Review. If $v_1, ..., v_n$ are orthogonal, the orthogonal projection of w onto $\mathrm{span}\{v_1, ..., v_n\}$ is

$$\hat{w} = \frac{w \cdot v_1}{v_1 \cdot v_1} v_1 + ... + \frac{w \cdot v_n}{v_n \cdot v_n} v_n.$$

Example 69.

(a) Project $\begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$ onto $W = \mathrm{span}\left\{ \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix} \right\}$.

(b) Express $\begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$ in terms of the basis $\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ -5 \end{bmatrix}$.

Solution.

(a) We note that the vectors $\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}$ are orthogonal to each other.

Therefore, the projection can be computed as $\dfrac{\begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}}{\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}} \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} + \dfrac{\begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}}{\begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}} \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix} = \dfrac{8}{6} \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} + \dfrac{4}{5} \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}$.

Comment. If we didn't have an orthogonal basis for $W = \mathrm{col}\left(\begin{bmatrix} 1 & 2 \\ 2 & -1 \\ 1 & 0 \end{bmatrix} \right)$, then we would have to solve the least squares problem $\begin{bmatrix} 1 & 2 \\ 2 & -1 \\ 1 & 0 \end{bmatrix} x = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$ instead to get the same final result (with more work).

(b) Note that this basis is orthogonal! Therefore, we can compute $\begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} = \dfrac{8}{6} \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} + \dfrac{4}{5} \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix} + \dfrac{5}{30} \begin{bmatrix} 1 \\ 2 \\ -5 \end{bmatrix}$.

(We proceed exactly as in the previous part to compute each coefficient as a quotient of dot products.)

Gram–Schmidt

(Gram–Schmidt orthogonalization)

Given a basis $w_1, w_2, ...$ for W, we produce an orthogonal basis $q_1, q_2, ...$ for W as follows:

- $q_1 = w_1$

- $q_2 = w_2 - \left(\begin{array}{c} \text{projection of} \\ w_2 \text{ onto } q_1 \end{array} \right)$

- $q_3 = w_3 - \left(\begin{array}{c} \text{projection of} \\ w_3 \text{ onto } q_1 \end{array} \right) - \left(\begin{array}{c} \text{projection of} \\ w_3 \text{ onto } q_2 \end{array} \right)$

- $q_4 = ...$

Note. Since q_1, q_2 are orthogonal, $\left(\begin{array}{c} \text{projection of} \\ w_3 \text{ onto } \mathrm{span}\{q_1, q_2\} \end{array} \right) = \left(\begin{array}{c} \text{projection of} \\ w_3 \text{ onto } q_1 \end{array} \right) + \left(\begin{array}{c} \text{projection of} \\ w_3 \text{ onto } q_2 \end{array} \right)$.

Important comment. When working numerically on a computer it actually saves time to compute an orthonormal basis $q_1, q_2, ...$ by the same approach but always normalizing each q_i along the way. The reason this saves time is that now the projections onto q_i only require a single dot product (instead of two). This is called **Gram–Schmidt orthonormalization**. When working by hand, it is usually simpler to wait until the end to normalize (so as to avoid working with square roots).

Note. When normalizing, the orthonormal basis $q_1, q_2, ...$ is the unique one (up to $\pm$ signs) with the property that $\mathrm{span}\{q_1, q_2, ..., q_k\} = \mathrm{span}\{w_1, w_2, ..., w_k\}$ for all $k = 1, 2, ...$.

Example 70. Using Gram–Schmidt, find an orthogonal basis for $W = \operatorname{span}\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \right\}$.

Solution. We already have the basis $w_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, $w_2 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$ for W. However, that basis is not orthogonal.

We can construct an orthogonal basis q_1, q_2 for W as follows:

- $q_1 = w_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

- $q_2 = w_2 - \left(\begin{matrix} \text{projection of} \\ w_2 \text{ onto } q_1 \end{matrix} \right) = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} - \frac{1}{3} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 2 \\ -4 \\ 2 \end{bmatrix}$

  Note. q_2 is the error of the projection of w_2 onto q_1. This guarantees that it is orthogonal to q_1.
  On the other hand, since q_2 is a combination of w_2 and q_1, we know that q_2 actually is in W.

We have thus found the orthogonal basis $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \frac{2}{3} \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$ for W (if we like, we can, of course, drop that $\frac{2}{3}$).

Important comment. By normalizing, we get an orthonormal basis for W: $\frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$.

Practical comment. When implementing Gram–Schmidt on a computer, it is beneficial (slightly less work) to normalize each q_i during the Gram–Schmidt process. This typically introduces square roots, which is why normalizing at the end is usually preferable when working by hand.

Comment. There are, of course, many orthogonal bases q_1, q_2 for W. Up to the length of the vectors, ours is the unique one with the property that $\operatorname{span}\{q_1\} = \operatorname{span}\{w_1\}$ and $\operatorname{span}\{q_1, q_2\} = \operatorname{span}\{w_1, w_2\}$.

Example 71. Using Gram–Schmidt, find an orthogonal basis for $W = \operatorname{span}\left\{ \begin{bmatrix} 0 \\ 3 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \right\}$.

Solution. We begin with the (not orthogonal) basis $w_1 = \begin{bmatrix} 0 \\ 3 \\ 0 \\ 0 \end{bmatrix}$, $w_2 = \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \end{bmatrix}$, $w_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$.

We then construct an orthogonal basis q_1, q_2, q_3:

- $q_1 = w_1 = \begin{bmatrix} 0 \\ 3 \\ 0 \\ 0 \end{bmatrix}$

- $q_2 = w_2 - \left(\begin{matrix} \text{projection of} \\ w_2 \text{ onto } q_1 \end{matrix} \right) = \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \end{bmatrix} - \frac{3}{9} \begin{bmatrix} 0 \\ 3 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ 0 \\ 0 \end{bmatrix}$

- $q_3 = w_3 - \left(\begin{matrix} \text{projection of } w_3 \\ \text{onto } \operatorname{span}\{q_1, q_2\} \end{matrix} \right) = w_3 - \left(\begin{matrix} \text{projection of} \\ w_3 \text{ onto } q_1 \end{matrix} \right) - \left(\begin{matrix} \text{projection of} \\ w_3 \text{ onto } q_2 \end{matrix} \right)$

  $= \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} - \frac{3}{9} \begin{bmatrix} 0 \\ 3 \\ 0 \\ 0 \end{bmatrix} - \frac{2}{4} \begin{bmatrix} 2 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}$

  Make sure you understand how q_3 was designed to be orthogonal to both q_1 and q_2!

  Also note that breaking up the projection onto $\operatorname{span}\{q_1, q_2\}$ into the projections onto q_1 and q_2 is only possible because q_1 and q_2 are orthogonal.

Hence, $\begin{bmatrix} 0 \\ 3 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}$ is an orthogonal basis of W.

Important. Normalizing, we obtain an orthonormal basis: $\begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}$.