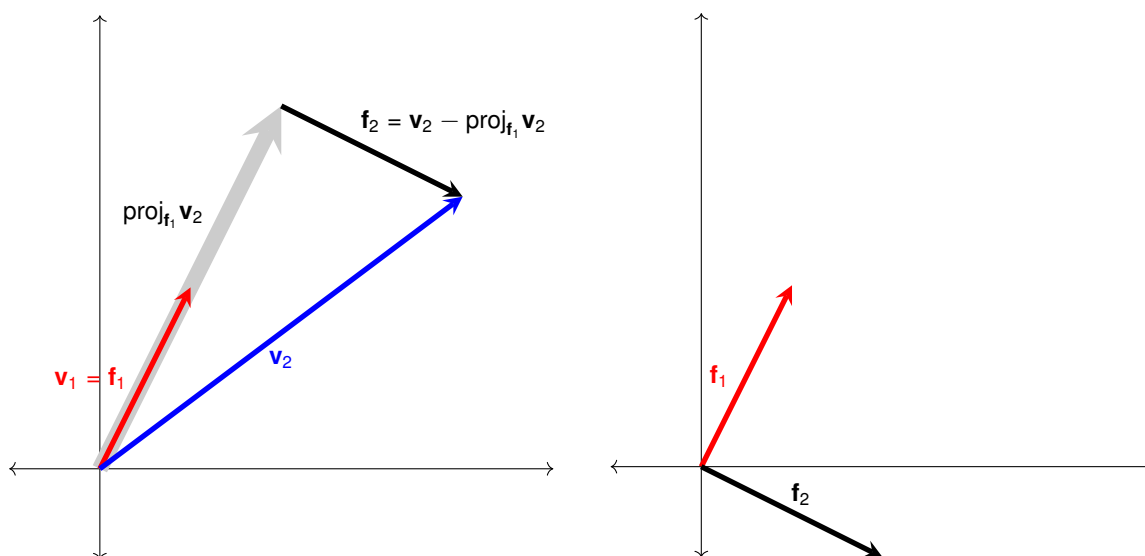

1 Gram-Schmidt Orthogonalization

In [Orthogonality and Projections](#) we said that a set $\{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_m\}$ of nonzero vectors in $\mathbb{R}^n$ is called an *orthogonal set* if $\mathbf{f}_i \cdot \mathbf{f}_j = 0$ for all $i \neq j$. In this section we will prove that every orthogonal set is linearly independent, and therefore it is a basis for its span. We have already seen that the expansion of a vector as a linear combination of orthogonal basis vectors is easy to obtain because formulas exist for the coefficients. Hence the orthogonal bases are the “nice” bases. Our next task is to show that every subspace of $\mathbb{R}^n$ has an orthogonal basis. We will start with intuitive explorations in lower dimensions, then proceed to formalize our results for subspaces of $\mathbb{R}^n$.

A Visual Guide to Creating an Orthogonal Basis

Given an arbitrary basis $\{\mathbf{v}_1, \mathbf{v}_2\}$ of $\mathbb{R}^2$, let's start building our orthogonal basis, $\{\mathbf{f}_1, \mathbf{f}_2\}$, by setting $\mathbf{f}_1 = \mathbf{v}_1$. To find the next element of our orthogonal basis, consider the orthogonal projection of $\mathbf{v}_2$ onto $\mathbf{f}_1$. (See the figure below.)



Next, let $\mathbf{f}_2 = \mathbf{v}_2 - \text{proj}_{\mathbf{f}_1} \mathbf{v}_2$. Observe that $\mathbf{f}_2$ is orthogonal to $\mathbf{f}_1$ (See Theorem ?? of [Orthogonality and Projections](#)). This gives us an orthogonal collection $\mathcal{B} = \{\mathbf{f}_1, \mathbf{f}_2\}$. It is intuitively clear that $\mathbf{f}_1$ and $\mathbf{f}_2$ are linearly independent. Therefore $\mathcal{B}$ is an orthogonal basis of $\mathbb{R}^2$.

The following exploration illustrates this process dynamically.

Exploration 1 Choose an arbitrary basis $\{\mathbf{v}_1, \mathbf{v}_2\}$ of $\mathbb{R}^2$ by dragging the tips of vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ to desired positions. Use the navigation bar at the bottom of the interactive window to go through the steps of constructing an orthogonal basis of $\mathbb{R}^2$.

Access GeoGebra interactives through the online version of this text at

<https://ximera.osu.edu/linearalgebradv3/LinearAlgebraInteractiveIntro>.

We can apply this process to any two-dimensional subset of $\mathbb{R}^n$. The following exploration will guide you through the process of constructing an orthogonal basis for a plane spanned by two arbitrary vectors in $\mathbb{R}^3$.

Exploration 2 Let $W = \text{span}(\mathbf{v}_1, \mathbf{v}_2)$. W is a plane through the origin in $\mathbb{R}^3$. Use the navigation bar at the bottom of the interactive window to go through the steps of constructing an orthogonal basis for W . RIGHT-CLICK and DRAG to rotate the image for a better view.

Access GeoGebra interactives through the online version of this text at

<https://ximera.osu.edu/linearalgebradv3/LinearAlgebraInteractiveIntro>.

In the next exploration, we take the process of constructing an orthogonal basis to the edge of the visual realm and construct an orthogonal basis for $\mathbb{R}^3$.

Exploration 3 In the GeoGebra interactive below $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is a basis of $\mathbb{R}^3$. Use check boxes to go through the steps for constructing an orthogonal basis starting with the given basis. RIGHT-CLICK and DRAG to rotate the image for a better view.

Access GeoGebra interactives through the online version of this text at

<https://ximera.osu.edu/linearalgebradv3/LinearAlgebraInteractiveIntro>.

Gram-Schmidt Orthogonalization Algorithm

In [Orthogonality and Projections](#) we have repeatedly assumed that our subspaces of $\mathbb{R}^n$ have an orthogonal basis. We will now prove that this is indeed the case. Recall that to be a basis of a subspace, a set of vectors must be linearly independent and it must span the subspace. We will start by demonstrating that a set of orthogonal vectors must be linearly independent.

Theorem 1. Let $\{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_k\}$ be an orthogonal set of non-zero vectors in $\mathbb{R}^n$. Then this set is linearly independent.

Proof To show that this set is linearly independent, we need to demonstrate that the only solution to the following equation is the trivial solution.

$$a_1\mathbf{w}_1 + a_2\mathbf{w}_2 + \dots + a_k\mathbf{w}_k = \mathbf{0}$$

To accomplish this, we need to show that all $a_i = 0$ for all $0 \leq i \leq k$. To do so we take the dot product of each side of the above equation with the vector $\mathbf{w}_i$ and obtain the following.

$$\begin{aligned}\mathbf{w}_i \cdot (a_1\mathbf{w}_1 + a_2\mathbf{w}_2 + \dots + a_k\mathbf{w}_k) &= \mathbf{w}_i \cdot \mathbf{0} \\ a_1(\mathbf{w}_i \cdot \mathbf{w}_1) + a_2(\mathbf{w}_i \cdot \mathbf{w}_2) + \dots + a_k(\mathbf{w}_i \cdot \mathbf{w}_k) &= 0\end{aligned}$$

Now since the set is orthogonal, $\mathbf{w}_i \cdot \mathbf{w}_m = 0$ for all $m \neq i$, so we have:

$$\begin{aligned}a_1(0) + \dots + a_i(\mathbf{w}_i \cdot \mathbf{w}_i) + \dots + a_k(0) &= 0 \\ a_i \|\mathbf{w}_i\|^2 &= 0\end{aligned}$$

We know that $\|\mathbf{w}_i\|^2 \neq 0$, so it follows that $a_i = 0$. Since i was chosen arbitrarily, $a_i = 0$ for all i ($0 \leq i \leq k$). This proves that $\{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_k\}$ is linearly independent. ■

The following theorem shows how to start with an arbitrary basis of a subspace W of $\mathbb{R}^n$ and find an orthogonal basis for W . To better understand the notation and the process presented in this theorem, you may want to match the steps of the theorem to the steps of Exploration ??.

Theorem 2 (Gram-Schmidt Orthogonalization). If $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m\}$ is any basis of a subspace W of $\mathbb{R}^n$, consider the following sequence of subspaces:

$$\begin{aligned} W_1 &= \text{span}\{\mathbf{v}_1\} \\ W_2 &= \text{span}\{\mathbf{v}_1, \mathbf{v}_2\} \\ W_3 &= \text{span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\} \\ &\vdots \\ W_m &= \text{span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \dots, \mathbf{v}_m\} \end{aligned}$$

Then we can construct an orthogonal basis $\{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_m\}$ for W_k for each $k = 2, 3, \dots, m$ by adding one vector at a time successively as follows:

$$\begin{aligned} \mathbf{f}_1 &= \mathbf{v}_1 \\ \mathbf{f}_2 &= \mathbf{v}_2 - \text{proj}_{W_1}(\mathbf{v}_2) \\ \mathbf{f}_3 &= \mathbf{v}_3 - \text{proj}_{W_2}(\mathbf{v}_3) \\ &\vdots \\ \mathbf{f}_m &= \mathbf{v}_m - \text{proj}_{W_{m-1}}(\mathbf{v}_m) \end{aligned}$$

Then, $\{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_m\}$ will be an orthogonal basis for W .

Proof Using the definition of projection onto a subspace, the iterative procedure above may be written:

$$\begin{aligned} \mathbf{f}_1 &= \mathbf{v}_1 \\ \mathbf{f}_2 &= \mathbf{v}_2 - \frac{\mathbf{v}_2 \cdot \mathbf{f}_1}{\|\mathbf{f}_1\|^2} \mathbf{f}_1 \\ \mathbf{f}_3 &= \mathbf{v}_3 - \frac{\mathbf{v}_3 \cdot \mathbf{f}_1}{\|\mathbf{f}_1\|^2} \mathbf{f}_1 - \frac{\mathbf{v}_3 \cdot \mathbf{f}_2}{\|\mathbf{f}_2\|^2} \mathbf{f}_2 \\ &\vdots \\ \mathbf{f}_k &= \mathbf{v}_k - \frac{\mathbf{v}_k \cdot \mathbf{f}_1}{\|\mathbf{f}_1\|^2} \mathbf{f}_1 - \frac{\mathbf{v}_k \cdot \mathbf{f}_2}{\|\mathbf{f}_2\|^2} \mathbf{f}_2 - \dots - \frac{\mathbf{v}_k \cdot \mathbf{f}_{k-1}}{\|\mathbf{f}_{k-1}\|^2} \mathbf{f}_{k-1} \\ &\vdots \end{aligned} \tag{1}$$

We see immediately that $\text{span}\{\mathbf{f}_1\} = W_1$ and that $\text{span}\{\mathbf{f}_1, \mathbf{f}_2\} = W_2$ because $\mathbf{f}_2$ is a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$. In fact, for any value of k , we see that $\text{span}\{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_k\} = W_k$, because each $\mathbf{f}_k$ is a linear combination of the vectors $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{k-1}\}$.

Repeated application of Corollary ?? shows that the set $\{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_m\}$ is orthogonal. Linear independence follows from orthogonality by Theorem ??.

We conclude that $\{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_m\}$ is a linearly independent orthogonal set that spans W . ■

Remark 1. Erhardt Schmidt (1876–1959) was a German mathematician who studied under the great David Hilbert. He first described the present algorithm in 1907. J rgen Pederson Gram (1850–1916) was a Danish actuary.

Example 1. Find an orthogonal basis of the row space of $A = \begin{bmatrix} 1 & 1 & -1 & -1 \\ 3 & 2 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix}$.

Explanation. Let $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ denote the rows of A and observe that $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is linearly independent. Take $\mathbf{f}_1 = \mathbf{v}_1$. The algorithm gives

$$\begin{aligned}\mathbf{f}_2 &= \mathbf{v}_2 - \frac{\mathbf{v}_2 \cdot \mathbf{f}_1}{\|\mathbf{f}_1\|^2} \mathbf{f}_1 = [3, 2, 0, 1] - \frac{4}{4}[1, 1, -1, -1] = [2, 1, 1, 2] \\ \mathbf{f}_3 &= \mathbf{v}_3 - \frac{\mathbf{v}_3 \cdot \mathbf{f}_1}{\|\mathbf{f}_1\|^2} \mathbf{f}_1 - \frac{\mathbf{v}_3 \cdot \mathbf{f}_2}{\|\mathbf{f}_2\|^2} \mathbf{f}_2 = \mathbf{v}_3 - \frac{0}{4} \mathbf{f}_1 - \frac{3}{10} \mathbf{f}_2 = \frac{1}{10}[4, -3, 7, -6]\end{aligned}$$

Hence $\{[1, 1, -1, -1], [2, 1, 1, 2], \frac{1}{10}[4, -3, 7, -6]\}$ is the orthogonal basis provided by the algorithm. In hand calculations it may be convenient to eliminate fractions (see the Remark below), so $\{[1, 1, -1, -1], [2, 1, 1, 2], [4, -3, 7, -6]\}$ is also an orthogonal basis for $\text{row}(A)$.

Remark 2. Observe that the vector $\frac{\mathbf{x} \cdot \mathbf{f}_i}{\|\mathbf{f}_i\|^2} \mathbf{f}_i$ is unchanged if a nonzero scalar multiple of $\mathbf{f}_i$ is used in place of $\mathbf{f}_i$. Hence, if a newly constructed $\mathbf{f}_i$ is multiplied by a nonzero scalar at some stage of the Gram-Schmidt algorithm, the subsequent $\mathbf{f}_s$ will be unchanged. This is useful in actual calculations.

The Gram-Schmidt algorithm demonstrates in a constructive way that every subspace of $\mathbb{R}^n$ has an orthogonal basis. We formalize this in one final theorem.

Theorem 3. Let W be a subspace of $\mathbb{R}^n$. Then W has an orthogonal basis. In fact, every orthogonal subset $\{\mathbf{f}_1, \dots, \mathbf{f}_m\}$ in W can be extended to an orthogonal basis for W .

Proof Suppose $\{\mathbf{f}_1, \dots, \mathbf{f}_m\}$ is an orthogonal subset of W . If $\text{span}(\mathbf{f}_1, \dots, \mathbf{f}_m) = W$, it is already a basis. Otherwise, there exists $\mathbf{x}$ in W outside $\text{span}(\mathbf{f}_1, \dots, \mathbf{f}_m)$. Using the Gram-Schmidt procedure we define $\mathbf{f}_{m+1} = \mathbf{x} - \text{proj}_{W_m}(\mathbf{x})$, where $W_m = \text{span}\{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_m\}$. If $\text{span}(\mathbf{f}_1, \dots, \mathbf{f}_m, \mathbf{f}_{m+1}) = W$, we are done. Otherwise, the process continues to create larger and larger orthogonal subsets of W . They are all linearly independent by Theorem ??, so we have a basis when we reach a subset containing $\dim(W)$ vectors. ■

The process described in the proof of this theorem is used in this final example.

Example 2. In Example ??, given $A = \begin{bmatrix} 1 & 1 & -1 & -1 \\ 3 & 2 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix}$, we showed that an orthogonal basis for $\text{row}(A)$ is given by

$$\{\mathbf{f}_1 = [1, 1, -1, -1], \mathbf{f}_2 = [2, 1, 1, 2], \mathbf{f}_3 = [4, -3, 7, -6]\}.$$

Choose any vector $\mathbf{v}_4 \in \mathbb{R}^4$ not in $\text{span}\{\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3\}$, and apply the Gram-Schmidt algorithm to produce a vector $\mathbf{f}_4$ such that $\{\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3, \mathbf{f}_4\}$ is an orthogonal basis for $\mathbb{R}^4$.

Explanation. Let $\mathbf{v}_4 = [1, 0, 0, 0]$. (How would you check that $\mathbf{v}_4$ is not in $\text{span}\{\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3\}$?) To get a vector $\mathbf{f}$

orthogonal to the row space, we perform an iteration of Gram-Schmidt:

$$\begin{aligned}\mathbf{f} &= \mathbf{v}_4 - \frac{\mathbf{v}_4 \cdot \mathbf{f}_1}{\|\mathbf{f}_1\|^2} \mathbf{f}_1 - \frac{\mathbf{v}_4 \cdot \mathbf{f}_2}{\|\mathbf{f}_2\|^2} \mathbf{f}_2 - \frac{\mathbf{v}_4 \cdot \mathbf{f}_3}{\|\mathbf{f}_3\|^2} \mathbf{f}_3 \\ &= [1, 0, 0, 0] - \frac{1}{4}[1, 1, -1, -1] - \frac{2}{10}[2, 1, 1, 2] - \frac{4}{110}[4, -3, 7, -6] \\ &= \frac{1}{44}[9, -15, -9, 3]\end{aligned}$$

Since any multiple of $\mathbf{f}$ will suffice, we are free to choose $\mathbf{f}_4 = 44\mathbf{f} = [9, -15, -9, 3]$ to get rid of the fraction. It is easy to check that $\{\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3, \mathbf{f}_4\}$ is an orthogonal set, and it follows from Theorem ?? that this set is a basis for $\mathbb{R}^4$.

Remark 3. Suppose instead of $[1, 0, 0, 0]$ we had started with $\mathbf{v}_4 = [7, -1, 7, -5]$. This vector $\mathbf{v}_4$ is in $\text{span}\{\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3\}$, as it is the sum of those three vectors. But if we were to try to proceed as above, we would get

$$\begin{aligned}\mathbf{f} &= \mathbf{v}_4 - \frac{\mathbf{v}_4 \cdot \mathbf{f}_1}{\|\mathbf{f}_1\|^2} \mathbf{f}_1 - \frac{\mathbf{v}_4 \cdot \mathbf{f}_2}{\|\mathbf{f}_2\|^2} \mathbf{f}_2 - \frac{\mathbf{v}_4 \cdot \mathbf{f}_3}{\|\mathbf{f}_3\|^2} \mathbf{f}_3 \\ &= [7, -1, 7, -5] - \frac{4}{4}[1, 1, -1, -1] - \frac{10}{10}[2, 1, 1, 2] - \frac{110}{110}[4, -3, 7, -6] \\ &= [0, 0, 0, 0]\end{aligned}$$

We could not add a multiple of $\mathbf{f}$ to $\{\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3\}$ to get an orthogonal basis for $\mathbb{R}^4$.

Practice Problems

Problem 4 Try Example ?? again starting with some other vector $\mathbf{v}_4 \in \mathbb{R}^4$.

Problems ??-??

In each case, use the Gram-Schmidt algorithm to convert the given basis $\mathcal{B}$ of V to an orthogonal basis.

Problem 5 $V = \mathbb{R}^2, \mathcal{B} = \left\{ \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix} \right\}$

Problem 6 $V = \mathbb{R}^2, \mathcal{B} = \left\{ \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\}$

Problem 7 $V = \mathbb{R}^3, \mathcal{B} = \left\{ \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} \right\}$

Problem 8 $V = \mathbb{R}^3, \mathcal{B} = \left\{ \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix} \right\}$

Text Source

This section was adapted from the first part of Section 8.1 of Keith Nicholson's *Linear Algebra with Applications*.
(CC-BY-NC-SA)

W. Keith Nicholson, *Linear Algebra with Applications*, Lyryx 2018, Open Edition, p. 415