Orthogonal Projections

Recall: to find the best approximate solution of $Ax=b$, want to find the closest vector $\hat{b}$ to b in $\mathrm{Col}(A) = \{Ax : x \in \mathbb{R}^n\}$

Q: If V is a subspace of $\mathbb{R}^n$ and $b \in \mathbb{R}^n$, which is the closest vector b_V to b in V?

Closest means: $\|b - b_V\|$ is minimized over all $b_V \in V$

A: $b - \hat{b}$ is orthogonal to V: $b - \hat{b} \in V^\perp$ [demo]

Def: Let V be a subspace of $\mathbb{R}^n$ and $b \in \mathbb{R}^n$. The orthogonal projection of b onto V is the closest vector b_V in V to b. It is defined by

$$\boxed{b - b_V \in V^\perp}$$

The orthogonal decomposition of b relative to V is

$$\boxed{b = b_V + b_{V^\perp}}$$

Here $b_{V^\perp} = b - b_V \in V^\perp$. Note that

$$b - b_{V^\perp} = b_V \in V = (V^\perp)^\perp$$

So that $b_{V^\perp}$ is projection onto $V^\perp$.

In other words, the orthogonal decomposition is

$$\boxed{\begin{aligned} b &= \begin{pmatrix}\text{closest vector} & b_V \\ \text{to } b \text{ in } V & \end{pmatrix} + \begin{pmatrix}\text{closest vector} & b_{V^\perp} \\ \text{to } b \text{ in } V^\perp & \end{pmatrix} \\ b &= \begin{pmatrix}\text{projection of } b \\ \text{onto } V\end{pmatrix} + \begin{pmatrix}\text{projection of } b \\ \text{onto } V^\perp\end{pmatrix}\end{aligned}}$$

[demos]

How to compute b_V?

Step 0: Write V as a column space or a null space.

$V = \mathrm{Col}(A)$: Then $V^\perp = \mathrm{Nul}(A^T)$, so

$$b - b_V \in \mathrm{Nul}(A^T) \implies A^T(b - b_V) = 0$$

If $b_V \in \mathrm{Col}(A)$ then $b_V = A\hat{x}$ for $\hat{x} \in \mathbb{R}^n$:

$$A^T(b - A\hat{x}) = 0 \implies A^TA\hat{x} = A^Tb$$

Solve this equation for $\hat{x}$ $\rightsquigarrow$ $b_V = A\hat{x}$

Eg: Let $b = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ and $V = \mathrm{Col}\begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 0 \end{pmatrix}$.

Find b_V = the orthogonal projection of b to V.

We set up the equations $A^TA\hat{x} = A^Tb$:

$$A^TA = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 3 & 2 \\ 2 & 2 \end{pmatrix}$$

$$A^Tb = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

In augmented matrix form, $A^TA\hat{x} = A^Tb$ is:

$$\left(\begin{array}{cc|c} 3 & 2 & 1 \\ 2 & 2 & 1 \end{array}\right) \xrightarrow{\text{rref}} \left(\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 1/2 \end{array}\right)$$

So $\hat{x} = \begin{pmatrix} 0 \\ 1/2 \end{pmatrix} \implies b_V = A\hat{x} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 \\ 1/2 \end{pmatrix} = \begin{pmatrix} 1/2 \\ 1/2 \\ 0 \end{pmatrix}$

[demo]

Check: $b_{V^\perp} = b - b_V = \begin{pmatrix} 1/2 \\ -1/2 \\ 0 \end{pmatrix}$

$$A^Tb_{V^\perp} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}\begin{pmatrix} 1/2 \\ -1/2 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\implies b_{V^\perp} \in \mathrm{Nul}(A^T) = \mathrm{Col}(A)^\perp \quad ✓$$

Distance from V: $\|b - b_V\| = \|b_{V^\perp}\| = \left\|\begin{pmatrix} 1/2 \\ -1/2 \\ 0 \end{pmatrix}\right\| = \frac{1}{\sqrt{2}}$

Orthogonal Decomposition: $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1/2 \\ 1/2 \\ 0 \end{pmatrix} + \begin{pmatrix} 1/2 \\ -1/2 \\ 0 \end{pmatrix}$

Procedure: To compute the orthogonal projection b_V of b onto $V = \mathrm{Col}(A)$:

(1) Solve the equation $A^TA\hat{x} = A^Tb$

(2) Then $b_V = A\hat{x}$ for any solution $\hat{x}$.

Then $b_{V^\perp} = b - b_V$, and the orthogonal decomposition of b relative to V is

$$b = b_V + b_{V^\perp}.$$

The distance from b to V is $\|b_{V^\perp}\|$.

Eg: Let $b = \begin{pmatrix}1\\1\\1\end{pmatrix}$ and $V = \mathrm{Col}\begin{pmatrix}1 & -1 & -1\\ 2 & 1 & 4\\ 1 & -1 & -1\end{pmatrix}$.

Find the orthogonal decomposition of b relative to V.

(1) $$A^TA = \begin{pmatrix}1 & 2 & 1\\ -1 & 1 & -1\\ -1 & 4 & -1\end{pmatrix}\begin{pmatrix}1 & -1 & -1\\ 2 & 1 & 4\\ 1 & -1 & -1\end{pmatrix} = \begin{pmatrix}6 & 0 & 6\\ 0 & 3 & 6\\ 6 & 6 & 18\end{pmatrix}$$

$$A^Tb = \begin{pmatrix}1 & 2 & 1\\ -1 & 1 & -1\\ -1 & 4 & -1\end{pmatrix}\begin{pmatrix}1\\1\\1\end{pmatrix} = \begin{pmatrix}4\\-1\\2\end{pmatrix}$$

Solve $A^TA\hat{x} = A^Tb$:

$$\left(\begin{array}{ccc|c}6 & 0 & 6 & 4\\ 0 & 3 & 6 & -1\\ 6 & 6 & 18 & 2\end{array}\right) \xrightarrow{\text{rref}} \left(\begin{array}{ccc|c}1 & 0 & 1 & 2/3\\ 0 & 1 & 2 & -1/3\\ 0 & 0 & 0 & 0\end{array}\right)$$

$\overset{PVF}{\rightsquigarrow}$ $$\hat{x} = \begin{pmatrix} 2/3 \\ -1/3 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} -1 \\ -2 \\ 1 \end{pmatrix}$$

(2) $b_V = A\hat{x}$ for any solution. Let's use the particular solution:

$$b_V = \begin{pmatrix} 1 & -1 & -1 \\ 2 & 1 & 4 \\ 1 & -1 & -1 \end{pmatrix} \begin{pmatrix} 2/3 \\ -1/3 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

NB: $b_V = b$: what does that mean? b was already in V! More on this later.

Def: The normal equation of $Ax = b$ is

$$A^T A \hat{x} = A^T b$$

Fact: $A^T A \hat{x} = A^T b$ is always consistent! (Otherwise the Procedure wouldn't work.)

Why? I claim $\mathrm{Col}(A^T) = \mathrm{Col}(A^T A)$.

From before: $\mathrm{Nul}(A) = \mathrm{Nul}(A^T A)$

Take $(\cdot)^\perp$: $\mathrm{Nul}(A)^\perp = \mathrm{Nul}(A^T A)^\perp$

$$\mathrm{Nul}(A)^\perp = \mathrm{Row}(A) = \mathrm{Col}(A^T)$$

$$\mathrm{Nul}(A^T A)^\perp = \mathrm{Row}(A^T A) = \mathrm{Col}((A^T A)^T)$$

$$= \mathrm{Col}(A^T A) \checkmark$$

Since $A^Tb \in \mathrm{Col}(A^T) = \mathrm{Col}(A^TA)$, the equation $A^TA\hat{x} = A^Tb$ is consistent.

NB: If $\hat{x}$ and $\hat{y}$ both solve

$$A^TA\hat{x} = A^T{x} = A^TA\hat{y}$$

then $0 = A^TA\hat{x} - A^TA\hat{y} = A^TA(\hat{x} - \hat{y})$

$$\Rightarrow \hat{x} - \hat{y} \in \mathrm{Nul}(A^TA) = \mathrm{Nul}(A) \Rightarrow A(\hat{x} - \hat{y}) = 0$$

$\Rightarrow b_V = A\hat{x} = A\hat{y}$. So any soln of $A^TA\hat{x} = A^Tb$ works.

Now we know how to project onto a column space. What if $V = \mathrm{Nul}(A)$?

Then $V^\perp = \mathrm{Nul}(A)^\perp = \mathrm{Row}(A) = \mathrm{Col}(A^T)$.

So first compute $b_{V^\perp}$ = projection onto a col space, then $b_V = b - b_{V^\perp}$.

Procedure: To compute the orthogonal projection b_V of b onto $V = \mathrm{Nul}(A)$:

(1) Compute $b_{V^\perp}$ = projection onto $V^\perp = \mathrm{Col}(A^T)$ using the normal equation:

$$AA^T\hat{x} = Ab \rightsquigarrow b_{V^\perp} = A^T\hat{x}$$

(2) $b_V = b - b_{V^\perp}$

Eg: Project $b=\begin{pmatrix}1\\0\\0\end{pmatrix}$ onto $V=\mathrm{Nul}\begin{pmatrix}1&1&1\\1&1&0\end{pmatrix}$.

We need to solve $AA^T\hat{x}=Ab$

$$AA^T=\begin{pmatrix}1&1&1\\1&1&0\end{pmatrix}\begin{pmatrix}1&1\\1&1\\1&0\end{pmatrix}=\begin{pmatrix}3&2\\2&2\end{pmatrix}$$

$$Ab=\begin{pmatrix}1&1&1\\1&1&0\end{pmatrix}\begin{pmatrix}1\\0\\0\end{pmatrix}=\begin{pmatrix}1\\1\end{pmatrix}$$

$$\rightsquigarrow \left(\begin{array}{cc|c}3&2&1\\2&2&1\end{array}\right) \overset{\text{rref}}{\rightsquigarrow} \left(\begin{array}{cc|c}1&0&0\\0&1&1/2\end{array}\right)$$

So $\hat{x}=\begin{pmatrix}0\\1/2\end{pmatrix} \Longrightarrow b_{V^\perp}=A^T\hat{x}=\begin{pmatrix}1&1\\1&1\\1&0\end{pmatrix}\begin{pmatrix}0\\1/2\end{pmatrix}=\begin{pmatrix}1/2\\1/2\\0\end{pmatrix}$

$$\Rightarrow b_V=\begin{pmatrix}1\\0\\0\end{pmatrix}-\begin{pmatrix}1/2\\1/2\\0\end{pmatrix}=\begin{pmatrix}1/2\\-1/2\\0\end{pmatrix} \checkmark$$

Projection onto a Line:

Suppose $V=\mathrm{Span}\{v\}$.

Then $V=\mathrm{Col}(A)$ where $A=v$ (one column).

$A^TA=v^Tv=v\cdot v$ is a scalar

$A^Tb=v^Tb=v\cdot b$

so the normal equation becomes

$$A^TA\hat{x}=A^Tb \rightsquigarrow (v\cdot v)\hat{x}=v\cdot b$$

Then $\hat{x} = \frac{v \cdot b}{v \cdot v} \rightsquigarrow b_V = A\hat{x} = \frac{v \cdot b}{v \cdot v} v$

Here's the formula:

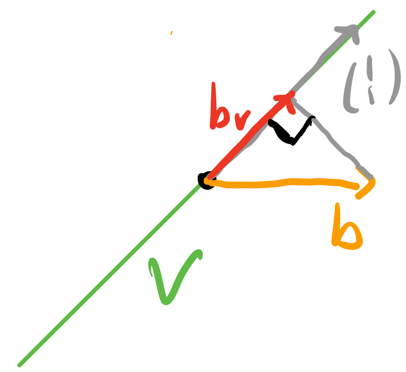
Projection onto the Line Span$\{v\}$

$$b_V = \frac{v \cdot b}{v \cdot v} v$$

Eg: Project $b = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ onto $V = \mathrm{Span}\left\{\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right\}$.

$$b_V = \frac{\begin{pmatrix} 1 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix}}{\begin{pmatrix} 1 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

[demo]



Properties of Projections: recall $b = b_V + b_{V^\perp}$

(1) $b_V = b \iff b_{V^\perp} = 0$

$\iff b \in V$

Think: the closest vector to b in V is b

$\iff$ b is already in V.

(2) $b_V = 0 \iff b = b_{V^\perp}$

$\iff b \in V^\perp$

(3) $(b_V)_V = b_V$ ($b_V \in V$; use (1))