The Gram-Schmidt Procedure

We are almost ready to start generating orthonormal bases for a subspace, but we need to make one last observation first.

Theorem 7.2.3

Suppose that $\vec{v}_1, \ldots, \vec{v}_k \in \mathbb{R}^n$. Then

$$\text{Span}\{\vec{v}_1, \ldots, \vec{v}_k\} = \text{Span}\{\vec{v}_1, \ldots, \vec{v}_{k-1}, \vec{v}_k + t_1\vec{v}_1 + \cdots + t_{k-1}\vec{v}_{k-1}\}$$

for any $t_1, \ldots, t_{k-1} \in \mathbb{R}$.

Proof

The secret to this theorem is to realize the following:

$$\begin{aligned} &a_1\vec{v}_1 + \cdots + a_{k-1}\vec{v}_{k-1} + a_k(\vec{v}_k + t_1\vec{v}_1 + \cdots + t_{k-1}\vec{v}_{k-1}) \\ &= (a_1 + t_1)\vec{v}_1 + \cdots + (a_{k-1} + t_{k-1})\vec{v}_{k-1} + a_k\vec{v}_k \end{aligned}$$

On the left, we have written a general vector of $\text{Span}\{\vec{v}_1, \ldots, \vec{v}_{k-1}, \vec{v}_k + t_1\vec{v}_1 + \cdots + t_{k-1}\vec{v}_{k-1}\}$, and have shown that it is in $\text{Span}\{\vec{v}_1, \ldots, \vec{v}_k\}$.

But if we set $b_i = a_i + t_i$ for $i = 1, \ldots, k-1$, and $b_k = a_k$, then we see that the right side of our equality is a general vector of $\text{Span}\{\vec{v}_1, \ldots, \vec{v}_k\}$, and our equality shows that it is in $\text{Span}\{\vec{v}_1, \ldots, \vec{v}_{k-1}, \vec{v}_k + t_1\vec{v}_1 + \cdots + t_{k-1}\vec{v}_{k-1}\}$.

At long last, we come to the Gram-Schmidt Procedure for using any spanning set to find an orthogonal basis for a subspace.

The Gram-Schmidt Procedure

The Gram-Schmidt Procedure

Set-up: Let $\{\vec{w}_1, \ldots, \vec{w}_k\}$ be a spanning set for a subspace $\mathbb{S}$ of $\mathbb{R}^n$. Yes, that's right. All we need is a spanning set. We don't need it to be linearly independent or have the correct number of vectors. If our set contains any dependent members, our procedure will produce the zero vector, which we will promptly toss aside and continue.

First step: Let $\vec{v}_1 = \vec{w}_1$. Then $\text{Span}\{\vec{v}_1\}$ is obviously equal to $\text{Span}\{\vec{w}_1\}$. We set $\mathbb{S}_1 = \text{Span}\{\vec{v}_1\}$.

Second step: Let $\vec{v}_2 = \text{perp}_{\mathbb{S}_1}\vec{w}_2$. Then we have

$$\vec{v}_2 = \vec{w}_2 - \text{proj}_{\mathbb{S}_1}\vec{w}_2 = \vec{w}_2 - \frac{\vec{w}_2 \cdot \vec{v}_1}{\|\vec{v}_1\|^2}\vec{v}_1$$

If $\vec{v}_2 = \vec{0}$, then $\vec{w}_2 = \text{proj}_{\mathbb{S}_1}\vec{w}_2$, which means that $\vec{w}_2 \in \mathbb{S}_1$.
That is, $\vec{w}_2 \in \text{Span}\{\vec{w}_1\}$.
Which means that $\text{Span}\{\vec{w}_1\} = \text{Span}\{\vec{w}_1, \vec{w}_2\}$, so we have $\text{Span}\{\vec{v}_1\} = \text{Span}\{\vec{w}_1, \vec{w}_2\}$.
This means that $\{\vec{v}_1\}$ is an orthogonal set such that $\text{Span}\{\vec{v}_1\} = \text{Span}\{\vec{w}_1, \vec{w}_2\}$, and we will set $\mathbb{S}_2 = \text{Span}\{\vec{v}_1\}$.

The Gram-Schmidt Procedure

If $\vec{v}_2 \neq \vec{0}$, then we note that since $\text{Span}\{\vec{v}_1\} = \text{Span}\{\vec{w}_1\}$, we know that $\text{Span}\{\vec{v}_1, \vec{w}_2\} = \text{Span}\{\vec{w}_1, \vec{w}_2\}$.

And, from Theorem 7.2.3, we get that $\text{Span}\{\vec{v}_1, \vec{w}_2\} = \text{Span}\{\vec{v}_1, \vec{w}_2 + t\vec{v}_1\}$.

Setting $t = -\dfrac{\vec{w}_2 \cdot \vec{v}_1}{\|\vec{v}_1\|^2}$, we have that $\{\vec{v}_1, \vec{w}_2 + t\vec{v}_1\} = \{\vec{v}_1, \vec{v}_2\}$.

So $\{\vec{v}_1, \vec{v}_2\}$ is an orthogonal set such that $\text{Span}\{\vec{v}_1, \vec{v}_2\} = \text{Span}\{\vec{w}_1, \vec{w}_2\}$, and we set $\mathbb{S}_2 = \text{Span}\{\vec{v}_1, \vec{v}_2\}$.

***i*-th step**: By this point we will have an orthogonal set $\{\vec{v}_1, \dots, \vec{v}_j\}$ (with the precise value of j depending on how many times we got the zero vector) such that $\text{Span}\{\vec{v}_1, \dots, \vec{v}_j\} = \text{Span}\{\vec{w}_1, \dots, \vec{w}_{i-1}\}$, and we will have $\mathbb{S}_{i-1} = \text{Span}\{\vec{v}_1, \dots, \vec{v}_j\}$.

If we set $\vec{v}_{j+1} = \text{perp}_{\mathbb{S}_{i-1}} \vec{w}_i$, then we have

$$\vec{v}_{j+1} = \vec{w}_i - \text{proj}_{\mathbb{S}_{i-1}} \vec{w}_i = \vec{w}_i - \frac{\vec{w}_i \cdot \vec{v}_1}{\|\vec{v}_1\|^2}\vec{v}_1 - \frac{\vec{w}_i \cdot \vec{v}_2}{\|\vec{v}_2\|^2}\vec{v}_2 - \cdots - \frac{\vec{w}_i \cdot \vec{v}_j}{\|\vec{v}_j\|^2}\vec{v}_j$$

As before, if $\vec{v}_{j+1} = \vec{0}$, then we know that $\vec{w}_i = \text{proj}_{\mathbb{S}_{i-1}} \vec{w}_i$, so $\vec{w}_i \in \text{Span}\{\vec{w}_1, \dots, \vec{w}_{i-1}\}$.
This means that $\text{Span}\{\vec{w}_1, \dots, \vec{w}_{i-1}\} = \text{Span}\{\vec{w}_1, \dots, \vec{w}_i\}$, and this means that $\{\vec{v}_1, \dots, \vec{v}_j\}$ is an orthogonal set such that $\text{Span}\{\vec{v}_1, \dots, \vec{v}_j\} = \text{Span}\{\vec{w}_1, \dots, \vec{w}_i\}$. So we set $\mathbb{S}_i = \text{Span}\{\vec{v}_1, \dots, \vec{v}_j\}$.

The Gram-Schmidt Procedure

But if $\vec{v}_{j+1} \neq 0$, then we notice that since $\text{Span}\{\vec{v}_1, \dots, \vec{v}_j\} = \text{Span}\{\vec{w}_1, \dots, \vec{w}_{i-1}\}$, we have $\text{Span}\{\vec{v}_1, \dots, \vec{v}_j, \vec{w}_i\} = \text{Span}\{\vec{w}_1, \dots, \vec{w}_{i-1}, \vec{w}_i\}$.

And then Theorem 7.2.3 tells us that $\text{Span}\{\vec{v}_1, \dots, \vec{v}_j, \vec{w}_i\} = \text{Span}\{\vec{v}_1, \dots, \vec{v}_j, \vec{w}_i + t_1\vec{v}_1 + \cdots + t_j\vec{v}_j\}$.

Setting $t_l = -\dfrac{(\vec{w}_i \cdot \vec{v}_l)}{\|\vec{v}_l\|^2}$ for $l = 1, \dots, j$, we have $\{\vec{v}_1, \dots, \vec{v}_j, \vec{w}_i + t_1\vec{v}_1 + \cdots + t_j\vec{v}_j\} = \{\vec{v}_1, \dots, \vec{v}_j, \vec{v}_{j+1}\}$.

So we've seen that $\{v_1, \dots, \vec{v}_{j+1}\}$ is an orthogonal set such that $\text{Span}\{v_1, \dots, \vec{v}_{j+1}\} = \text{Span}\{\vec{w}_1, \dots, \vec{w}_i\}$. So we set $\mathbb{S}_i = \text{Span}\{v_1, \dots, \vec{v}_{j+1}\}$.

***k*-th step**: This proceeds the same as the i-th step, but at the end we will have an orthogonal set $\{\vec{v}_1, \dots, \vec{v}_j\}$ such that $\text{Span}\{\vec{v}_1, \dots, \vec{v}_j\} = \text{Span}\{\vec{w}_1, \dots, \vec{w}_k\} = \mathbb{S}$.

Since we made sure that $\{\vec{v}_1, \dots, \vec{v}_j\}$ does not contain any zero vectors, we know that $\{\vec{v}_1, \dots, \vec{v}_j\}$ is also linearly independent (by Theorem 7.1.1), which means that $\{\vec{v}_1, \dots, \vec{v}_j\}$ is an orthogonal basis for $\mathbb{S}$.

Of course, this will make more sense once we work through an example.

The Gram-Schmidt Procedure

Example

Find an orthogonal basis for $\text{Span}\left\{\begin{bmatrix}1\\-2\\-1\\-3\end{bmatrix},\begin{bmatrix}2\\-4\\-2\\-6\end{bmatrix},\begin{bmatrix}3\\-5\\0\\-4\end{bmatrix},\begin{bmatrix}-1\\4\\2\\3\end{bmatrix},\begin{bmatrix}-6\\-1\\-8\\3\end{bmatrix}\right\}$.

We start by setting $\vec{v}_1 = \vec{w}_1 = \begin{bmatrix}1\\-2\\-1\\-3\end{bmatrix}$ and $\mathbb{S}_1 = \text{Span}\left\{\begin{bmatrix}1\\-2\\-1\\-3\end{bmatrix}\right\}$.

Next, we set $\vec{v}_2 = \text{perp}_{\mathbb{S}_1}\vec{w}_2 = \vec{w}_2 - \text{proj}_{\mathbb{S}_1}\vec{w}_2$.

The Gram-Schmidt Procedure

Doing the calculations, we get

$$\begin{aligned}\vec{v}_2 &= \begin{bmatrix}2\\-4\\-2\\-6\end{bmatrix} - \frac{\vec{w}_2\cdot\vec{v}_1}{\|\vec{v}_1\|^2}\begin{bmatrix}1\\-2\\-1\\-3\end{bmatrix}\\ &= \begin{bmatrix}2\\-4\\-2\\-6\end{bmatrix} - \frac{2+8+2+18}{1^2+(-2)^2+(-1)^2+(-3)^2}\begin{bmatrix}1\\-2\\-1\\-3\end{bmatrix}\\ &= \begin{bmatrix}2\\-4\\-2\\-6\end{bmatrix} - \frac{30}{15}\begin{bmatrix}1\\-2\\-1\\-3\end{bmatrix}\\ &= \begin{bmatrix}0\\0\\0\\0\end{bmatrix}\end{aligned}$$

This means that $\vec{w}_2 = \begin{bmatrix}2\\-4\\-2\\-6\end{bmatrix}$ is in the span of $\mathbb{S}_1$, so we can discard it.

The Gram-Schmidt Procedure

We set $\mathbb{S}_2 = \mathbb{S}_1 = \text{Span}\left\{ \begin{bmatrix} 1 \\ -2 \\ -1 \\ -3 \end{bmatrix} \right\}$, and we now set $\vec{v}_2 = \textbf{perp}_{\mathbb{S}_2} \vec{w}_3 = \vec{w}_3 - \textbf{proj}_{\mathbb{S}_2} \vec{w}_3$. Doing the calculations, we get

$$\begin{aligned} \vec{v}_2 &= \begin{bmatrix} 3 \\ -5 \\ 0 \\ 4 \end{bmatrix} - \frac{\vec{w}_3 \cdot \vec{v}_1}{||\vec{v}_1||^2} \begin{bmatrix} 1 \\ -2 \\ -1 \\ -3 \end{bmatrix} \\ &= \begin{bmatrix} 3 \\ -5 \\ 0 \\ 4 \end{bmatrix} - \frac{3+10+0+12}{1^2+(-2)^2+(-1)^2+(-3)^2} \begin{bmatrix} 1 \\ -2 \\ -1 \\ -3 \end{bmatrix} \\ &= \begin{bmatrix} 3 \\ -5 \\ 0 \\ 4 \end{bmatrix} - \frac{25}{15} \begin{bmatrix} 1 \\ -2 \\ -1 \\ -3 \end{bmatrix} \\ &= \begin{bmatrix} 4/3 \\ -5/3 \\ 5/3 \\ 1 \end{bmatrix} \end{aligned}$$

The Gram-Schmidt Procedure

Instead of simply leaving $\vec{v}_2 = \begin{bmatrix} 4/3 \\ -5/3 \\ 5/3 \\ 1 \end{bmatrix}$, we note that if $\{\vec{v}_1, \vec{v}_2\}$ is an orthogonal set, then so is $\{\vec{v}_1, s\vec{v}_2\}$ for any scalar s.

So, to simplify our calculations, we can multiply our $\vec{v}_2$ by 3, and set $\mathbb{S}_3 = \text{Span}\left\{ \begin{bmatrix} 1 \\ -2 \\ -1 \\ -3 \end{bmatrix}, \begin{bmatrix} 4 \\ -5 \\ 5 \\ 3 \end{bmatrix} \right\}$.

Next, we set $\vec{v}_3 = \text{perp}_{\mathbb{S}_3} \vec{w}_4 = \vec{w}_4 - \text{proj}_{\mathbb{S}_3} \vec{w}_4$.

The Gram-Schmidt Procedure

$$\vec{v}_3 = \begin{bmatrix} -1 \\ 4 \\ 2 \\ 3 \end{bmatrix} - \frac{\vec{w}_4 \cdot \vec{v}_1}{||\vec{v}_1||^2} \begin{bmatrix} 1 \\ -2 \\ -1 \\ -3 \end{bmatrix} - \frac{\vec{w}_4 \cdot \vec{v}_2}{||\vec{v}_2||^2} \begin{bmatrix} 4 \\ -5 \\ 5 \\ 3 \end{bmatrix}$$

$$= \begin{bmatrix} -1 \\ 4 \\ 2 \\ 3 \end{bmatrix} - \frac{-1-8-2-9}{1^2+(-2)^2+(-1)^2+(-3)^2} \begin{bmatrix} 1 \\ -2 \\ -1 \\ -3 \end{bmatrix} - \frac{-4-20+10+9}{4^2+(-5)^2+5^2+3^2} \begin{bmatrix} 4 \\ -5 \\ 5 \\ 3 \end{bmatrix}$$

$$= \begin{bmatrix} -1 \\ 4 \\ 2 \\ 3 \end{bmatrix} + \frac{20}{15} \begin{bmatrix} 1 \\ -2 \\ -1 \\ -3 \end{bmatrix} + \frac{5}{75} \begin{bmatrix} 4 \\ -5 \\ 5 \\ 3 \end{bmatrix}$$

$$= \begin{bmatrix} 3/5 \\ 1 \\ 1 \\ -4/5 \end{bmatrix}$$

$$= \frac{1}{5} \begin{bmatrix} 3 \\ 5 \\ 5 \\ -4 \end{bmatrix}$$

The Gram-Schmidt Procedure

So we set $\vec{v}_3 = \begin{bmatrix} 3 \\ 5 \\ 5 \\ -4 \end{bmatrix}$ (factoring out the $\frac{1}{5}$ to make our calculations easier), and we set

$$\mathbb{S}_4 = \text{Span}\left\{ \begin{bmatrix} 1 \\ -2 \\ -1 \\ -3 \end{bmatrix}, \begin{bmatrix} 4 \\ -5 \\ 5 \\ 3 \end{bmatrix}, \begin{bmatrix} 3 \\ 5 \\ 5 \\ -4 \end{bmatrix} \right\}.$$

Finally, we set $\vec{v}_4 = \text{perp}_{\mathbb{S}_4} \vec{w}_5 = \vec{w}_5 - \text{proj}_{\mathbb{S}_4} \vec{w}_5$.

The Gram-Schmidt Procedure

$$\begin{aligned}
\vec{v}_4 &= \begin{bmatrix} -6 \\ -1 \\ -8 \\ 3 \end{bmatrix} - \frac{\vec{w}_5 \cdot \vec{v}_1}{||\vec{v}_1||^2} \begin{bmatrix} 1 \\ -2 \\ -1 \\ -3 \end{bmatrix} - \frac{\vec{w}_5 \cdot \vec{v}_2}{||\vec{v}_2||^2} \begin{bmatrix} 4 \\ -5 \\ 5 \\ 3 \end{bmatrix} - \frac{\vec{w}_5 \cdot \vec{v}_3}{||\vec{v}_3||^2} \begin{bmatrix} 3 \\ 5 \\ 5 \\ -4 \end{bmatrix} \\
&= \begin{bmatrix} -6 \\ -1 \\ -8 \\ 3 \end{bmatrix} - \frac{-6+2+8-9}{1^2+(-2)^2+(-1)^2+(-3)^2} \begin{bmatrix} 1 \\ -2 \\ -1 \\ -3 \end{bmatrix} - \frac{-24+5-40+9}{4^2+(-5)^2+5^2+3^2} \begin{bmatrix} 4 \\ -5 \\ 5 \\ 3 \end{bmatrix} \\
&\qquad - \frac{-18-5-40-12}{3^2+5^2+5^2+(-4)^2} \begin{bmatrix} 3 \\ 5 \\ 5 \\ -4 \end{bmatrix} \\
&= \begin{bmatrix} -6 \\ -1 \\ -8 \\ 3 \end{bmatrix} + \frac{5}{15} \begin{bmatrix} 1 \\ -2 \\ -1 \\ -3 \end{bmatrix} + \frac{50}{75} \begin{bmatrix} 4 \\ -5 \\ 5 \\ 3 \end{bmatrix} + \frac{75}{75} \begin{bmatrix} 3 \\ 5 \\ 5 \\ -4 \end{bmatrix} \\
&= \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}
\end{aligned}$$

The Gram-Schmidt Procedure

This means that $\vec{w}_5 = \begin{bmatrix} -6 \\ -1 \\ -8 \\ 3 \end{bmatrix}$ is in the span of $\mathbb{S}_4$, so we can discard it and set

$$\mathbb{S}_5 = \mathbb{S}_4 = \text{Span}\left\{ \begin{bmatrix} 1 \\ -2 \\ -1 \\ -3 \end{bmatrix}, \begin{bmatrix} 4 \\ -5 \\ 5 \\ 3 \end{bmatrix}, \begin{bmatrix} 3 \\ 5 \\ 5 \\ -4 \end{bmatrix} \right\}.$$

And since $\vec{w}_5$ was the last vector in our spanning set, our algorithm comes to an end, and we have that

$$\left\{ \begin{bmatrix} 1 \\ -2 \\ -1 \\ -3 \end{bmatrix}, \begin{bmatrix} 4 \\ -5 \\ 5 \\ 3 \end{bmatrix}, \begin{bmatrix} 3 \\ 5 \\ 5 \\ -4 \end{bmatrix} \right\} \text{ is an orthogonal basis for Span} \left\{ \begin{bmatrix} 1 \\ -2 \\ -1 \\ -3 \end{bmatrix}, \begin{bmatrix} 2 \\ -4 \\ -2 \\ -6 \end{bmatrix}, \begin{bmatrix} 3 \\ -5 \\ 0 \\ -4 \end{bmatrix}, \begin{bmatrix} -1 \\ 4 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} -6 \\ -1 \\ -8 \\ 3 \end{bmatrix} \right\}.$$