \forall \mathbf{u} \in \mathcal{S}\}.$$

Theorem 6.7 (Orthogonal Decomposition Theorem)

Let $W \subset \mathcal{R}^n$ be a subspace of $\mathcal{R}^n$. Then, for any vector $\mathbf{u}$ in $\mathcal{R}^n$, there exist unique vectors $\mathbf{w} \in W$ and $\mathbf{z} \in W^\perp$ such that $\mathbf{u} = \mathbf{w} + \mathbf{z}$.

In addition, if $\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k\}$ is an orthonormal basis for W, then

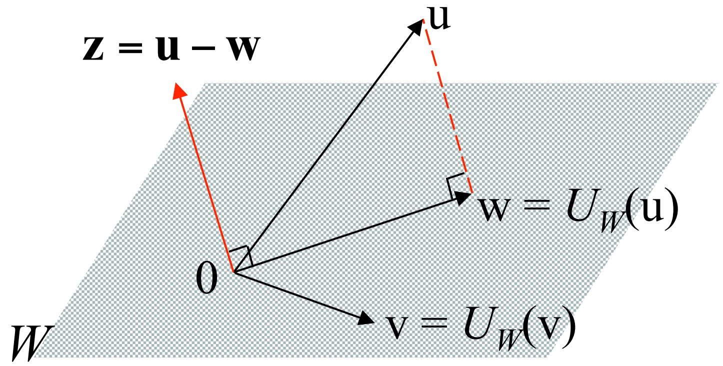
$$\mathbf{w} = (\mathbf{u} \cdot \mathbf{v}_1)\mathbf{v}_1 + (\mathbf{u} \cdot \mathbf{v}_2)\mathbf{v}_2 + \cdots + (\mathbf{u} \cdot \mathbf{v}_k)\mathbf{v}_k.$$

Definitions

Let W be a subspace of $\mathcal{R}^n$ and $\mathbf{u} \in \mathcal{R}^n$. The **orthogonal projection of u on W** is the unique vector $\mathbf{w}$ such that $\mathbf{u} - \mathbf{w} \in W^{\perp}$.

The function $U_W : \mathcal{R}^n \to \mathcal{R}^n$ such that $U_W(\mathbf{u})$ is the orthogonal projection of $\mathbf{u}$ on W for every $\mathbf{u} \in \mathcal{R}^n$ is called the **orthogonal projection operator** on W.

For $n = 3$ and W a 2-dimensional subspace of $\mathcal{R}^3$, the above definitions match the orthogonal projection defined previously.



Proposition: U_W is linear.

Proof If $U_W(\mathbf{u}_1) = \mathbf{w}_1$ and $U_W(\mathbf{u}_2) = \mathbf{w}_2$, then $\exists$ unique $\mathbf{z}_1, \mathbf{z}_2 \in W^{\perp}$ such that $\mathbf{u}_1 = \mathbf{w}_1 + \mathbf{z}_1$ and $\mathbf{u}_2 = \mathbf{w}_2 + \mathbf{z}_2 \Rightarrow \mathbf{u}_1 + \mathbf{u}_2 = \mathbf{w}_1 + \mathbf{w}_2 + \mathbf{z}_1 + \mathbf{z}_2$
$\Rightarrow U_W(\mathbf{u}_1 + \mathbf{u}_2) = \mathbf{w}_1 + \mathbf{w}_2$ since $\mathbf{w}_1 + \mathbf{w}_2 \in W$ and $\mathbf{z}_1 + \mathbf{z}_2 \in W^{\perp}$
Similarly $U_W(c\mathbf{u}) = c\mathbf{w}\ \forall c, \mathbf{u}$

Example: Let W be the solution space of $x_1 - x_2 + 2x_3 = 0$, and have an orthonormal basis

$$\mathcal{B} = \{\mathbf{w}_1, \mathbf{w}_2\} = \left\{ \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \frac{1}{\sqrt{3}} \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} \right\}$$

$\mathbf{u} = [\ 1\ \ 3\ \ 4\ ]^T \in \mathcal{R}^3$ can be uniquely decomposed as $\mathbf{u} = \mathbf{w} + \mathbf{z}$, where

$$\mathbf{w} = (\mathbf{u} \cdot \mathbf{w}_1)\mathbf{w}_1 + (\mathbf{u} \cdot \mathbf{w}_2)\mathbf{w}_2 = \frac{4}{\sqrt{2}}\mathbf{w}_1 + \frac{6}{\sqrt{3}}\mathbf{w}_2$$

$$= 2 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + 2 \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 4 \\ 2 \end{bmatrix} \in W, \text{ and}$$

$$\mathbf{z} = \mathbf{u} - \mathbf{w} = \begin{bmatrix} 1 \\ 3 \\ 4 \end{bmatrix} - \begin{bmatrix} 0 \\ 4 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} \in W^{\perp}.$$

Definition

The **standard matrix** of **orthogonal projection operator** U_W on a subspace W of $\mathcal{R}^n$ is called the **orthogonal projection matrix** for W and is denoted as P_W.

Question

Given a subspace W of $\mathcal{R}^n$, how do we find its orthogonal projection matrix P_W?

Lemma

Let C be a matrix with **linearly independent** columns. Then C^TC is **invertible**.

Proof Suppose $C^TC\mathbf{b} = \mathbf{0}$ for some $\mathbf{b}$.

$$\Rightarrow \ \|C\mathbf{b}\|^2 = (C\mathbf{b})\cdot(C\mathbf{b}) = (C\mathbf{b})^T C\mathbf{b} = \mathbf{b}^T C^T C\mathbf{b} = \mathbf{b}^T\mathbf{0} = 0.$$

$\Rightarrow \ C\mathbf{b} = \mathbf{0} \ \Rightarrow \ \mathbf{b} = \mathbf{0}$ since C has L.I. columns.

Thus C^TC is invertible.

Theorem 6.8
Let C be an $n \times k$ matrix whose columns form a basis for a subspace W of $\mathcal{R}^n$. Then $P_W = C(C^T C)^{-1} C^T$.

Proof Let $\mathbf{u} \in \mathcal{R}^n$ and $\mathbf{w} = U_W(\mathbf{u})$. Since $W = \text{Col } C$, $\mathbf{w} = C\mathbf{v}$ for some $\mathbf{v} \in \mathcal{R}^k$ and $\mathbf{u} - \mathbf{w} \in W^\perp = (\text{Col } C)^\perp = (\text{Row } C^T)^\perp = \text{Null } C^T$.

$\Rightarrow \mathbf{0} = C^T(\mathbf{u} - \mathbf{w}) = C^T\mathbf{u} - C^T\mathbf{w} = C^T\mathbf{u} - C^T C\mathbf{v}$.

$\Rightarrow C^T\mathbf{u} = C^T C\mathbf{v}$.

$\Rightarrow \mathbf{v} = (C^T C)^{-1} C^T\mathbf{u}$ and $\mathbf{w} = C(C^T C)^{-1} C^T\mathbf{u}$ as $C^T C$ is invertible.

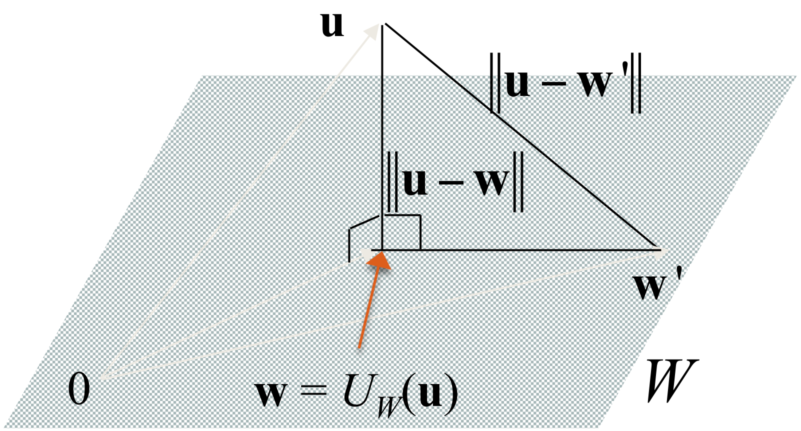
Example: Let W be the 2-dimensional subspace of $\mathscr{R}^3$ with equation $x_1 - x_2 + 2x_3 = 0$.

$$\Rightarrow W \text{ has a basis } \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} \right\}, \text{ and with } C = \begin{bmatrix} 1 & -2 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$P_W = \frac{1}{6} \begin{bmatrix} 5 & 1 & -2 \\ 1 & 5 & 2 \\ -2 & 2 & 2 \end{bmatrix}. \text{ Note } P_W [\ 1 \ \ 3 \ \ 4\ ]^T = [\ 0 \ \ 4 \ \ 2\ ]^T.$$

Closest Vector Property

Let W be a subspace of $\mathcal{R}^n$ and $\mathbf{u}$ be a vector in $\mathcal{R}^n$. Among all vectors in W, the vector **closest** to $\mathbf{u}$ is the **orthogonal projection** $U_W(\mathbf{u})$ of $\mathbf{u}$ on W.



$$\forall\ \mathbf{w}' \in W,\ \mathbf{w} - \mathbf{w}' \in W.$$

$$\Rightarrow (\mathbf{u} - \mathbf{w}) \cdot (\mathbf{w} - \mathbf{w}') = 0.$$

$$\begin{aligned}
\Rightarrow \left\|\mathbf{u} - \mathbf{w}'\right\|^2 \\
&= \left\|(\mathbf{u} - \mathbf{w}) + (\mathbf{w} - \mathbf{w}')\right\|^2 \\
&= \left\|\mathbf{u} - \mathbf{w}\right\|^2 + \left\|\mathbf{w} - \mathbf{w}'\right\|^2 \\
&> \left\|\mathbf{u} - \mathbf{w}\right\|^2 \quad \forall\ \mathbf{w}' \neq \mathbf{w}.
\end{aligned}$$

Closest Vector Property

Let W be a subspace of $\mathcal{R}^n$ and $\mathbf{u}$ be a vector in $\mathcal{R}^n$. Among all vectors in W, the vector closest to $\mathbf{u}$ is the orthogonal projection $U_W(\mathbf{u})$ of $\mathbf{u}$ on W.

Definition

The **distance from a vector $\mathbf{u} \in \mathcal{R}^n$ to a subspace W of $\mathcal{R}^n$** is the distance between $\mathbf{u}$ and the orthogonal projection of $\mathbf{u}$ on W.

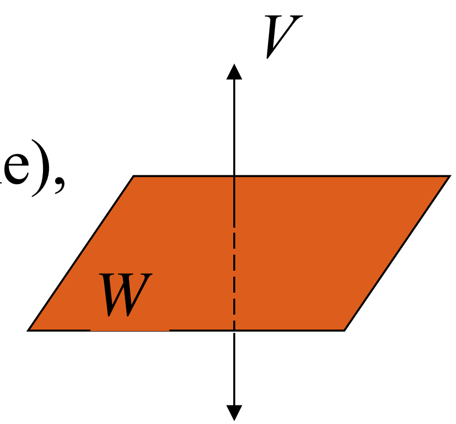
Example: For the solution space W of $x_1 - x_2 + 2x_3 = 0$ and the vector $\mathbf{v} = [\ 1\ \ 3\ \ 4\ ]^T \in \mathcal{R}^3$, the orthogonal projection of $\mathbf{v}$ onto W is $\mathbf{w} = [\ 0\ \ 4\ \ 2\ ]^T$. Thus the distance from $\mathbf{v}$ to W is

$$\|\mathbf{v} - \mathbf{w}\| = \|\mathbf{z}\| = \|\ [\ 1\ -1\ \ 2\ ]^T\ \| = \sqrt{6}$$

Homework Set for Section 6.3

- Section 6.3: Problems 2, 4, 5, 9, 11, 14, 21, 22, 24, 58, 62, 65-67, 71, 72, 75, 79, 82.

Important: An orthogonal complement of a subspace W is NOT just any shape (e.g., a line) that is "perpendicular" to W (e.g., a plane).



Ex. For $W = \{ [\, w_1 \;\; w_2 \;\; 0 \,]^T \mid w_1,\; w_2 \in \mathcal{R} \}$ (xy-plane),

$V = \{ [\, 0 \;\; 0 \;\; v_3 \,]^T \mid v_3 \in \mathcal{R} \}$ (z-axis)

is the orthogonal complement of W.

But $V' = \{ [\, 1 \;\; 0 \;\; v_3 \,]^T \mid v_3 \in \mathcal{R} \}$ is NOT an orthogonal complement of W even though the shape of V' looks "perpendicular" to the xy-plane!