## The Gram-Schmidt Process

Let $\{\vec{v}_1, \cdots, \vec{v}_k\} \subset \mathbb{R}^n$. We say that $\{\vec{v}_1, \cdots, \vec{v}_k\}$ is an orthonormal (o.n.) set of vectors provided

$$\vec{v}_i \cdot \vec{v}_j = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases}.$$

Put another way, $|\vec{v}_i| = 1$ for $i = 1, \cdots, k$ and $\vec{v}_i$ and $\vec{v}_j$ are orthogonal whenever $i \neq j$. The standard basis for $\mathbb{R}^n$ is an example of an o.n. set of vectors in $\mathbb{R}^n$. As we shall see, an o.n. subset of $\mathbb{R}^n$ has much in common with the standard basis for $\mathbb{R}^n$.

<u>Proposition #41</u>: Suppose $\{\vec{v}_1, \cdots, \vec{v}_k\}$ is an o.n. set of vectors in $\mathbb{R}^n$. Then $\{\vec{v}_1, \cdots, \vec{v}_k\}$ is a linearly independent set of vectors.

<u>Proof</u>: Suppose $\sum_{i=1}^{k} c_i \vec{v}_i = \vec{0}$. Then, if $j \in \{1, \cdots, k\}$, we have

$$
\begin{aligned}
0 &= \vec{0} \cdot \vec{v}_j \\
&= \left( \sum_{i=1}^{k} c_i \vec{v}_i \right) \cdot \vec{v}_j \\
&= \sum_{i=1}^{k} c_i \left( \vec{v}_i \cdot \vec{v}_j \right).
\end{aligned}
$$

Since $\vec{v}_i \cdot \vec{v}_j = 0$ if $i \neq j$ and $\vec{v}_j \cdot \vec{v}_j = 1$, $\sum_{i=1}^{k} c_i (\vec{v}_i \cdot \vec{v}_j) = c_j$.

It follows that $c_j = 0$, for $j = 1, \cdots, k$.

Thus, $\{\vec{v}_1, \cdots, \vec{v}_k\}$ is linearly independent.

**Proposition #42:** Suppose $\{\vec{v}_1, \cdots, \vec{v}_k\}$ is an o.n. set of vectors in $\mathbb{R}^n$ and $\vec{w} \in \text{span}(\{\vec{v}_1, \cdots, \vec{v}_k\})$. Then

$$\vec{w} = \sum_{i=1}^{k} (\vec{w} \cdot \vec{v}_i)\, \vec{v}_i \,.$$

**Proof:** Since $\vec{w} \in \text{span}(\{\vec{v}_1, \cdots, \vec{v}_k\})$ there exist $c_1, \cdots, c_k \in \mathbb{R}$ such that

$$\vec{w} = \sum_{i=1}^{k} c_i \vec{v}_i \,.$$

If $j \in \{1, \cdots, k\}$, we the have

$$\begin{aligned} \vec{w} \cdot \vec{v}_j &= \left( \sum_{i=1}^{k} c_i \vec{v}_i \right) \cdot \vec{v}_j \\ &= \sum_{i=1}^{k} c_i \,(\vec{v}_i \cdot \vec{v}_j) \\ &= c_j \,. \end{aligned}$$

Thus, $c_j = \vec{w} \cdot \vec{v}_j$ for $j = 1, \cdots, k$. Therefore

$$\vec{w} = \sum_{i=1}^{k} (\vec{w} \cdot \vec{v}_i)\, \vec{v}_i \,.$$

<u>Proposition #43</u>: Suppose $\{\vec{v}_1, \cdots, \vec{v}_k\}$ is an o.n. set of vectors in $\mathbb{R}^n$ and $\vec{w} \in \mathbb{R}^n$. Then $\vec{w} - \sum_{i=1}^{k} (\vec{w} \cdot \vec{v}_i)\vec{v}_i$ is perpendicular to every vector in $\text{span}(\{\vec{v}_1, \cdots, \vec{v}_k\})$.

<u>Proof</u>: Let us make an initial observation.

<u>Claim</u>: Suppose $\vec{u} \in \mathbb{R}^n$ and $\vec{u} \cdot \vec{v}_i = 0$, for $i = 1, \cdots, k$. Then $\vec{u}$ is perpendicular to every vector in $\text{span}(\{\vec{v}_1, \cdots, \vec{v}_k\})$.

<u>Proof of Claim</u>: The typical element in $\text{span}(\{\vec{v}_1, \cdots, \vec{v}_k\})$ is of the form $\sum_{i=1}^{k} c_i \vec{v}_i$, where $c_1, \cdots, c_k \in \mathbb{R}$. Since

$$\left(\sum_{i=1}^{k} c_i \vec{v}_i\right) \cdot \vec{u} = \sum_{i=1}^{k} c_i (\vec{v}_i \cdot \vec{u}) = 0,$$

it follows that every element in $\operatorname{span}(\{\vec{v}_1, \cdots, \vec{v}_k\})$ is perpendicular to $\vec{u}$. This establishes the Claim.

We now complete the proof of Proposition #43. According to the Claim, we need only show

$$\left(\vec{w} - \sum_{i=1}^{k} (\vec{w} \cdot \vec{v}_i) \vec{v}_i\right) \cdot \vec{v}_j = 0$$

for $j = 1, \cdots, k$. This follows from

$$\left(\vec{w} - \sum_{i=1}^{k} (\vec{w} \cdot \vec{v}_i) \vec{v}_i\right) \cdot \vec{v}_j$$

$$= \vec{w} \cdot \vec{v}_j - \sum_{i=1}^{k} (\vec{w} \cdot \vec{v}_i)(\vec{v}_i \cdot \vec{v}_j)$$

$$= \vec{w} \cdot \vec{v}_j - \vec{w} \cdot \vec{v}_j = 0.$$

Proposition #43 reveals some significant geometry. Note that

$$\sum_{i=1}^{k} (\vec{w} \cdot \vec{v}_i)\, \vec{v}_i \in \operatorname{span}(\{\vec{v}_1, \cdots, \vec{v}_k\})$$

and, by Proposition #41, $\operatorname{span}(\{\vec{v}_1, \cdots, \vec{v}_k\})$ is a subspace of $\mathbb{R}^n$ of dimension $k$. Now, if $\vec{w} \in \operatorname{span}(\{\vec{v}_1, \cdots, \vec{v}_k\})$, then Proposition #42 tells us that

$$\vec{w} = \sum_{i=1}^{k} (\vec{w} \cdot \vec{v}_i)\, \vec{v}_i$$

— so

$$\vec{w} - \sum_{i=1}^{k} (\vec{w} \cdot \vec{v}_i)\, \vec{v}_i = \vec{0}.$$

If $\vec{w} \notin \operatorname{span}(\{\vec{v}_1, \cdots, \vec{v}_k\})$, then $\operatorname{span}(\{\vec{v}_1, \cdots, \vec{v}_k, \vec{w}\})$ is a $(k+1)$-dimensional subspace of $\mathbb{R}^n$

and, inside of $\text{span}(\{\vec{v}_1, \cdots, \vec{v}_k, \vec{w}\})$, $\text{span}(\{\vec{v}_1, \cdots, \vec{v}_k\})$ looks roughly like a plane in $\mathbb{R}^3$ – where the plane passes through the origin.

This is because

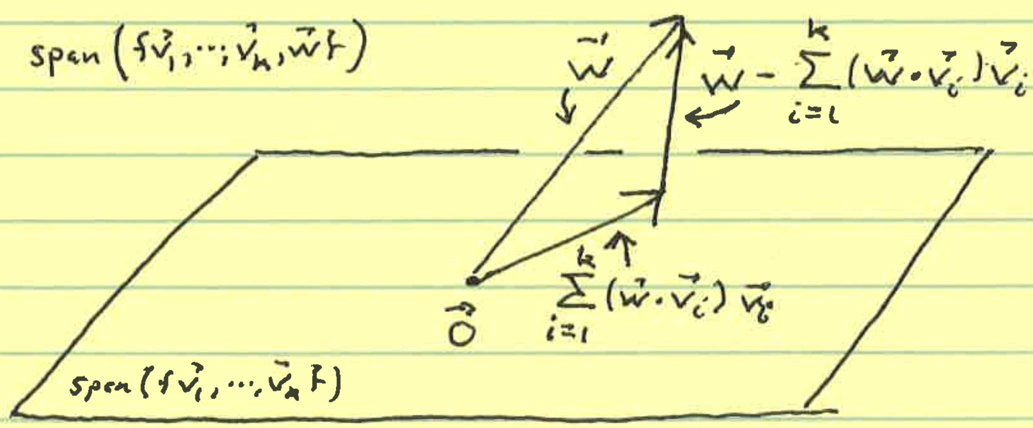
$$\dim(\mathbb{R}^3) - \dim(\text{plane}) = 3 - 2 = 1$$

and

$$\dim\left(\text{span}(\{\vec{v}_1, \cdots, \vec{v}_k, \vec{w}\})\right) - \dim\left(\text{span}(\{\vec{v}_1, \cdots, \vec{v}_k\})\right)$$
$$= k+1-k = 1.$$

Proposition #43 tells us that

$\vec{w} - \sum_{i=1}^{k} (\vec{w} \cdot \vec{v}_i)\vec{v}_i$ is a "normal vector" to $\text{span}(\{\vec{v}_1, \cdots, \vec{v}_k\})$ inside of $\text{span}(\{\vec{v}_1, \cdots, \vec{v}_k, \vec{w}\})$. Thus, if we position ourself inside of span$(\{\vec{v}_1, \cdots, \vec{v}_k, \vec{w}\})$ so that span$(\{\vec{v}_1, \cdots, \vec{v}_k\})$ appears as a horizontal plane – we see a picture as depicted below.




Thus, the distance from the head of $\vec{w}$ to span$(\{\vec{v}_1, \cdots, \vec{v}_k\})$ is $\left| \vec{w} - \sum_{i=1}^{k} (\vec{w} \cdot \vec{v}_i) \vec{v}_i \right|$ and $\sum_{i=1}^{k} (\vec{w} \cdot \vec{v}_i) \vec{v}_i$ is the orthogonal projection of $\vec{w}$ into $\operatorname{span}(\{\vec{v}_1, \cdots, \vec{v}_k\})$.

Prior to stating the next Proposition, we'll prove the following

<u>Claim</u>: Suppose $\{\vec{v}_1, \cdots, \vec{v}_k\}$ is an o.n. set of vectors in $\mathbb{R}^n$ and $\vec{w} \in \mathbb{R}^n$ such that $\vec{w} \notin \operatorname{span}(\{\vec{v}_1, \cdots, \vec{v}_k\})$. Then

$$\vec{w} - \sum_{i=1}^{k} (\vec{w} \cdot \vec{v}_i)\vec{v}_i \neq \vec{0}.$$

<u>Proof of Claim</u>: If $\vec{w} - \sum_{i=1}^{k} (\vec{w} \cdot \vec{v}_i)\vec{v}_i = \vec{0}$, then $\vec{w} = \sum_{i\in 1}^{k} (\vec{w} \cdot \vec{v}_i)\vec{v}_i \in \operatorname{span}(\{\vec{v}_1, \cdots, \vec{v}_k\})$ - which contradicts $\vec{w} \notin \operatorname{span}(\{\vec{v}_1, \cdots, \vec{v}_k\})$.

So $\vec{w} - \sum_{i=1}^{k} (\vec{w} \cdot \vec{v}_i)\vec{v}_i \neq \vec{0}$ and the Claim is established.

<u>Proposition #44</u>: Suppose $\{\vec{v}_1, \cdots, \vec{v}_k\}$ is an o.n. set of vectors in $\mathbb{R}^n$ and $\vec{w} \in \mathbb{R}^n$ such that $\vec{w} \notin \operatorname{span}(\{\vec{v}_1, \cdots, \vec{v}_k\})$.

Set $\vec{z} = \vec{w} - \sum_{i=1}^{k} (\vec{w} \cdot \vec{v}_i)\vec{v}_i$ and, noting $\vec{z} \neq \vec{0}$ by the Claim, set $\vec{\varepsilon} = \frac{1}{|\vec{z}|}\vec{z}$. Then:

(i) $\{\vec{v}_1, \cdots, \vec{v}_k, \vec{\varepsilon}\}$ is an o.n. set of vectors in $\mathbb{R}^n$,

(ii) $\operatorname{span}(\{\vec{v}_1, \cdots, \vec{v}_k, \vec{w}\}) = \operatorname{span}(\{\vec{v}_1, \cdots, \vec{v}_k, \vec{\varepsilon}\})$, and

(iii) $\vec{w} \cdot \vec{\varepsilon} = \left| \vec{w} - \sum_{i=1}^{k} (\vec{w} \cdot \vec{v}_i)\vec{v}_i \right|$.

<u>Proof of parts (ii) and (iii)</u>: We begin with the proof of (ii). Clearly the

vectors $\vec{v}_1, \cdots, \vec{v}_k$ are elements in both $\mathrm{span}(\{\vec{v}_1, \cdots, \vec{v}_k, \vec{w}\})$ and $\mathrm{span}(\{\vec{v}_1, \cdots, \vec{v}_k, \vec{\varepsilon}\})$. Thus (ii) follows by showing $\vec{\varepsilon} \in \mathrm{span}(\{\vec{v}_1, \cdots, \vec{v}_k, \vec{w}\})$ and $\vec{w} \in \mathrm{span}(\{\vec{v}_1, \cdots, \vec{v}_k, \vec{\varepsilon}\})$ – since $\vec{\varepsilon} \in \mathrm{span}(\{\vec{v}_1, \cdots, \vec{v}_k, \vec{w}\})$ implies $\mathrm{span}(\{\vec{v}_1, \cdots, \vec{v}_k, \vec{\varepsilon}\}) \subset \mathrm{span}(\{\vec{v}_1, \cdots, \vec{v}_k, \vec{w}\})$ and $\vec{w} \in \mathrm{span}(\{\vec{v}_1, \cdots, \vec{v}_k, \vec{\varepsilon}\})$ implies $\mathrm{span}(\{\vec{v}_1, \cdots, \vec{v}_k, \vec{w}\}) \subset \mathrm{span}(\{\vec{v}_1, \cdots, \vec{v}_k, \vec{\varepsilon}\})$.

To see that $\vec{\varepsilon} \in \mathrm{span}(\{\vec{v}_1, \cdots, \vec{v}_k, \vec{w}\})$ recall that $\vec{\varepsilon} = \frac{1}{|\vec{z}|}\vec{z}$, where

$$\vec{z} = \vec{w} - \sum_{i=1}^{k} (\vec{w} \cdot \vec{v}_i)\vec{v}_i$$

$$= \sum_{i=1}^{k} (-\vec{w} \cdot v_i)\vec{v}_i + \vec{w} \in \mathrm{span}(\{\vec{v}_1, \cdots, \vec{v}_k, \vec{w}\}).$$

Thus $\vec{z} \in \text{span}(\{\vec{v}_1, \cdots, \vec{v}_k, \vec{w}\})$.

Since $\text{span}(\{\vec{v}_1, \cdots, \vec{v}_k, \vec{w}\})$ is a subspace of $\mathbb{R}^n$, it is closed under scalar multiplication — so

$\vec{\varepsilon} = \frac{1}{|\vec{z}|}\vec{z} \in \text{span}(\{\vec{v}_1, \cdots, \vec{v}_k, \vec{w}\})$.

We now show $\vec{w} \in \text{span}(\{\vec{v}_1, \cdots, \vec{v}_k, \vec{\varepsilon}\})$.

Again recall $\vec{\varepsilon} = \frac{1}{|\vec{z}|}\vec{z}$ — so

$$|\vec{z}|\vec{\varepsilon} = \vec{z} = \vec{w} - \sum_{i=1}^{k} (\vec{w} \cdot \vec{v}_i)\vec{v}_i .$$

It follows that

$$\vec{w} = \sum_{i=1}^{k} (\vec{w} \cdot \vec{v}_i)\vec{v}_i + |\vec{z}|\vec{\varepsilon} .$$

Since $\sum_{i=1}^{k} (\vec{w} \cdot \vec{v}_i)v_i + |\vec{z}|\vec{\varepsilon}$ is an element in $\text{span}(\{\vec{v}_1, \cdots, \vec{v}_k, \vec{\varepsilon}\})$, it follows that $\vec{w} \in \text{span}(\{\vec{v}_1, \cdots, \vec{v}_k, \vec{\varepsilon}\})$.

As for the proof of (iii), we have already shown that

$$\vec{w} = \sum_{i=1}^{k} (\vec{w} \cdot \vec{v}_i)\vec{v}_i + |\vec{z}|\vec{\varepsilon}.$$

Also, since $\vec{w} \in \text{span}(\{\vec{v}_1, \cdots, \vec{v}_k, \vec{\varepsilon}\})$ and $\{\vec{v}_1, \cdots, \vec{v}_k, \vec{\varepsilon}\}$ is an o.n. set of vectors, by (i), it follows that

$$\vec{w} = \sum_{i=1}^{k} (\vec{w} \cdot \vec{v}_i)\vec{v}_i + (\vec{w} \cdot \vec{\varepsilon})\vec{\varepsilon}$$

- by Proposition #42.

We now know that

$$\sum_{i=1}^{k} (\vec{w} \cdot \vec{v}_i)\vec{v}_i + |\vec{z}|\vec{\varepsilon} = \sum_{i=1}^{k} (\vec{w} \cdot \vec{v}_i)\vec{v}_i + (\vec{w} \cdot \vec{\varepsilon})\vec{\varepsilon}$$

so

$$|\vec{z}|\vec{\varepsilon} = (\vec{w} \cdot \vec{\varepsilon})\vec{\varepsilon}.$$

But this implies that

$$\left( (\vec{w} \cdot \vec{\varepsilon}) - |\vec{z}| \right) \vec{\varepsilon} = \vec{0}.$$

Since $\vec{\varepsilon} \neq \vec{0}$, we conclude that

$$(\vec{w} \cdot \vec{\varepsilon}) - |\vec{z}| = 0.$$

So

$$\begin{aligned} \vec{w} \cdot \vec{\varepsilon} &= |\vec{z}| \\ &= \left| \vec{w} - \sum_{i=1}^{k} (\vec{w} \cdot \vec{v}_i) \vec{v}_i \right|. \end{aligned}$$

<u>HW #50</u>: Prove part (i) of Proposition #44.

<u>Theorem #19 (The Gram-Schmidt Process)</u>:

Suppose that $\{\vec{v}_1, \cdots, \vec{v}_m\}$ is a linearly independent set of vectors in $\mathbb{R}^n$. Then there is a set of o.n. vectors in $\mathbb{R}^n$, $\{\vec{\varepsilon}_1, \cdots, \vec{\varepsilon}_m\}$ say, such that

$$\operatorname{span}\left(\{\vec{v}_1, \cdots, \vec{v}_k\}\right) = \operatorname{span}\left(\{\vec{\varepsilon}_1, \cdots, \vec{\varepsilon}_k\}\right)$$

for each $k \in \{1, \cdots, m\}$. Moreover

$$\vec{v}_k \cdot \vec{\varepsilon}_k = \left| \vec{v}_k - \sum_{i=1}^{k-1} (\vec{v}_k \cdot \vec{\varepsilon}_i)\, \vec{\varepsilon}_i \right|$$

for each $k \in \{1, \cdots, m\}$.

Note: Typically, the moreover statement,

$$\vec{v}_k \cdot \vec{\varepsilon}_k = \left| \vec{v}_k - \sum_{i=1}^{k-1} (\vec{v}_k \cdot \vec{\varepsilon}_i)\, \vec{\varepsilon}_i \right| ,$$

is not included as part of the Gram-Schmidt Process. However, as we shall see in the next section, the moreover statement is also significant.

Proof of the Gram-Schmidt Process:

The proof is by induction.

Base Case: We must construct a unit vector, $\vec{\varepsilon}_1$, such that

$$\text{span}(\{\vec{v}_1\}) = \text{span}(\{\vec{\varepsilon}_1\})$$

and

$$\vec{v}_1 \cdot \vec{\varepsilon}_1 = |\vec{v}_1|.$$

To this end, note that $\{\vec{v}_1, \cdots, \vec{v}_m\}$ being linearly independent implies $\vec{v}_1 \neq \vec{0}$. Set $\vec{\varepsilon}_1 = \frac{1}{|\vec{v}_1|}\vec{v}_1$. Then $\vec{\varepsilon}_1$ is certainly a unit vector. Also, since

$$\vec{\varepsilon}_1 = \frac{1}{|\vec{v}_1|}\vec{v}_1 \in \text{span}(\{\vec{v}_1\})$$

it follows that $\text{span}(\{\vec{\varepsilon}_1\}) \subset \text{span}(\{\vec{v}_1\})$.

On the other hand

$$\vec{v}_1 = |\vec{v}_1|\,\vec{\varepsilon}_1 \in \text{span}(\{\vec{\varepsilon}_1\}),$$

so $\text{span}(\{\vec{v}_1\}) \subset \text{span}(\{\vec{\varepsilon}_1\})$.

It follows that $\text{span}(\{\vec{v}_1\}) = \text{span}(\{\vec{\varepsilon}_1\})$.

Finally,

$$\vec{v}_1 \cdot \vec{\varepsilon}_1 = \vec{v}_1 \cdot \left(\frac{1}{|\vec{v}_1|}\vec{v}_1\right) = \frac{1}{|\vec{v}_1|}(\vec{v}_1 \cdot \vec{v}_1)$$

$$= \frac{1}{|\vec{v}_1|}|\vec{v}_1|^2 = |\vec{v}_1| .$$

This completes the Base Case.

<u>Inductive Step</u>: Let $l \in \{1, \cdots, m-1\}$ and suppose we have constructed an o.n. set of vectors, $\{\vec{\varepsilon}_1, \cdots, \vec{\varepsilon}_l\}$, such that

$$\text{span}(\{\vec{v}_1, \cdots, \vec{v}_k\}) = \text{span}(\{\vec{\varepsilon}_1, \cdots, \vec{\varepsilon}_k\})$$

for each $k \in \{1, \cdots, l\}$ and

$$\vec{v}_k \cdot \vec{\varepsilon}_k = \left| \vec{v}_k - \sum_{i=1}^{k-1} (\vec{v}_k \cdot \vec{\varepsilon}_i)\, \vec{\varepsilon}_i \right|$$

for each $k \in \{1, \cdots, \ell\}$.

We must extend the o.n. set $\{\vec{\varepsilon}_1, \cdots, \vec{\varepsilon}_\ell\}$ to an o.n. set $\{\vec{\varepsilon}_1, \cdots, \vec{\varepsilon}_\ell, \vec{\varepsilon}_{\ell+1}\}$ such that

$$\operatorname{span}(\{\vec{v}_1, \cdots, \vec{v}_{\ell+1}\}) = \operatorname{span}(\{\vec{\varepsilon}_1, \cdots, \vec{\varepsilon}_{\ell+1}\})$$

and

$$\vec{v}_{\ell+1} \cdot \vec{\varepsilon}_{\ell+1} = \left| \vec{v}_{\ell+1} - \sum_{i=1}^{\ell} (\vec{v}_{\ell+1} \cdot \vec{\varepsilon}_i)\, \vec{\varepsilon}_i \right|.$$

To get things started, we'll show that $\vec{v}_{\ell+1} \notin \operatorname{span}(\{\vec{\varepsilon}_1, \cdots, \vec{\varepsilon}_\ell\})$. However, if $\vec{v}_{\ell+1} \in \operatorname{span}(\{\vec{\varepsilon}_1, \cdots, \vec{\varepsilon}_\ell\})$, then the inductive hypothesis, $\operatorname{span}(\{\vec{v}_1, \cdots, \vec{v}_\ell\}) = \operatorname{span}(\{\vec{\varepsilon}_1, \cdots, \vec{\varepsilon}_\ell\})$, implies that $\vec{v}_{\ell+1} \in \operatorname{span}(\{\vec{v}_1, \cdots, \vec{v}_\ell\})$.

This implies that $\{\vec{v}_1, \cdots, \vec{v}_{\ell+1}\}$ is linearly dependent - so $\{\vec{v}_1, \cdots, \vec{v}_m\}$ must also be linearly dependent - a contradiction. It follows that $\vec{v}_{\ell+1} \notin \text{span}(\{\vec{\varepsilon}_1, \cdots, \vec{\varepsilon}_\ell\})$.

Now, according to Proposition #44, if we set $\vec{z} = \vec{v}_{\ell+1} - \sum_{i=1}^{\ell} (\vec{v}_{\ell+1} \cdot \vec{\varepsilon}_i)\vec{\varepsilon}_i$ and then set $\vec{\varepsilon}_{\ell+1} = \frac{1}{|\vec{z}|}\vec{z}$, then

(i) $\{\vec{\varepsilon}_1, \cdots, \vec{\varepsilon}_\ell, \vec{\varepsilon}_{\ell+1}\}$ is an o.n. set,

(ii) $\text{span}(\{\vec{\varepsilon}_1, \cdots, \vec{\varepsilon}_\ell, \vec{v}_{\ell+1}\}) = \text{span}(\{\vec{\varepsilon}_1, \cdots, \vec{\varepsilon}_\ell, \vec{\varepsilon}_{\ell+1}\})$,

and

(iii) $\vec{v}_{\ell+1} \cdot \vec{\varepsilon}_{\ell+1} = \left|\vec{v}_{\ell+1} - \sum_{i=1}^{\ell} (\vec{v}_{\ell+1} \cdot \vec{\varepsilon}_i)\vec{\varepsilon}_i\right|$.

As such, to complete the proof of the

Inductive Step we need only show

$$\text{span}(\{\vec{v}_1, \cdots, \vec{v}_{\ell+1}\}) = \text{span}(\{\vec{\varepsilon}_1, \cdots, \vec{\varepsilon}_\ell, \vec{v}_{\ell+1}\}).$$

To this end, initially recall that

$$\text{span}(\{\vec{v}_1, \cdots, \vec{v}_\ell\}) = \text{span}(\{\vec{\varepsilon}_1, \cdots, \vec{\varepsilon}_\ell\}).$$

This implies that

$$\vec{\varepsilon}_1, \cdots, \vec{\varepsilon}_\ell \in \text{span}(\{\vec{v}_1, \cdots, \vec{v}_{\ell+1}\}) \text{ and}$$

$$\vec{v}_1, \cdots, \vec{v}_\ell \in \text{span}(\{\vec{\varepsilon}_1, \cdots, \vec{\varepsilon}_\ell, \vec{v}_{\ell+1}\}).$$

Clearly, $\vec{v}_{\ell+1}$ is an element in both

$\text{span}(\{\vec{v}_1, \cdots, \vec{v}_{\ell+1}\})$ and $\text{span}(\{\vec{\varepsilon}_1, \cdots, \vec{\varepsilon}_\ell, \vec{v}_{\ell+1}\})$

it follows that

$$\vec{\varepsilon}_1, \cdots, \vec{\varepsilon}_\ell, \vec{v}_{\ell+1} \in \text{span}(\{\vec{v}_1, \cdots, \vec{v}_{\ell+1}\}) \text{ and}$$

$$\vec{v}_1, \cdots, \vec{v}_\ell, \vec{v}_{\ell+1} \in \text{span}(\{\vec{\varepsilon}_1, \cdots, \vec{\varepsilon}_\ell, \vec{v}_{\ell+1}\}).$$

As such, we must have

$$\mathrm{span}(\{\vec{\varepsilon}_1, \cdots, \vec{\varepsilon}_\ell, \vec{v}_{\ell+1}\}) \subset \mathrm{span}(\{\vec{v}_1, \cdots, \vec{v}_{\ell+1}\})$$

and

$$\mathrm{span}(\{\vec{v}_1, \cdots, \vec{v}_{\ell+1}\}) \subset \mathrm{span}(\{\vec{\varepsilon}_1, \cdots, \vec{\varepsilon}_\ell, \vec{v}_{\ell+1}\}).$$

It follows that

$$\mathrm{span}(\{\vec{v}_1, \cdots, \vec{v}_{\ell+1}\}) = \mathrm{span}(\{\vec{\varepsilon}_1, \cdots, \vec{\varepsilon}_\ell, \vec{v}_{\ell+1}\}).$$

This completes the proof of the Inductive Step — and the Theorem is established.

I'll write G-S for the Gram-Schmidt Process. The G-S has a number of applications. As a warm-up we'll prove

<u>Proposition #45</u>: Suppose $W$ is a subspace of $\mathbb{R}^n$. Then $W$ can be realized

as the solution space of a system of homogeneous, linear equations in $n$ variables.

<u>Proof</u>: If $W = \mathbb{R}^n$, then $W$ is the solution space of the equation

$$0x_1 + \cdots + 0x_n = 0.$$

If $W = \{\vec{0}\}$, then $W$ is the solution space of the system

$$\begin{array}{cccccc} x_1 & & & & = & 0 \\ & x_2 & & & = & 0 \\ & & \ddots & & & \vdots \\ & & & x_n & = & 0 \end{array}.$$

Suppose now that $\dim(W) \in \{1, \cdots, n-1\}$.

Recall that the homogeneous linear equation, $a_1x_1 + \cdots + a_nx_n = 0$, can be written as $[a_1, \cdots, a_n] \cdot [\vec{x}_1, \cdots, \vec{x}_n] = 0$. It follows that, if we set $\vec{x} = [x_1, \cdots, x_n]$, then

$$\vec{w}_1 \cdot \vec{x} = 0$$
$$\vec{w}_2 \cdot \vec{x} = 0$$
$$\vdots$$
$$\vec{w}_m \cdot \vec{x} = 0,$$

where $\vec{w}_1, \cdots, \vec{w}_m \in \mathbb{R}^n$ is a system of $m$ homogeneous, linear equations in $n$ variables.

We now appeal to the G-S as follows: Begin with a basis for $W$, $\{\vec{v}_1, \cdots, \vec{v}_k\}$ say.

Then extend $\{\vec{v}_1, \cdots, \vec{v}_k\}$ to a basis for $\mathbb{R}^n$, $\{\vec{v}_1, \cdots, \vec{v}_n\}$ say. Apply the G-S to $\{\vec{v}_1, \cdots, \vec{v}_n\}$. This yields an o.n. basis, $\{\vec{\varepsilon}_1, \cdots, \vec{\varepsilon}_n\}$ say, for $\mathbb{R}^n$ such that

$$\text{span}(\{\vec{\varepsilon}_1, \cdots, \vec{\varepsilon}_k\}) = \text{span}(\{\vec{v}_1, \cdots, \vec{v}_k\}) = W.$$

<u>Claim</u>: $W$ is the solution space to the system of homogeneous equations:

$$\mathcal{S}: \quad \begin{aligned} \vec{\varepsilon}_{k+1} \cdot \vec{x} &= 0 \\ &\vdots \\ \vec{\varepsilon}_n \cdot \vec{x} &= 0 \end{aligned}$$

where $\vec{x} = [x_1, \cdots, x_n]$.

To prove the Claim, first note that

since $\{\vec{\varepsilon}_1, \cdots, \vec{\varepsilon}_n\}$ is an o.n. set of vectors and $j > k$, then

$$\vec{\varepsilon}_j \cdot \vec{\varepsilon}_i = 0$$

for $i = 1, \cdots, k$. This implies that $\vec{\varepsilon}_j$ is orthogonal to every vector in $\text{span}(\{\vec{\varepsilon}_1, \cdots, \vec{\varepsilon}_k\}) = W$. Thus, if $\vec{x} \in W$, then

$$\begin{gathered} \vec{\varepsilon}_{k+1} \cdot \vec{x} = 0 \\ \vdots \\ \vec{\varepsilon}_n \cdot \vec{x} = 0. \end{gathered}$$

Thus, $W$ is contained in the solution set of $\mathcal{S}$.

Suppose now that $\vec{x}$ is in the solution space for $\mathcal{S}$. Then $\vec{\varepsilon}_j \cdot \vec{x} = 0$

for $j = k+1, \cdots, n$. Since $\vec{x} \in \mathbb{R}^n$, it follows that

$$\vec{x} = \sum_{i=1}^{n} (\vec{\varepsilon}_i \cdot \vec{x})\, \vec{\varepsilon}_i ,$$

by Proposition #42. Since $\vec{\varepsilon}_i \cdot \vec{x} = 0$ if $i > k$, we have

$$\vec{x} = \sum_{i=1}^{k} (\vec{\varepsilon}_i \cdot \vec{x})\, \vec{\varepsilon}_i \in \operatorname{span}(\{\vec{\varepsilon}_1, \cdots, \vec{\varepsilon}_k\}) = W.$$

This implies that the solution space of $\mathcal{S}$ is contained in $W$.

Consequently, $W$ equals the solution space of $\mathcal{S}$ — and the proof is complete.

<u>Ex</u>: Let $\vec{v}_1 = [1,1]$ and $\vec{v}_2 = [1,2]$. Since $\vec{v}_2$ is not a scalar multiple of $\vec{v}_1$, $\{\vec{v}_1, \vec{v}_2\}$ is a basis for $\mathbb{R}^2$.

Let's apply the G-S to $\{\vec{v}_1, \vec{v}_2\}$.

Initially note that

$$|\vec{v}_1| = |[1,1]| = \sqrt{1^2+1^2} = \sqrt{2}.$$

Thus $\vec{\varepsilon}_1 = \frac{1}{\sqrt{2}}[1,1]$.

Next

$$\vec{z} = \vec{v}_2 - (\vec{v}_2 \cdot \vec{\varepsilon}_1)\,\vec{\varepsilon}_1$$

$$= [1,2] - \left([1,2] \cdot \tfrac{1}{\sqrt{2}}[1,1]\right)\tfrac{1}{\sqrt{2}}[1,1]$$

$$= [1,2] - \tfrac{1}{2}\left([1,2]\cdot[1,1]\right)[1,1]$$

$$= [1,2] - \tfrac{1}{2}(3)[1,1]$$

$$= [1,2] - [3/2, 3/2]$$

$$= [-1/2, 1/2]$$

$$= \tfrac{1}{2}[-1,1].$$

So $|\vec{z}| = \frac{1}{2}|[-1,1]| = \frac{1}{2}\sqrt{1^2+1^2} = \frac{\sqrt{2}}{2} = \frac{1}{\sqrt{2}}$.

Thus $\vec{\varepsilon}_2 = \frac{1}{|\vec{z}|}\vec{z}$

$$= \frac{1}{1/\sqrt{2}} \frac{1}{2}[-1,1]$$

$$= \frac{\sqrt{2}}{2}[-1,1]$$

$$= \frac{1}{\sqrt{2}}[-1,1].$$

Thus, the o.n. basis for $\mathbb{R}^2$ obtained by applying the G-S to

$\{\vec{v}_1 = [1,1], \vec{v}_2 = [1,2]\}$

is

$\{\vec{\varepsilon}_1 = \frac{1}{\sqrt{2}}[1,1], \vec{\varepsilon}_2 = \frac{1}{\sqrt{2}}[-1,1]\}.$