# More on orthogonality

## Projecting onto 1-dimensional spaces

When we project onto a $1$-dimensional space $\operatorname{span}\{\boldsymbol{w}\}$, we usually just say that we are projecting onto $\boldsymbol{w}$.

**Example 50.** What is the orthogonal projection of $\boldsymbol{v} = (1,2,3)^T$ onto $\boldsymbol{w} = (1,1,1)^T$?

**Solution.** To project $\boldsymbol{b} = \boldsymbol{v}$ onto $\operatorname{col}(A)$, with $A = \boldsymbol{w}$, we first find a least squares solution to $A^T A\boldsymbol{x} = A^T\boldsymbol{b}$, which is $\boldsymbol{w}^T\boldsymbol{w}\boldsymbol{x} = \boldsymbol{w}^T\boldsymbol{v}$, that is, $3\boldsymbol{x} = 6$. It follows that $\hat{\boldsymbol{x}} = 2$.

The orthogonal projection is $A\hat{\boldsymbol{x}} = \boldsymbol{w}\hat{\boldsymbol{x}} = 2\boldsymbol{w} = \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix}$.

**Makes sense.** Make sure you realize how much sense this answer makes, even if we didn't know anything about projections. Namely, we are looking for a vector closest to $(1,2,3)^T$ of the form $(x,x,x)^T$. The best we can do is choose the mean of the values $1,2,3$, which is $2$.

**Let's check.** Recall that we can check that this is true by verifying that the error $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} - \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$ is orthogonal to $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$. Indeed, $\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = 0$.

**Example 51.** Find a formula for the projection of a vector $\boldsymbol{v}$ onto another vector $\boldsymbol{w}$.

**Solution.** As in the previous example, we first find a least squares solution to $\boldsymbol{w}^T\boldsymbol{w}\boldsymbol{x} = \boldsymbol{w}^T\boldsymbol{v}$, which results in $\hat{\boldsymbol{x}} = \frac{\boldsymbol{w}^T\boldsymbol{v}}{\boldsymbol{w}^T\boldsymbol{w}}$. (Note that $\boldsymbol{w}^T\boldsymbol{v}$ and $\boldsymbol{w}^T\boldsymbol{w}$ are $1\times 1$ matrices, i.e. just numbers!)

Hence, the projection of $\boldsymbol{v}$ onto $\boldsymbol{w}$ is $\boldsymbol{w}\hat{\boldsymbol{x}} = \boldsymbol{w}\frac{\boldsymbol{w}^T\boldsymbol{v}}{\boldsymbol{w}^T\boldsymbol{w}} = \left(\frac{\boldsymbol{w}\cdot\boldsymbol{v}}{\|\boldsymbol{w}\|^2}\right)\boldsymbol{w}$. In summary:

> The (orthogonal) projection of $\boldsymbol{v}$ onto $\boldsymbol{w}$ is $\dfrac{\boldsymbol{w}\cdot\boldsymbol{v}}{\|\boldsymbol{w}\|^2}\boldsymbol{w}$.

**Comment.** If you have taken Calculus 3, you have seen that formula before. Most likely, you were deriving it using angles at that time. Namely, the dot product has the following connection to angles:

> $\boldsymbol{v}\cdot\boldsymbol{w} = \|\boldsymbol{v}\|\,\|\boldsymbol{w}\|\cos\theta$ where $\theta \in [0,\pi]$ is the angle between $\boldsymbol{v}$ and $\boldsymbol{w}$

**Why?** You can derive this by repeating what we did, right after Definition 20 to show that $\boldsymbol{v}$ and $\boldsymbol{w}$ are orthogonal if and only if $\boldsymbol{v}\cdot\boldsymbol{w} = 0$. Just replace Pythagoras with the law of cosines ($c^2 = a^2 + b^2 - 2ab\cos\theta$ holds in any triangle!).

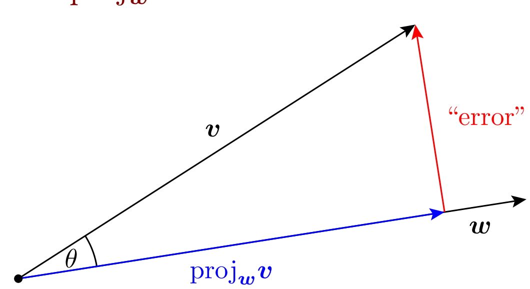
**Two obvious cases.** Observe that the cases $\theta = 0$ and $\theta = 90°$ are clearly true.

We will not discuss angles much further in this class. Just in case it is helpful, here is the typical argument given in Calculus 3 to determine the projection $\operatorname{proj}_{\boldsymbol{w}}\boldsymbol{v}$ of $\boldsymbol{v}$ onto $\boldsymbol{w}$:

From the sketch, we see that "error" $= \boldsymbol{v} - \operatorname{proj}_{\boldsymbol{w}}\boldsymbol{v}$ and that this error is orthogonal to $\boldsymbol{w}$.

Basic trigonometry tells us that the length of $\operatorname{proj}_{\boldsymbol{w}}\boldsymbol{v}$ is $\|\boldsymbol{v}\|\cos\theta$. Hence:

$$\begin{aligned}\operatorname{proj}_{\boldsymbol{w}}\boldsymbol{v} &= \underbrace{\|\boldsymbol{v}\|\cos\theta}_{\text{length}}\ \underbrace{\frac{\boldsymbol{w}}{\|\boldsymbol{w}\|}}_{\text{direction}} \\ &= \frac{\|\boldsymbol{v}\|\,\|\boldsymbol{w}\|\cos\theta}{\|\boldsymbol{w}\|}\frac{\boldsymbol{w}}{\|\boldsymbol{w}\|} = \left(\frac{\boldsymbol{v}\cdot\boldsymbol{w}}{\|\boldsymbol{w}\|^2}\right)\boldsymbol{w}\end{aligned}$$




Armin Straub
straub@southalabama.edu

**Review.** We can compute the orthogonal projection $\hat{\boldsymbol{b}}$ of $\boldsymbol{b}$ onto $W$ as follows:

- Write $W = \operatorname{col}(A)$, where the columns of $A$ are a basis of $W$.
  Then, $\hat{\boldsymbol{b}} = A\hat{\boldsymbol{x}}$ where $\hat{\boldsymbol{x}}$ is the least squares solution to $A\boldsymbol{x} = \boldsymbol{b}$ (i.e. $A^T A\hat{\boldsymbol{x}} = A^T\boldsymbol{b}$).

Assuming $A^T A$ is invertible (which, by Example 53, is automatically the case if the columns of $A$ are independent), we have $\hat{\boldsymbol{x}} = (A^T A)^{-1} A^T \boldsymbol{b}$ and hence:

**(projection matrix)** The projection $\hat{\boldsymbol{b}}$ of $\boldsymbol{b}$ onto $\operatorname{col}(A)$ is (assuming cols of $A$ are independent)

$$\hat{\boldsymbol{b}} = \underbrace{A(A^T A)^{-1} A^T}_{P} \boldsymbol{b}.$$

The matrix $P = A(A^T A)^{-1} A^T$ is the **projection matrix** for projecting onto $\operatorname{col}(A)$.

**Comment.** Replace $\boldsymbol{b}$ with $\boldsymbol{v}$ and $A$ with $\boldsymbol{w}$, and see how this formula reduces to precisely the projection formula for one vector onto another (in that case $\boldsymbol{w}^T\boldsymbol{w}$ is just a number, so no need for a matrix inverse).

**Example 52.**

(a) Determine the projection matrix $P$ for projecting onto $W = \operatorname{span}\left\{ \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$

(b) Determine the projection of $\begin{bmatrix} 3 \\ 7 \\ 4 \end{bmatrix}$ onto $W$ using the projection matrix.

**Solution.**

(a) Choosing $A = \begin{bmatrix} 1 & 0 \\ -1 & 0 \\ 0 & 1 \end{bmatrix}$, the projection matrix $P$ is $A(A^T A)^{-1} A^T = \begin{bmatrix} 1 & 0 \\ -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$
$= \begin{bmatrix} 1 & 0 \\ -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1/2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1/2 & -1/2 & 0 \\ -1/2 & 1/2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

(b) The projection of $\begin{bmatrix} 3 \\ 7 \\ 4 \end{bmatrix}$ onto $W$ is $\begin{bmatrix} 1/2 & -1/2 & 0 \\ -1/2 & 1/2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 7 \\ 4 \end{bmatrix} = \begin{bmatrix} -2 \\ 2 \\ 4 \end{bmatrix}$.

**Comment.** You can choose $A$ in any way such that its columns are a basis for $W$. The final projection matrix will always be the same.

**Example 53. (bonus challenge!)** Can you show that, if the columns of a matrix $A$ are independent, then $A^T A$ is always invertible?

**For instance.** In our example, the columns of $A = \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix}$ are indepedent. The bonus claim is that it follows that $A^T A = \begin{bmatrix} 17 & 1 \\ 1 & 5 \end{bmatrix}$ is invertible. That's certainly true here. Why is it always true?

*Hint:* Assume $A^T A$ is not invertible, so that $A^T A\boldsymbol{x} = \boldsymbol{0}$ for some $\boldsymbol{x}$. Multiply both sides with $\boldsymbol{x}^T$ and …

Send me an email (or submit your solution in class) before 2/16 to collect up to two bonus points!

Armin Straub
straub@southalabama.edu

**Example 54. (extra)**

(a) What is the matrix $P$ for projecting onto $W = \operatorname{span}\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \right\}$?

(b) Using the projection matrix, project $\begin{bmatrix} 2 \\ 3 \\ 3 \end{bmatrix}$ onto $W = \operatorname{span}\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \right\}$.

**Solution.**

(a) Choosing $A = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix}$, the projection matrix $P$ is $A(A^T A)^{-1} A^T = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \end{bmatrix}$

$= \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} \frac{1}{8} \begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \end{bmatrix} = \frac{1}{8} \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 4 & 2 \\ 2 & -4 & 2 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{bmatrix}$.

(b) The projection is $\frac{1}{2} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \\ 3 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 5 \\ 6 \\ 5 \end{bmatrix}$.

**Check.** The error $\begin{bmatrix} 2 \\ 3 \\ 3 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 5 \\ 6 \\ 5 \end{bmatrix} = \begin{bmatrix} -1/2 \\ 0 \\ 1/2 \end{bmatrix}$ is indeed orthogonal to $W$.

Armin Straub
straub@southalabama.edu