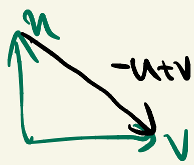
Week 13 : Inner product.

Goal : Introduce "product" of vectors $u, v \in \mathbb{R}^n$.

Motivation : Pythagorean thm.

2D:



$\rightarrow \|v-u\|^2 = \|u\|^2 + \|v\|^2$

where $\|\cdot\|$ = length of vector.

think of $\|u\|^2 = \langle u, u \rangle$ product with itself.

In $\mathbb{R}^n$, $u = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} \in \mathbb{R}^n \rightarrow \|u\|$ = length of u

$= \sqrt{\sum_{i=1}^{n} u_i^2}$ (Pyth. Thm)

Defn : (Inner product) For $u, v \in \mathbb{R}^n$, we define

$$\langle u, v \rangle = \sum_{i=1}^{n} u_i v_i = u^t v,$$ (the inner product of u with v)

(Some ppl denote it using $u \cdot v$, called dot product.)

Then : $\|u\|^2 = \langle u, u \rangle = \sum_{i=1}^{n} u_i^2 = u^t u$. (= square of length).

Thm : For any $u, v, w \in \mathbb{R}^n$

① $\langle u, v \rangle = \langle v, u \rangle = \sum_{i=1}^{n} u_i v_i$

② $\langle \alpha u + \beta v, w \rangle = \alpha \langle u, w \rangle + \beta \langle v, w \rangle$, for $\alpha, \beta \in \mathbb{R}$.

③ $\|u\| \geq 0$. Moreover if $\|u\| = 0$, then $u = 0$.

$\rightarrow$ therefore, we also call $\|u\|$ as norm of vector u.

pf: By expanding everything ....

Thm (More properties) For all $u, v \in \mathbb{R}^n$

① $\|u \pm v\|^2 = \|u\|^2 + \|v\|^2 \pm 2\langle u, v\rangle$

② $\|u+v\|^2 + \|u-v\|^2 = 2\left(\|u\|^2 + \|v\|^2\right)$.

---

pf: ①: $\|u \pm v\|^2 = \langle u \pm v, u \pm v\rangle$

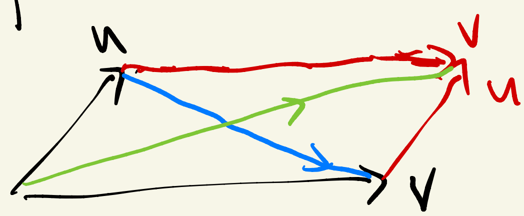
$= \langle u \pm v, u\rangle + \langle u \pm v, v\rangle$

$= \langle u, u\rangle \pm \langle v, u\rangle + \langle u, v\rangle \pm \langle v, v\rangle$

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

②: Adding up $\pm$ case in ① #

Geometric picture of ②:



In 2D, ② follows from cosine law

(Before we move on) Q: why care??

In 2D,

$\|u - v\|^2 = \|u\|^2 + \|v\|^2$ (Pyth Thm)

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

$$\boxed{* \quad u \perp v \Leftrightarrow \langle u, v\rangle = 0.}$$

Defn: Let $u, v \in \mathbb{R}^n$, we say that u is orthogonal to v ($u \perp v$) if $\langle u, v\rangle = 0$.

Thm: For $u, v \in \mathbb{R}$

① $u \perp v \Leftrightarrow v \perp u$

② $u \perp u \Leftrightarrow u = 0$

③ $u \perp w \;\; \forall w \in \mathbb{R}^n \Leftrightarrow u = 0$

④ $\|u+v\|^2 = \|u\|^2 + \|v\|^2 \Leftrightarrow u \perp v$.

In 2D: $\langle u, v\rangle = \|u\| \cdot \|v\| \cdot \cos\theta$ (seen by cosine law)

where 

$\Rightarrow |\langle u, v\rangle| \leq \|u\| \cdot \|v\|$ since $|\cos\theta| \leq 1$.

Moreover equality holds iff $\cos\theta = 1$ or -1

iff $u \parallel v$.

In higher dimension:

Thm (Cauchy-Schwarz inequality)

Let $u, v \in \mathbb{R}^n$, we have $|\langle u, v\rangle| \leq \|u\| \cdot \|v\|$.

Moreover, equality holds iff u, v are linearly dependent.

---

pf: If $\|u\| = 0$ or $\|v\| = 0$, then trivially true.

May assume $u, v \neq 0$.

Let $f(t) = \langle u + tv, u + tv\rangle$ for $t \in \mathbb{R}$.

Then $f(t) = \|u\|^2 + 2t\langle u, v\rangle + t^2\|v\|^2 \geq 0$, $\forall t \in \mathbb{R}$

$\Rightarrow |2\langle u, v\rangle|^2 - 4 \cdot \|u\|^2 \|v\|^2 \leq 0$.

$\Rightarrow |\langle u, v\rangle| \leq \|u\| \cdot \|v\|$ #

Moreover, equality holds $\Rightarrow f(t)$ has a unique real root, λ.

$\Rightarrow u + \lambda v = 0 \Rightarrow u, v$ are linearly dep.

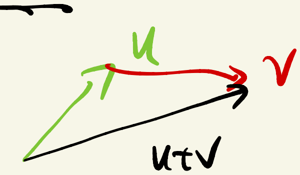
($\Leftarrow$): if u, v are linearly dep., then it is trivial.

Another important properties:

Thm (Δ inequality) For $u, v \in \mathbb{R}^n$,

$\|u + v\| \leq \|u\| + \|v\|$

Picture:



Moreover, equality holds iff u, v are non-neg. multiple of each other.

pf: $\|u+v\|^2 = \|u\|^2 + \|v\|^2 + 2\langle u, v\rangle$
$\leq \|u\|^2 + \|v\|^2 + 2\|u\| \cdot \|v\|$ (Cauchy-Schwarz ineq.)
$= (\|u\| + \|v\|)^2$

Equality holds only if $\langle u, v\rangle = \|u\| \cdot \|v\|$.

If $u, v \neq 0 \Rightarrow u = \lambda v$ for some $\lambda \in \mathbb{R}$.
if $\lambda < 0$, then impossible since RHS ≥ 0.
$\therefore \lambda \geq 0$.

If $u = 0$ OR $v = 0$, then conclusion holds.

($\Leftarrow$): If u, v are nonneg multiple of each other, conclusion holds trivially.

☆☆ Relation to linear independence.

Thm: Suppose $u_1, u_2, \cdots, u_m$ are non-zero vectors in $\mathbb{R}^n$ s.t. $\langle u_i, u_j\rangle = 0 \ \forall i \neq j$ (pairwise orthogonal)

Then, ① $u_1, u_2, \cdots, u_m$ are linearly indep

② (Unique representation using inner product)

if $v \in \text{span}\{u_1, u_2, \cdots, u_m\}$, then

$$v = \sum_{i=1}^{m} \frac{\langle v, u_i\rangle}{\|u_i\|} \cdot \boxed{\frac{u_i}{\|u_i\|}}$$

unit vector along the direction of u_i

pf: ① Suppose $\alpha_i, (i=1,2,\cdots,m) \in \mathbb{R}$ s.t.

$$\sum_{i=1}^{m} \alpha_i u_i = 0$$

$$\Rightarrow \quad \alpha_1 u_1 = -\sum_{i=2}^{m} \alpha_i u_i$$

$$\Rightarrow \quad \alpha_1 \langle u_1, u_1 \rangle = \alpha_1 \|u_1\|^2 = -\sum_{i=2}^{m} \alpha_i \langle u_i, u_1 \rangle = 0$$

$$\Rightarrow \quad \alpha_1 \|u_1\|^2 = 0 \quad \Rightarrow \quad \alpha_1 = 0$$

Replace $\alpha_1 u_1$ by $\alpha_j u_j$ for other j $\Rightarrow$ $\alpha_i = 0 \ \forall i = 1, 2 \ldots m$. #

② If $v \in \text{span}\{u_1, u_2, \cdots, u_m\}$ ← dim = m by ①

then $v = \sum_{i=1}^{m} \alpha_i u_i$ for some (unique) $\alpha_i \in \mathbb{R}$.

Find α_i: $\langle v, u_j \rangle = \sum_{i=1}^{m} \alpha_i \langle u_i, u_j \rangle = \alpha_j \|u_j\|^2$

$$\Rightarrow \quad \alpha_j = \frac{\langle v, u_j \rangle}{\|u_j\|^2}. \qquad \forall j = 1, 2, \cdots, m$$

$$\Rightarrow \quad v = \sum_{j=1}^{m} \frac{\langle v, u_j \rangle}{\|u_j\|} \frac{u_j}{\|u_j\|}. \quad \#$$

---

typical eg: in $\mathbb{R}^n$, $u_i = e_i = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix}$ ← i-th entry

then $\{e_1, e_2, \cdots, e_n\}$ is a basis of $\mathbb{R}^n$

s.t. $\langle e_i, e_j \rangle = 0$ if $i \neq j$ and $\|e_i\| = 1 \ \forall i = 1, \cdots n$.

Defn: Let $u_1, u_2, \cdots, u_k \in \mathbb{R}^n$.

(a) We say that $\{u_1, u_2, \cdots, u_k\}$ is an orthonormal set in $\mathbb{R}^n$

if $\quad \|u_i\| = 1 \ \forall i = 1, 2, \cdots, k \quad$ and

$\langle u_i, u_j \rangle = 0 \ \forall i \neq j$.

(b) Suppose V = subspace of $\mathbb{R}^n$. then $u_1, u_2, \cdots, u_k$ is an orthonormal basis if (a) holds and

$\{u_1, u_2, \cdots, u_k\}$ is a basis of V.

Why so special??

Thm: If $\{v_1, v_2, \cdots, v_k\}$ is an orthonormal basis of V, a subspace of $\mathbb{R}^n$. If $u, w \in V$,

then
$$\begin{cases} u = \sum_{i=1}^{k} \langle u, v_i \rangle v_i \\ w = \sum_{i=1}^{k} \langle w, v_i \rangle v_i \end{cases}$$

① $\|u\|^2 = \sum_{i=1}^{k} |\langle u, v_i \rangle|^2$

② $\langle u, w \rangle = \sum_{i=1}^{k} \langle u, v_i \rangle \langle w, v_i \rangle$.

And

If $V = \mathbb{R}^n$

$$u = \sum_{i=1}^{n} u_i e_i$$

$$w = \sum_{i=1}^{n} w_i e_i$$

$$\langle u, w \rangle = \sum_{i=1}^{n} u_i w_i$$

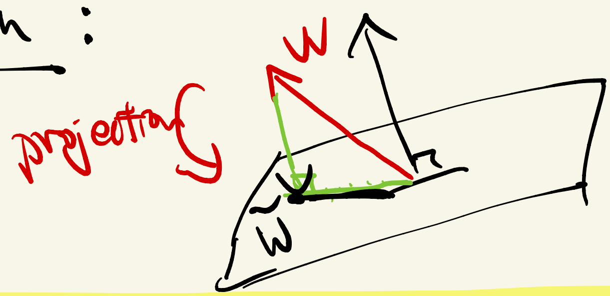
$$\|u\|^2 = \sum_{i=1}^{n} u_i^2$$

pf: let
$$\begin{cases} u = \sum_{i=1}^{k} u_i v_i , & u_i \in \mathbb{R} \\ w = \sum_{i=1}^{k} w_i v_i & w_i \in \mathbb{R} \end{cases}$$

- $$\|u\|^2 = \left\langle \sum_{i=1}^{k} u_i v_i , \sum_{j=1}^{k} u_j v_j \right\rangle$$
$$= \sum_{i=1}^{k} \sum_{j=1}^{k} u_i u_j \langle v_i , v_j \rangle = \sum_{i=1}^{k} u_i^2 \quad \left(\because \langle v_i , v_j \rangle = \delta_{ij} \right)$$

- $$\langle u , w \rangle = \sum_{i=1}^{k} \sum_{j=1}^{k} u_i w_j \langle v_i , v_j \rangle = \sum_{i=1}^{n} u_i w_i \qquad \#$$

Projection :



$V \subseteq \mathbb{R}^3$

$\dim(V) = 2$

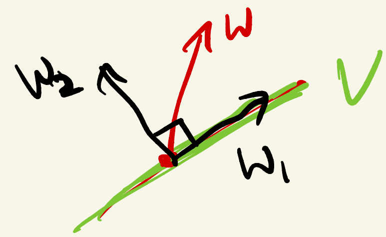
Goal: Given $w \in \mathbb{R}^3$, find the "Best" $\tilde{w} \in V$ which approximate w ??

Need inner product :

Idea : decompose $w = w_1 + w_2$ s.t. $w_1 \in V$, $w_2 \notin V$.

But there are ∞ many such choice !!

Best one: $w_1 \perp w_2$.



How to find w_1, w_2 ??

Thm: Let V be a subspace in $\mathbb{R}^n$ and $v_1, v_2, \cdots, v_k$ be an orthonormal basis of V. Let $w \in \mathbb{R}^n$ and define

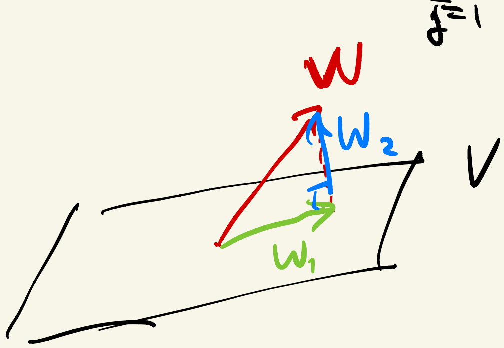
$$w_1 = \sum_{j=1}^{k} \langle w, v_j \rangle v_j \in V \quad \text{and} \quad w_2 = w - w_1$$

Then: ① $w_2 \perp v \quad \forall v \in V$.

② for $s \in V$, we have

$\|w - s\| \geq \|w - w_1\|$ and "equality hold iff $s = w_1$"

③ $\|w\|^2 \geq \sum_{j=1}^{k} |\langle w, v_j \rangle|^2$.



Meaning: $\mathrm{dist}(w, w_1)$ = distance from w to V.

pf: ①: Suffice to show that $\langle v_j, w_2 \rangle = 0 \quad \forall j = 1, 2, \cdots, k$

Since $\operatorname{span}\{v_1, v_2, \cdots, v_k\} = V$.

$$\langle v_j, w_2\rangle = \langle v_j, w - w_1\rangle$$
$$= \langle v_j, w\rangle - \langle v_j, \sum_{i=1}^{k} \langle w, v_i\rangle v_i\rangle$$
$$= \langle v_j, w\rangle - \cancelto{1}{\langle v_j, v_j\rangle} \langle w, v_j\rangle = 0.$$

② : $\|w - s\|^2 = \|w_1 + w_2 - s\|^2$
$$= \|w_1 - s\|^2 + \|w_2\|^2 \quad (\text{since } w_2 \perp V)$$
$$\geq \|w_2\|^2 = \|w - w_1\|^2.$$

Equality holds iff $w_1 = s$.

③ : $\|w\|^2 = \|w_1\|^2 + \|w_2\|^2 \geq \|w_1\|^2 = \sum_{j=1}^{k} |\langle w, v_j\rangle|^2$.

Defn: If V is a subspace of $\mathbb{R}^n$.

$$V^\perp = \{x \in \mathbb{R}^n \mid \langle x, v\rangle = 0 \ \forall v \in V\}$$
$$= \text{orthogonal complement of } V \text{ in } \mathbb{R}^n$$

normal to plane V with $\vec{n} \perp V$.

picture:

$V \in \mathbb{R}^3$

$V^\perp = \operatorname{span}\{\vec{n}\}$.

We call $w_2 = w - w_1$, the orthogonal projection on V

$= pr_V(W).$

<u>Example</u>:

$$v_1 = \frac{1}{2}\begin{bmatrix}1\\1\\1\\1\end{bmatrix}, \quad v_2 = \frac{1}{2}\begin{bmatrix}-1\\-1\\1\\1\end{bmatrix} \in \mathbb{R}^4$$

$$\|v_1\|^2 = \frac{1}{4}(1+1+1+1) = 1 = \|v_2\|$$

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

$\longrightarrow$ $\{v_1, v_2\}$ is orthonormal basis of span$\{v_1, v_2\}$

let $v_3 = \begin{bmatrix}x\\y\\z\\w\end{bmatrix} \in \mathbb{R}^4$, Then $pr_V(v_3) = \langle v_3, v_1\rangle v_1 + \langle v_3, v_2\rangle v_2$

$$\langle v_3, v_1\rangle = \frac{1}{2}(x+y+z+w)$$

$$\langle v_3, v_2\rangle = \frac{1}{2}(-x-y+z+w)$$

$$\therefore pr_V(v_3) = \frac{1}{4}(x+y+z+w)\begin{bmatrix}1\\1\\1\\1\end{bmatrix} + \frac{1}{4}(-x-y+z+w)\begin{bmatrix}-1\\-1\\1\\1\end{bmatrix}$$

<u>Q. left</u>: How to find orthonormal basis for a given vector subspace $V \in \mathbb{R}^n$??

<u>Thm</u>: Any non-zero vector subspace $V \subseteq \mathbb{R}^n$ admits an orthonormal basis. (Not unique !!)

<u>pf</u>: By <u>construction</u>: (called Gram-Schmidt process)

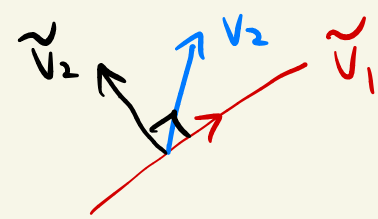
take a basis of $V = \text{span}\{v_1, v_2, \cdots, v_m\}$

where $\dim(V) = m > 0$. $\Rightarrow \|v_i\| \neq 0, \forall i = 1, 2, \cdots m$

$\boxed{\text{Goal}}$: Refine $v_1, v_2, \cdots, v_m$ s.t. they become orthonormal.

- when $k=1$: define $\tilde{v}_1 = \dfrac{v_1}{\|v_1\|}$ s.t. $\|\tilde{v}_1\| = 1$.

- when $k=2$, $\rightarrow$ If $v_2 \perp \tilde{v}_1$, then take $\tilde{v}_2 = \dfrac{v_2}{\|v_2\|}$



More generally, consider $\overline{v}_2 = v_2 - \langle v_2, \tilde{v}_1 \rangle \tilde{v}_1 = pr_{\text{span}\{\tilde{v}_1\}}(v_2)$

Note: $\overline{v}_2 \neq 0$ since otherwise $v_2 \in \text{span}\{v_1\}$

$\Rightarrow$ Not basis!!

take $\tilde{v}_2 = \dfrac{\overline{v}_2}{\|\overline{v}_2\|}$.

Induction: Suppose $\exists\ \tilde{v}_1, \tilde{v}_2, \cdots, \tilde{v}_\ell \in V$ s.t.

$$\text{span}\{v_1, v_2, \cdots, v_\ell\} = \text{span}\{\tilde{v}_1, \tilde{v}_2, \cdots, \tilde{v}_\ell\}$$

for some $\ell \in \{1, 2, \cdots, m\}$ s.t.

$$\langle \tilde{v}_i, \tilde{v}_j \rangle = 0 \quad \forall i \neq j \quad \text{and} \quad \|\tilde{v}_i\| = 1.$$

Then we construct $\tilde{v}_{l+1}$ (if $l < m$) as follow:

$$\bar{v}_{l+1} = v_{l+1} - \sum_{j=1}^{l} \langle v_{l+1}, \tilde{v}_j \rangle \tilde{v}_j$$

- $\bar{v}_{l+1} \neq 0$ since otherwise $v_{l+1} \in \operatorname{span}\{v_1, \cdots, v_l\}$ which is Impossible.

- take $\tilde{v}_{l+1} = \dfrac{\bar{v}_{l+1}}{\|\bar{v}_{l+1}\|}$.

Checking: $\forall\, i \neq j$, $i, j \in \{1, 2, \cdots, l+1\}$.

$\|\tilde{v}_i\| = 1$, (i) $\langle \tilde{v}_i, \tilde{v}_j \rangle = 0$ if $i, j < l+1$

(ii) $$\langle \tilde{v}_i, \tilde{v}_{l+1} \rangle = \frac{1}{\|\bar{v}_{l+1}\|} \langle \tilde{v}_i, \bar{v}_{l+1} \rangle$$

$$= \frac{1}{\|\bar{v}_{l+1}\|} \langle \tilde{v}_i, v_{l+1} - \sum_{m=1}^{l} \langle v_{l+1}, \tilde{v}_m \rangle \tilde{v}_m \rangle$$

$$= \frac{1}{\|\bar{v}_{l+1}\|} \left(\langle \tilde{v}_i, v_{l+1} \rangle - \langle v_{l+1}, \tilde{v}_i \rangle \right)$$

$$= 0.$$

(iii) $\operatorname{span}\{v_1, v_2, \cdots, v_\ell, v_{\ell+1}\}$

$= \operatorname{span}\{\tilde{v}_1, \tilde{v}_2, \cdots, \tilde{v}_\ell, \tilde{v}_{\ell+1}\}$ since.

(A) $\supseteq$: By construction and induction hypothesis

$\subseteq$: since $v_{\ell+1} = \tilde{v}_{\ell+1} + \sum_{m=1}^{\ell} \langle v_{\ell+1}, \tilde{v}_i \rangle \tilde{v}_i$

$\in \operatorname{span}\{v_1, v_2, \cdots, v_\ell, \tilde{v}_{\ell+1}\}$

By induction, we can find

$\{\tilde{v}_1, \tilde{v}_2, \cdots, \tilde{v}_m\}$ s.t. they span V

and B an orthonormal basis of V. #

Eg: $V = \operatorname{span}\left\{ \overset{v_1}{\begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}}, \overset{v_2}{\begin{bmatrix} 0 \\ 1 \\ 2 \\ 1 \end{bmatrix}}, \overset{v_3}{\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}} \right\}$

linearly independent (check by RREF)

Goal: find O.N. basis of V.

Step 1 : $\|v_1\| = \sqrt{1+1} = \sqrt{2}$

$\therefore$ take $u_1 = \frac{1}{\sqrt{2}} v_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}$.

Step 2 : $\langle v_2, u_1 \rangle = \frac{1}{\sqrt{2}} \langle v_1, v_2 \rangle = \frac{2}{\sqrt{2}} = \sqrt{2}$.

$$w_2 = v_2 - \langle v_2, u_1 \rangle u_1$$
$$= \begin{bmatrix} 0 \\ 1 \\ 2 \\ 1 \end{bmatrix} - \cancel{\sqrt{2}} \cdot \frac{1}{\cancel{\sqrt{2}}} \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

and $\|w_2\| = 2$

$\therefore$ take $u_2 = \frac{1}{2} w_2 = \frac{1}{2} \begin{bmatrix} -1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$.

Step 3 :

$$w_3 = v_3 - \langle v_3, u_1 \rangle u_1 - \langle v_3, u_2 \rangle u_2$$

$\langle v_3, u_1 \rangle = \frac{1}{\sqrt{2}} \langle v_3, v_1 \rangle = \sqrt{2}$.

$\langle v_3, u_2 \rangle = \frac{1}{2} \left\langle \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \right\rangle = 1$.

$\therefore w_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} - \cancel{\sqrt{2}} \cdot \frac{1}{\cancel{\sqrt{2}}} \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} -1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$

$$= \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} - \frac{1}{2}\begin{bmatrix} -1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 \\ 1 \\ -1 \\ 1 \end{bmatrix}.$$

and $\|W_3\| = \frac{1}{2}\cdot\sqrt{1+1+1+1} = 1$

$\therefore$ take $U_3 = W_3$

$\therefore V = \text{span}\{U_1, U_2, U_3\}$.

Moreover, $\cdot$ $\text{span}\{U_1\} = \text{span}\{V_1\}$

$\cdot$ $\text{span}\{U_1, U_2\} = \text{span}\{V_1, V_2\}$

$\cdot$ $\text{span}\{U_1, U_2, U_3\} = \text{span}\{V_1, V_2, V_3\}$