Review.

- To find a least squares solution $\hat{\boldsymbol{x}}$ of $A\boldsymbol{x}=\boldsymbol{b}$, we solve $A^TA\boldsymbol{x}=A^T\boldsymbol{b}$.

  Comment. In all of our examples (and this is typical), there is a unique least squares solution $\hat{\boldsymbol{x}}$. However, that doesn't need to be the case. For instance, in the extreme case $A=0$, any vector would be a least squares solution. Note that there is a unique least squares solution $\hat{\boldsymbol{x}}$ if and only if A^TA is invertible.

  Bonus challenge. The bonus challenge in Example 45 shows that A^TA is invertible if and only if the columns of A are linearly independent.

  [It is easy to see that, if A^TA is invertible, then the columns of A are linearly independent. Namely, suppose they were dependent. That means $A\boldsymbol{x}=\boldsymbol{0}$ for some nontrivial $\boldsymbol{x}$. But then $A^TA\boldsymbol{x}=\boldsymbol{0}$, which cannot be the case since A^TA is invertible.]

- The orthogonal projection of $\boldsymbol{b}$ onto $\mathrm{col}(A)$ is $A\hat{\boldsymbol{x}}$.

  Comment. Even if there is several least squares solutions $\hat{\boldsymbol{x}}$, they would all produce the same $A\hat{\boldsymbol{x}}$.

- To find a,b so that $y=a+bx$ best fits given points $(x_1,y_1),(x_2,y_2),\ldots$, we determine a least squares solution to

$$\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \\ \vdots & \vdots \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix}=\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \end{bmatrix}.$$

- Likewise, to find a,b,c so that $y=a+bx+cx^2$ best fits given points $(x_1,y_1),(x_2,y_2),\ldots$, we determine a least squares solution to

$$\begin{bmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \\ \vdots & \vdots & \vdots \end{bmatrix}\begin{bmatrix} a \\ b \\ c \end{bmatrix}=\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \end{bmatrix}.$$

Example 53. (homework)

(a) Find the least squares solution to $\begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 2 & 4 \\ 1 & 0 & 0 \end{bmatrix}\boldsymbol{x}=\begin{bmatrix} 3 \\ 10 \\ -10 \\ 3 \end{bmatrix}$.

(b) What is the orthogonal projection of $\begin{bmatrix} 3 \\ 10 \\ -10 \\ 3 \end{bmatrix}$ onto $\mathrm{span}\left\{\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix},\begin{bmatrix} 1 \\ -1 \\ 2 \\ 0 \end{bmatrix},\begin{bmatrix} 1 \\ 1 \\ 4 \\ 0 \end{bmatrix}\right\}$?

(c) Find a,b,c so that $y=a+bx+cx^2$ best fits the points $(1,3),(-1,10),(2,-10),(0,3)$.

Solution. (final answers only)

(a) $\hat{\boldsymbol{x}}=\begin{bmatrix} 6 \\ -9/2 \\ -3/2 \end{bmatrix}$

(b) The orthogonal projection is $[0,9,-9,6]^T$.

(c) $y=6-\frac{9}{2}x-\frac{3}{2}x^2$

Comment. Note that any three points (of course, they should have different x coordinates) uniquely characterize a quadratic function $y=a+bx+cx^2$. We have four points.

Armin Straub
straub@southalabama.edu

5 More on orthogonality

5.1 Projecting onto 1-dimensional spaces

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

Example 54. 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 $\text{col}(A)$, with $A = \boldsymbol{w}$, we first find a least squares solution to $A^TA\boldsymbol{x} = A^T\boldsymbol{b}$, 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 55. 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 1×1 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 $\left(\frac{\boldsymbol{w}\cdot\boldsymbol{v}}{\|\boldsymbol{w}\|^2}\right)\boldsymbol{w}$.

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!).

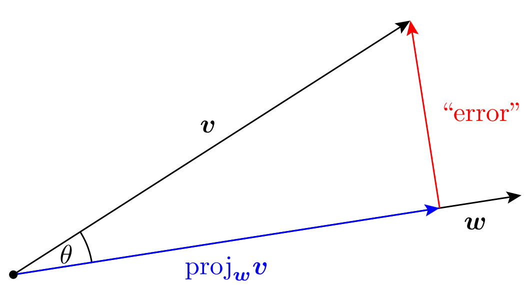
Two obvious cases. Observe that the cases $\theta = 0$ and $\theta = 90°$ are clearly true.

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 $\text{proj}_{\boldsymbol{w}}\boldsymbol{v}$ of $\boldsymbol{v}$ onto $\boldsymbol{w}$:

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

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

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



Armin Straub
straub@southalabama.edu