MTH 100 : Lecture 36

- Last time we defined inner product and inner product space. Gave various examples of inner product space.
- We have defined orthogonality of two vectors.

We have shown that an orthogonal set of nonzero vectors is linearly independent.

Definition: If W is a subspace of V, then a vector $v \in V$ is said to be orthogonal to W if v is orthogonal to every vector in W. The set of all vectors orthogonal to W is called the orthogonal complement of W and is denoted by $W^{\perp}$ ('W-perp').

So, $W^{\perp} = \{v \in V : v \perp w \ \forall\, w \in W\}$

Proposition:

(a) v belongs to $W^{\perp}$ if and only if v is orthogonal to every vector in a spanning set for W.

(b) $W^{\perp}$ is a subspace of V and $W \cap W^{\perp} = \{0\}$

Proof:

(a) $\Rightarrow$: If $v \in W^{\perp}$, v is actually orthogonal to every vector of W and hence orthogonal to every vector in a spanning set of W.

$\Leftarrow$: Suppose v is orthogonal to every vector in a spanning set K for W

Let w be any vector of W.

Then w can be written as

$$w = c_1 w_1 + c_2 w_2 + \cdots + c_p w_p$$

where $w_1, w_2, \dots, w_p \in K$

Now $\langle w, v \rangle = \langle c_1 w_1 + c_2 w_2 + \cdots + c_p w_p, v \rangle$

$= c_1 \langle w_1, v \rangle + c_2 \langle w_2, v \rangle + \cdots + c_p \langle w_p, v \rangle$

$= 0$ (By hypothesis)

Therefore $v \in W^{\perp}$

(b) Since zero vector is orthogonal to every vector,

$\underset{\text{(zero vector)}}{0} \in W^{\perp}$

Now let $v_1, v_2 \in W^{\perp}$

Then for any vector $w \in W$,

$\langle v_1 + v_2, w \rangle = \langle v_1, w \rangle + \langle v_2, w \rangle = 0 + 0 = 0$

and so $v_1 + v_2 \in W^{\perp}$

Now let $v \in W^{\perp}$ and $c \in \mathbb{R}$

Then for any vector $w \in W$,

$\langle cv, w \rangle = c \langle v, w \rangle = c.0 = 0$

and so $cv \in W^{\perp}$.

Therefore $W^{\perp}$ is a subspace of V.

If $\omega \in W \cap W^{\perp}$, then $\omega \in W$ and $\omega \in W^{\perp}$
$\Rightarrow \langle \omega, \omega \rangle = 0 \Rightarrow \omega = 0$

Thus $\boxed{W \cap W^{\perp} = \{0\}}$

<u>Note:</u> Actually if S is any subset of V, then
$S^{\perp} = \{ v \in V : v \perp u \ \forall u \in S \}$ is
a subspace of V (even if S is not a subspace).

<u>Proof:</u> Since zero vector is orthogonal to every vector, $\underset{\text{(zero vector)}}{0} \in S^{\perp}$

Now let $v_1, v_2 \in S^{\perp}$
Then for any $u \in S$, $\langle v_1 + v_2, u \rangle = \langle v_1, u \rangle + \langle v_2, u \rangle$
$= 0 + 0 = 0$

$\Rightarrow v_1 + v_2 \in S^{\perp}$

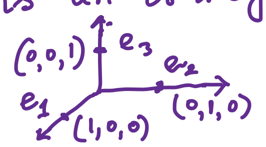
Now let $v \in S^{\perp}$ and $c \in \mathbb{R}$
Then for any $u \in S$, $\langle cv, u \rangle = c \langle v, u \rangle = c.0 = 0$
$\Rightarrow cv \in S^{\perp}$

Therefore $S^{\perp}$ is a subspace of V.

<u>Orthogonal Bases</u>

<u>Definition:</u>
- An orthogonal basis for a vector space V is a basis which is also an orthogonal set.
- An orthogonal basis for a subspace W is a <u>basis</u> which is also <u>an orthogonal set.</u>

<u>Example:</u> $\{e_1, e_2, e_3\}$ is an orthogonal basis for $\mathbb{R}^3$.



Note: • A set of vectors in a vector space is called Orthonormal if it is an orthogonal set and norm of each vector in the set is 1.

• A set of vectors in a vector space V is called an orthonormal basis if it is an orthogonal basis and norm of each vector in the basis is 1.

example: $\{e_1, e_2, e_3\}$ is an orthonormal basis for $\mathbb{R}^3$.

Proposition:

Let $\{u_1, u_2, \ldots, u_p\}$ be an orthogonal basis for a subspace W. Then if $y = c_1u_1 + c_2u_2 + \cdots + c_pu_p$ is any vector in W,

we have: $$c_j = \frac{\langle y, u_j\rangle}{\langle u_j, u_j\rangle} \quad \text{for } j = 1, \ldots, p$$

Proof: Since $y \in W$, y is uniquely expressible as a linear combination of the basis vectors.

i.e. $y = c_1u_1 + c_2u_2 + \cdots + c_pu_p$

To determine the coefficients $c_1, c_2, \ldots, c_p$

let us take the innerproduct with u_j for $1 \le j \le p$.

Then $$\langle y, u_j\rangle = c_1\langle u_1, u_j\rangle + c_2\langle u_2, u_j\rangle + \cdots + c_j\langle u_j, u_j\rangle + \cdots + \cdots + c_p\langle u_p, u_j\rangle$$

$$= 0 + 0 + \cdots + c_j\langle u_j, u_j\rangle + 0 + \cdots + 0$$

$$= c_j\langle u_j, u_j\rangle$$

$$\Rightarrow \boxed{c_j = \frac{\langle y, u_j\rangle}{\langle u_j, u_j\rangle}}$$

Note: • The above proposition shows that it is easy to find the coordinates of a vector relative to an orthogonal basis it is only needed to take an inner product and divide by the inner product of the basis vector with itself.

• If it is an orthonormal basis, then the length of each basis vector is 1 and even the step of division is avoided.

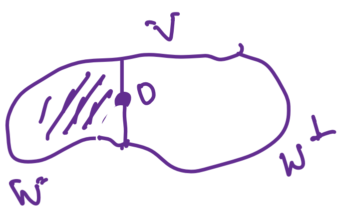
Orthogonal Decomposition

Theorem (Orthogonal Decomposition Theorem):

Let W be any finite-dimensional subspace of a vector space V. Then each vector y in V can be written uniquely in the form $y = \hat{y} + z$ where $\hat{y} \in W$ and $z \in W^\perp$

In fact, if $\{u_1, u_2, \dots, u_p\}$ is any orthogonal basis of W, then $\hat{y} = c_1 u_1 + c_2 u_2 + \dots + c_p u_p$

where $c_j = \dfrac{\langle y, u_j \rangle}{\langle u_j, u_j \rangle}$ for $j = 1, 2, \dots, p$



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

$y \in V$.

$y = \underbrace{\hat{y}}_{\in W} + z \in W^\perp$

Alternative Statement:

Given any finite-dimensional subspace W of V, we can then express $V = W + W^\perp$ with $W \cap W^\perp = \{0\}$

i.e. $V = W \oplus W^\perp$

Note: Thus every vector $v \in V$ can be uniquely expressed as a sum of a vector in W and a vector in $W^\perp$ i.e. as the sum of two vectors which are orthogonal to each other.

Note: The vector $\hat{y}$ is called the orthogonal projection of y onto W and written as

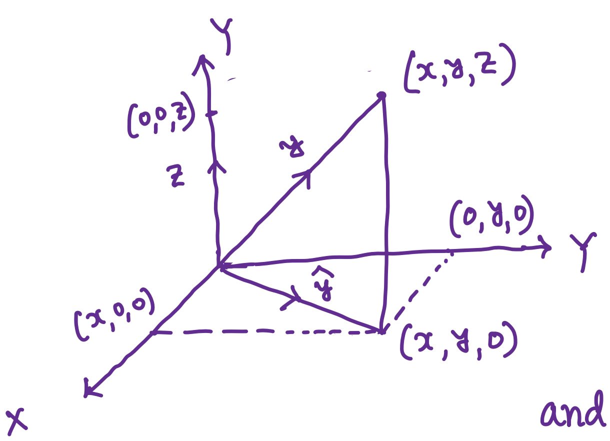
$\text{Proj}_W\, y = \hat{y}$ where $y = \hat{y} + z$ as in the Orthogonal Decomposition Theorem.

- In case $W = \text{Span}\{u\}$ is a one dimensional subspace of W, the expression for $\hat{y}$ is simplified to: $\hat{y} = \dfrac{\langle y, u\rangle}{\langle u, u\rangle} u$, which is simply called the orthogonal projection of y onto u.

Note: In case $y \in W$, its orthogonal projection onto W is y itself, i.e. $\boxed{\hat{y} = y \text{ for } y \in W}$

Example

Let $V = \mathbb{R}^3$, $W = \{(u, v, 0) : u, v \in \mathbb{R}\}$
i.e. W is the XY-plane.



For any $(x, y, z) \in \mathbb{R}^3$,

$(x, y, z) = (x, y, 0) + (0, 0, z)$

$= (x e_1 + y e_2) + z e_3$

Here $(x, y, 0) \in W$ and $(0, 0, z) \in W^{\perp}$

and $\boxed{\text{Proj}_W (x, y, z) = (x, y, 0)}$

Theorem (The Gram-Schmidt Process):

Given a basis $\{x_1, x_2, \dots, x_p\}$ for a subspace W of V, we can generate an orthogonal basis $\{v_1, v_2, \dots, v_p\}$ for W such that
$\text{Span}\{v_1, v_2, \dots, v_k\} = \text{Span}\{x_1, x_2, \dots, x_k\}$ for $k = 1, 2, \dots, p$.

In fact the vectors v_j are defined as follows:

$$v_1 = x_1$$

$$v_2 = x_2 - \left(\frac{\langle x_2, v_1\rangle}{\langle v_1, v_1\rangle}\right) v_1$$

$$v_3 = x_3 - \left(\frac{\langle x_3, v_1\rangle}{\langle v_1, v_1\rangle}\right) v_1 - \left(\frac{\langle x_3, v_2\rangle}{\langle v_2, v_2\rangle}\right) v_2$$

$$\vdots \qquad \vdots$$

$$v_p = x_p - \left(\frac{\langle x_p, v_1\rangle}{\langle v_1, v_1\rangle}\right) v_1 - \left(\frac{\langle x_p, v_2\rangle}{\langle v_2, v_2\rangle}\right) v_2 - \cdots\cdots \left(\frac{\langle x_p, v_{p-1}\rangle}{\langle v_{p-1}, v_{p-1}\rangle}\right) v_{p-1}$$

Note:

- At each stage, we subtract from the original basis vector x_i its projection onto the span of the previously obtained orthogonal vectors $v_1, v_2, \dots, v_{i-1}$.

- The process uses the idea we already used in Orthogonal Decomposition Theorem, of subtracting the orthogonal projection onto a subspace from the original vector.
- A formal proof that the vectors $\{v_1, v_2, \dots, v_k\}$ form an orthogonal set and that $\text{Span}\{v_1, v_2, \dots, v_k\} = \text{Span}\{x_1, x_2, \dots, x_k\}$ can be done by induction on k.
- We can obtain an <u>orthonormal basis</u> for every subspace W of V by normalizing each vector in an <u>orthogonal basis</u> (dividing each of the vectors by its norm). This step is usually left to the end because square roots can emerge.

<u>Ex:</u>
Construct an orthonormal basis for $\mathbb{R}^3$ starting with the basis using Gram-Schmidt process:

$$x_1 = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}, \quad x_2 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \quad x_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

First put $v_1 = x_1 = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$

Then $v_2 = x_2 - \dfrac{\langle v_1, x_2 \rangle}{\langle v_1, v_1 \rangle} v_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} - \dfrac{13}{14} \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$

$$= \begin{bmatrix} -\frac{6}{7} \\ \frac{15}{14} \\ \frac{3}{14} \end{bmatrix}$$

$$v_3 = x_3 - \frac{\langle v_1, x_3 \rangle}{\langle v_1, v_1 \rangle} v_1 - \frac{\langle v_2, x_3 \rangle}{\langle v_2, v_2 \rangle} v_2$$

$$= \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \frac{6}{14} \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} - \frac{-\frac{6}{7} + \frac{15}{14} + \frac{3}{14}}{\frac{12^2 + 15^2 + 3^2}{14^2}} \begin{bmatrix} -\frac{6}{7} \\ \frac{15}{14} \\ \frac{3}{14} \end{bmatrix}$$

$$= \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \frac{3}{7} \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} - \frac{2}{9} \begin{bmatrix} -\frac{6}{7} \\ \frac{15}{14} \\ \frac{3}{14} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{3} \\ \frac{1}{3} \\ -\frac{1}{3} \end{bmatrix}$$

<u>Check</u>: v_1, v_2 and v_3 are orthogonal (whereas the original basis vectors x_1, x_2, x_3 were not).

$$\langle v_1, v_2 \rangle = \left\langle \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}, \begin{bmatrix} -\frac{6}{7} \\ \frac{15}{14} \\ \frac{3}{14} \end{bmatrix} \right\rangle = -\frac{12}{7} + \frac{15}{14} + \frac{9}{14}$$

$$= \frac{-24 + 24}{14} = 0$$

$$\langle v_2, v_3 \rangle = \left\langle \begin{bmatrix} -\frac{6}{7} \\ \frac{15}{14} \\ \frac{3}{14} \end{bmatrix}, \begin{bmatrix} \frac{1}{3} \\ \frac{1}{3} \\ -\frac{1}{3} \end{bmatrix} \right\rangle = -\frac{6}{21} + \frac{15}{42} - \frac{3}{42}$$

$$= \frac{-12+15-3}{42} = 0$$

and $$\langle v_1, v_3 \rangle = \left\langle \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}, \begin{bmatrix} \frac{1}{3} \\ \frac{1}{3} \\ -\frac{1}{3} \end{bmatrix} \right\rangle = \frac{2}{3} + \frac{1}{3} - 1$$

$$= 1 - 1 = 0$$

<u>Note:</u> If we want an orthonormal basis, we divide each vector v_i by its length $\|v_i\|$ to get:

$$v_1' = \frac{1}{\sqrt{14}} \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}, \quad v_2' = \frac{1}{\sqrt{42}} \begin{bmatrix} -4 \\ 5 \\ 1 \end{bmatrix}, \quad v_3' = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$$

<u>Ex:</u> Given an orthogonal basis β:

$$v_1 = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}, \quad v_2 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \text{ and } v_3 = \begin{bmatrix} -1 \\ 4 \\ -1 \end{bmatrix},$$

find the coordinate of $v = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ with respect to this basis β.

<u>First check</u>

$$\langle v_1, v_2 \rangle = 2 \times 1 + 1 \times 0 + 2 \times (-1) = 0$$

$$\langle v_2, v_3 \rangle = 1 \times (-1) + 0 \times 4 + (-1) \times (-1) = 0$$

$$\langle v_1, v_3 \rangle = 2 \times (-1) + 1 \times 4 + 2 \times (-1) = 0$$

Now if $v = c_1 v_1 + c_2 v_2 + c_3 v_3$

then $c_1 = \dfrac{\langle v, v_1\rangle}{\langle v_1, v_1\rangle} = \dfrac{2\times 1 + 1\times 2 + 2\times 3}{2^2 + 1^2 + 2^2}$

$= \dfrac{10}{9}$

$c_2 = \dfrac{\langle v, v_2\rangle}{\langle v_2, v_2\rangle} = \dfrac{1\times 1 + 0\times 2 + (-1)\times 3}{1^2 + 0^2 + 1^2} = -\dfrac{2}{2}$

$= -1$

$c_3 = \dfrac{\langle v, v_3\rangle}{\langle v_3, v_3\rangle} = \dfrac{1\times(-1) + 2\times 4 + 3\times(-1)}{(-1)^2 + 4^2 + (-1)^2}$

$= \dfrac{4}{18} = \dfrac{2}{9}$

So, $[v]_\beta = \begin{bmatrix} \frac{10}{9} \\ -1 \\ \frac{2}{9} \end{bmatrix}$

Check:

$$\frac{10}{9}\begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix} + (-1)\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} + \frac{2}{9}\begin{bmatrix} -1 \\ 4 \\ -1 \end{bmatrix} = \begin{bmatrix} \frac{20}{9} - 1 - \frac{2}{9} \\ \frac{10}{9} + \frac{8}{9} \\ \frac{20}{9} - \frac{2}{9} \end{bmatrix}$$

$$= \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

Some additional Results

Proposition (Pythagorean Theorem):

u and v are orthogonal to each other if and only if $\|u+v\|^2 = \|u\|^2 + \|v\|^2$

Proof:

$$\|u+v\|^2 = \langle u+v, u+v \rangle$$
$$= \langle u,u \rangle + \langle u,v \rangle + \langle v,u \rangle + \langle v,v \rangle$$
$$= \|u\|^2 + \|v\|^2 + 2\langle u,v \rangle$$

Therefore $\|u+v\|^2 = \|u\|^2 + \|v\|^2$

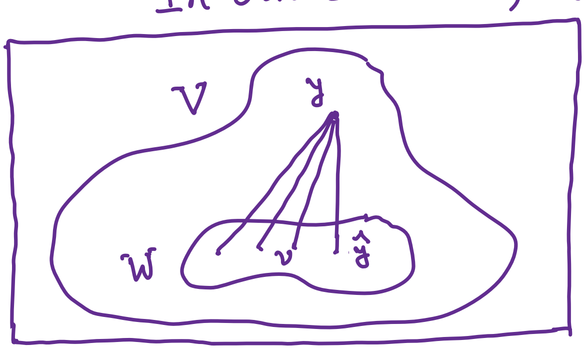
$$\Leftrightarrow \langle u,v \rangle = 0 \Leftrightarrow u \perp v$$

Proposition (Best Approximation Theorem):

Let W be any finite dimensional subspace of V, y any vector in V and $\hat{y}$ be the orthogonal projection of y onto W.

Then $\|y-\hat{y}\| < \|y-v\|$ for all $v \in W$ distinct from $\hat{y}$.

In otherwords, $\hat{y}$ is the closest vector (point) in W to y



Proof: Let $v \in W$ which is distinct from $\hat{y}$

Then $\|y - v\|^2 = \langle y-v, y-v \rangle$

$$= \langle y - \hat{y} + \hat{y} - v, y - \hat{y} + \hat{y} - v \rangle$$

$$= \langle (y-\hat{y}) + (\hat{y}-v), (y-\hat{y}) + (\hat{y}-v) \rangle$$

$$= \langle y-\hat{y}, y-\hat{y} \rangle + \langle \hat{y}-v, \hat{y}-v \rangle + \langle y-\hat{y}, \hat{y}-v \rangle + \langle \hat{y}-v, y-\hat{y} \rangle$$

$$= \|y-\hat{y}\|^2 + \|\hat{y}-v\|^2 + 2\langle y-\hat{y}, \hat{y}-v \rangle$$

Now $y - \hat{y} \in W^{\perp}$ (Note that $y = \hat{y} + (y-\hat{y})$ and $\hat{y} \in W$)

and $\hat{y} - v \in W$ (Since $\hat{y}, v \in W$ and W is a subspace of V)

So, $\langle \hat{y}-v, y-\hat{y} \rangle = 0$

Therefore $\|y-v\|^2 = \|y-\hat{y}\|^2 + \|\hat{y}-v\|^2$

Now if $y = v$, then $y \in W \Rightarrow y = \hat{y} = v$ which is not allowed as v is distinct from $\hat{y}$.

Hence $\|y-v\|^2 > 0$

and therefore $\|y-v\| > \|y-\hat{y}\|$

i.e. $\boxed{\|y-\hat{y}\| < \|y-v\|}$

Corollary: If y is any vector and W is a finite-dimensional subspace, then

$$\|\text{proj}_W y\| \leq \|y\|$$

Proof: We know that
$$y = \text{Proj}_W y + z \quad \text{where } z \in W^\perp$$
By Pythagorean Theorem,
$$\|y\|^2 = \|\text{Proj}_W y + z\|^2 = \|\text{Proj}_W y\|^2 + \|z\|^2$$
Since $\|z\|^2 \geq 0$, we get $\|y\|^2 \geq \|\text{Proj}_W y\|^2$
$$\Rightarrow \boxed{\|\text{Proj}_W y\| \leq \|y\|}$$

Proof of Cauchy-Schwarz Inequality (using the above corollary)

- Want to prove that $|\langle u,v\rangle| \leq \|u\|\|v\| \ \forall u,v \in V$.

Clearly the result holds if either $u=0$ or $v=0$. So, we may assume that both u and v are non-zero and apply the corollary above taking $W = \text{span}\{v\}$

Then $\|\text{Proj}_W u\| \leq \|u\|$

Now $\text{Proj}_W u = \dfrac{\langle u,v\rangle}{\langle v,v\rangle} v$

So, $\left\| \dfrac{\langle u,v\rangle}{\langle v,v\rangle} v \right\| \leq \|u\|$
$$\Rightarrow \frac{|\langle u,v\rangle|\,\|v\|}{\|v\|^2} \leq \|u\|$$

$$\Rightarrow \quad \frac{|\langle u,v\rangle|}{\|v\|} \leq \|u\|$$

$$\Rightarrow \quad \boxed{|\langle u,v\rangle| \leq \|u\|\,\|v\|}$$

Proof of Orthogonal Decomposition Theorem:

Note: In this proof we assume that any finite-dimensional subspace W of an inner product space has an orthogonal basis.

This assumption is Gram-Schmidt Process which has been covered later.

However the proof of Gram-Schmidt Process does not depend on Orthogonal Decomposition Theorem and so the assumption is logically valid.

- First we prove the uniqueness of the decomposition.

Suppose $y \in V$ and $y = \hat{y} + z$ and $y = \hat{y}_1 + z_1$ where $\hat{y}, \hat{y}_1 \in W$ and $z, z_1 \in W^{\perp}$

Subtracting, $0 = (\hat{y} - \hat{y}_1) + (z - z_1)$

$$\Rightarrow \quad \hat{y}_1 - \hat{y} = z - z_1$$

Now $\hat{y}_1 - \hat{y} \in W$ and $z - z_1 \in W^{\perp}$

So, $\underset{(= z - z_1)}{\hat{y}_1 - \hat{y}} \in W \cap W^{\perp}$

Since $W \cap W^{\perp} = \{0\}$, we get $\left.\begin{array}{l}\hat{y}_1 - \hat{y} = 0 \Rightarrow \hat{y} = \hat{y}_1 \\ \text{and } z - z_1 = 0 \Rightarrow z = z_1\end{array}\right\}$

Thus the decomposition is unique.

- Now we will prove that such a decomposition exists.

i.e. any $y \in V$ can be written as

$$y = \hat{y} + z \quad \text{where } \hat{y} \in W \text{ and } z \in W^{\perp}.$$

Let $\{u_1, u_2, \ldots, u_p\}$ be an orthogonal basis of W.

Let $\hat{y} = c_1 u_1 + \cdots + c_p u_p$

$$\left.\text{where } c_j = \frac{\langle y, u_j \rangle}{\langle u_j, u_j \rangle}\right\} \text{ for } j = 1, 2, \ldots, p.$$

Now $\hat{y} \in W$.

Let $z = y - \hat{y}$ Then $y = \hat{y} + z$

$$\begin{aligned}
\text{Now } \langle z, u_j \rangle &= \langle y - \hat{y}, u_j \rangle = \langle y, u_j \rangle - \langle \hat{y}, u_j \rangle \\
&= \langle y, u_j \rangle - \langle c_1 u_1 + \cdots + c_p u_p, u_j \rangle \\
&= \langle y, u_j \rangle - c_1 \langle u_1, u_j \rangle - \cdots - c_p \langle u_p, u_j \rangle \\
&= \langle y, u_j \rangle - c_j \langle u_j, u_j \rangle \\
&= \langle y, u_j \rangle - \frac{\langle y, u_j \rangle}{\langle u_j, u_j \rangle} \langle u_j, u_j \rangle \\
&= \langle y, u_j \rangle - \langle y, u_j \rangle = 0
\end{aligned}$$

for $j = 1, 2, \ldots, p$.

So, $z \perp u_j$ for $j = 1, 2, \ldots, p$

Since $\{u_1, u_2, \ldots, u_p\}$ is a basis of W,

$z \perp W \Rightarrow z \in W^{\perp}$

Therefore $\boxed{y = \hat{y} + z}$ where $\hat{y} \in W$ and $z \in W^{\perp}$.

$\boxed{\text{QED}}$