Example 61. (warmup) Determine the orthogonal projection of $\begin{bmatrix} 3 \\ 7 \\ 4 \end{bmatrix}$ onto $\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$.

Solution. The orthogonal projection of $\begin{bmatrix} 3 \\ 7 \\ 4 \end{bmatrix}$ onto $\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$ is $\frac{\begin{bmatrix} 3 \\ 7 \\ 4 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}}{\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} = -2 \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$.

Let $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ be a basis of V. That means that we can write any vector $\boldsymbol{w}$ in V as a linear combination of the basis vectors, i.e. $\boldsymbol{w} = c_1 \boldsymbol{v}_1 + \ldots + c_n \boldsymbol{v}_n$. Moreover, the values c_i are unique.

We can find the values c_i simply by solving the system $\boldsymbol{w} = c_1 \boldsymbol{v}_1 + \ldots + c_n \boldsymbol{v}_n$.

If, in fact, $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ is an orthogonal basis of V, then there is a way to determine the values c_i without solving any system! Can you see how?

Solution. Take the dot product of $\boldsymbol{v}_1$ with both sides of $\boldsymbol{w} = c_1 \boldsymbol{v}_1 + \ldots + c_n \boldsymbol{v}_n$:

$$\begin{aligned} \boldsymbol{v}_1 \cdot \boldsymbol{w} &= \boldsymbol{v}_1 \cdot (c_1 \boldsymbol{v}_1 + \ldots + c_n \boldsymbol{v}_n) \\ &= c_1 \boldsymbol{v}_1 \cdot \boldsymbol{v}_1 + c_2 \boldsymbol{v}_1 \cdot \boldsymbol{v}_2 + \ldots + c_n \boldsymbol{v}_1 \cdot \boldsymbol{v}_n \\ &= c_1 \boldsymbol{v}_1 \cdot \boldsymbol{v}_1 \end{aligned}$$

Hence, $c_1 = \frac{\boldsymbol{v}_1 \cdot \boldsymbol{w}}{\boldsymbol{v}_1 \cdot \boldsymbol{v}_1}$. In general, $c_j = \frac{\boldsymbol{v}_j \cdot \boldsymbol{w}}{\boldsymbol{v}_j \cdot \boldsymbol{v}_j}$.

Important observation. $c_1 \boldsymbol{v}_1$ is the orthogonal projection of $\boldsymbol{w}$ onto $\boldsymbol{v}_1$.

In conclusion, we have found the following:

(express in terms of an orthogonal basis)

If $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ is an orthogonal basis of V, and $\boldsymbol{w}$ is in V, then

$$\boldsymbol{w} = \underbrace{c_1 \boldsymbol{v}_1}_{\substack{\text{proj of } \boldsymbol{w} \\ \text{onto } \boldsymbol{v}_1}} + \ldots + \underbrace{c_n \boldsymbol{v}_n}_{\substack{\text{proj of } \boldsymbol{w} \\ \text{onto } \boldsymbol{v}_n}} \qquad \text{with} \quad c_j = \frac{\boldsymbol{w} \cdot \boldsymbol{v}_j}{\boldsymbol{v}_j \cdot \boldsymbol{v}_j}.$$

In other words, $\boldsymbol{w}$ decomposes as the sum of its projections onto each basis vector.

Note. If $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ is an orthonormal basis, then this simplifies to $c_j = \boldsymbol{w} \cdot \boldsymbol{v}_j$.

If $\boldsymbol{w}$ is not in V, then the above formula produces the orthogonal projection of $\boldsymbol{w}$ onto V:

(orthogonal projection if we have an orthogonal basis)

If $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ is an orthogonal basis of V, then the orthogonal projection of $\boldsymbol{w}$ onto V is

$$\left(\frac{\boldsymbol{w} \cdot \boldsymbol{v}_1}{\boldsymbol{v}_1 \cdot \boldsymbol{v}_1} \right) \boldsymbol{v}_1 + \ldots + \left(\frac{\boldsymbol{w} \cdot \boldsymbol{v}_n}{\boldsymbol{v}_n \cdot \boldsymbol{v}_n} \right) \boldsymbol{v}_n.$$

Why? Think about adding extra vectors $\boldsymbol{v}_{n+1}, \boldsymbol{v}_{n+2}, \ldots$ to our orthogonal basis until we have an orthogonal basis for the ambient space (where $\boldsymbol{w}$ lives). Then we can express $\boldsymbol{w}$ as in the previous box, and the formula for the orthogonal projection is what we get when dropping terms involving $\boldsymbol{v}_{n+1}, \boldsymbol{v}_{n+2}, \ldots$ But these dropped terms, the "error", is orthogonal to V. Hence, this must be the orthogonal projection.

Armin Straub
straub@southalabama.edu

Example 62. Express $\underbrace{\begin{bmatrix} 3 \\ 7 \\ 4 \end{bmatrix}}_{\boldsymbol{x}}$ in terms of the basis $\underbrace{\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}}_{\boldsymbol{v}_1}, \underbrace{\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}}_{\boldsymbol{v}_2}, \underbrace{\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}}_{\boldsymbol{v}_3}$.

Solution. Because $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3$ is an orthogonal basis of $\mathbb{R}^3$, we get:

$$\begin{aligned}\begin{bmatrix} 3 \\ 7 \\ 4 \end{bmatrix} &= c_1 \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + c_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \\ &= \underbrace{\frac{\begin{bmatrix} 3 \\ 7 \\ 4 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}}{\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}}_{\text{projection of } \boldsymbol{x} \text{ onto } \boldsymbol{v}_1} + \underbrace{\frac{\begin{bmatrix} 3 \\ 7 \\ 4 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}}{\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}}_{\text{projection of } \boldsymbol{x} \text{ onto } \boldsymbol{v}_2} + \underbrace{\frac{\begin{bmatrix} 3 \\ 7 \\ 4 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}}{\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}}_{\text{projection of } \boldsymbol{x} \text{ onto } \boldsymbol{v}_3} \\ &= \frac{-4}{2} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + \frac{10}{2} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + \frac{4}{1} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\end{aligned}$$

Because we spelled out all the details this looks more involved than it is. We only computed 6 dot products!

Note. Of course, this again features $-2\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$, the orthogonal projection of $\begin{bmatrix} 3 \\ 7 \\ 4 \end{bmatrix}$ onto $\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$.

Alternative. We could have solved $\begin{bmatrix} 1 & 1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 3 \\ 7 \\ 4 \end{bmatrix}$ to also find $\begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} -2 \\ 5 \\ 4 \end{bmatrix}$.

The numbers are particularly easy here but in general, to find this solution, we have to go through the entire process of Gaussian elimination. On the other hand, if we have an orthogonal basis, the former approach requires less work, because it is just computing a few dot products.

Example 63. Determine the projection of $\begin{bmatrix} 3 \\ 7 \\ 4 \end{bmatrix}$ onto $W = \operatorname{span}\left\{\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right\}$.

Comment. We know how to do this using least squares.

However, realizing that $\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ are orthogonal allows us to save some time.

Solution. (using orthogonality) As in Example 62, the projection of $\begin{bmatrix} 3 \\ 7 \\ 4 \end{bmatrix}$ onto $\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$ is $-2\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$ and the projection of $\begin{bmatrix} 3 \\ 7 \\ 4 \end{bmatrix}$ onto $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ is $4\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$.

Hence, the orthogonal projection of $\begin{bmatrix} 3 \\ 7 \\ 4 \end{bmatrix}$ onto $W = \operatorname{span}\left\{\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right\}$ is $-2\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + 4\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -2 \\ 2 \\ 4 \end{bmatrix}$.

Important. This "shortcut" only works because we have an orthogonal basis for W!

Armin Straub
straub@southalabama.edu

Example 64. Find an orthogonal basis for $W = \text{span}\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \right\}$.

Solution. We already have the basis $\boldsymbol{w}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, $\boldsymbol{w}_2 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$ for W. However, that basis is not orthogonal.

We can construct an orthogonal basis $\boldsymbol{q}_1, \boldsymbol{q}_2$ for W as follows:

- $\boldsymbol{q}_1 = \boldsymbol{w}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

  Since this is our first basis vector, we don't yet have other basis vectors it needs to be orthogonal to.

- $\boldsymbol{q}_2 = \boldsymbol{w}_2 - \begin{pmatrix} \text{projection of} \\ \boldsymbol{w}_2 \text{ onto } \boldsymbol{q}_1 \end{pmatrix} = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} - \frac{1}{3}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2/3 \\ -4/3 \\ 2/3 \end{bmatrix}$

  Make sure our way to construct $\boldsymbol{q}_2$ makes sense to you!

  $\boldsymbol{q}_2$ is the error of the projection of $\boldsymbol{w}_2$ onto $\boldsymbol{q}_1$. This guarantees that it is orthogonal to $\boldsymbol{q}_1$.

  On the other hand, since $\boldsymbol{q}_2$ is a combination of $\boldsymbol{w}_2$ and $\boldsymbol{q}_1$, we know that $\boldsymbol{q}_2$ actually is in W.

We have thus found the orthogonal basis $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2/3 \\ -4/3 \\ 2/3 \end{bmatrix}$ for W.

Important comment. Normalizing these, we get $\frac{1}{\sqrt{3}}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \frac{1}{\sqrt{6}}\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$, which is an orthonormal basis for W.

Comment. There are, of course, many orthogonal bases $\boldsymbol{q}_1, \boldsymbol{q}_2$ for W. Up to the length of the vectors, ours is the unique one with the property that $\text{span}\{\boldsymbol{q}_1\} = \text{span}\{\boldsymbol{w}_1\}$ and $\text{span}\{\boldsymbol{q}_1, \boldsymbol{q}_2\} = \text{span}\{\boldsymbol{w}_1, \boldsymbol{w}_2\}$.

(Gram–Schmidt orthogonalization)

Given a basis $\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots$ for W, produce an orthogonal basis $\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots$ for W.

- $\boldsymbol{q}_1 = \boldsymbol{w}_1$
- $\boldsymbol{q}_2 = \boldsymbol{w}_2 - \begin{pmatrix} \text{projection of} \\ \boldsymbol{w}_2 \text{ onto } \boldsymbol{q}_1 \end{pmatrix}$
- $\boldsymbol{q}_3 = \boldsymbol{w}_3 - \begin{pmatrix} \text{projection of} \\ \boldsymbol{w}_3 \text{ onto } \boldsymbol{q}_1 \end{pmatrix} - \begin{pmatrix} \text{projection of} \\ \boldsymbol{w}_3 \text{ onto } \boldsymbol{q}_2 \end{pmatrix}$
- $\boldsymbol{q}_4 = \ldots$

Comment. Since $\boldsymbol{q}_1, \boldsymbol{q}_2$ are orthogonal, $\begin{pmatrix} \text{projection of} \\ \boldsymbol{w}_3 \text{ onto span}\{\boldsymbol{q}_1, \boldsymbol{q}_2\} \end{pmatrix} = \begin{pmatrix} \text{projection of} \\ \boldsymbol{w}_3 \text{ onto } \boldsymbol{q}_1 \end{pmatrix} + \begin{pmatrix} \text{projection of} \\ \boldsymbol{w}_3 \text{ onto } \boldsymbol{q}_2 \end{pmatrix}$.

Important comment. When working numerically it actually saves time to compute an orthonormal basis $\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots$ by the same approach but always normalizing each $\boldsymbol{q}_i$ along the way. The reason this saves time is that now the projections onto $\boldsymbol{q}_i$ only require a single dot product (instead of two). This is called **Gram–Schmidt orthonormalization**.

Note. When normalizing, the orthonormal basis $\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots$ is the unique one with the property that $\text{span}\{\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots, \boldsymbol{q}_k\} = \text{span}\{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_k\}$ for all $k = 1, 2, \ldots$.