LECTURE 22: CLOSEST VECTOR & LEAST SQUARES

We claimed $\text{Orth}_W(x) \cdot \text{Proj}_W(x) = 0$ for all $x \in \mathbb{R}^n$ provided we construct Proj_W and Orth_W as follows, if $\beta = \{w_1, \dots, w_n\}$ is orthonormal basis for $\mathbb{R}^n$ and $W = \text{span}\{w_1, \dots, w_k\}$ then $\text{Proj}_W(x) = \sum_{j=1}^{k} (x \cdot w_j) w_j$ and $\text{Orth}_W(x) = x - \text{Proj}_W(x) = \sum_{j=k+1}^{n} (x \cdot w_j) w_j$ thus

$$\begin{aligned}
\text{Orth}_W(x) \cdot \text{Proj}_W(x) &= \left(\sum_{i=k+1}^{n} (x \cdot w_i) w_i\right) \cdot \left(\sum_{j=1}^{k} (x \cdot w_j) w_j\right) \\
&= \sum_{i=k+1}^{n} \sum_{j=1}^{k} (x \cdot w_i)(x \cdot w_j) \underbrace{w_i \cdot w_j}_{\delta_{ij} = 0 \text{ since } i \neq j} \\
&= 0.
\end{aligned}$$

Thus $\text{Proj}_W(x)$ and $\text{Orth}_W(x)$ are orthogonal and we also have $x = \text{Proj}_W(x) + \text{Orth}_W(x)$ hence

$$\boxed{\|x\|^2 = \|\text{Proj}_W(x)\|^2 + \|\text{Orth}_W(x)\|^2}$$

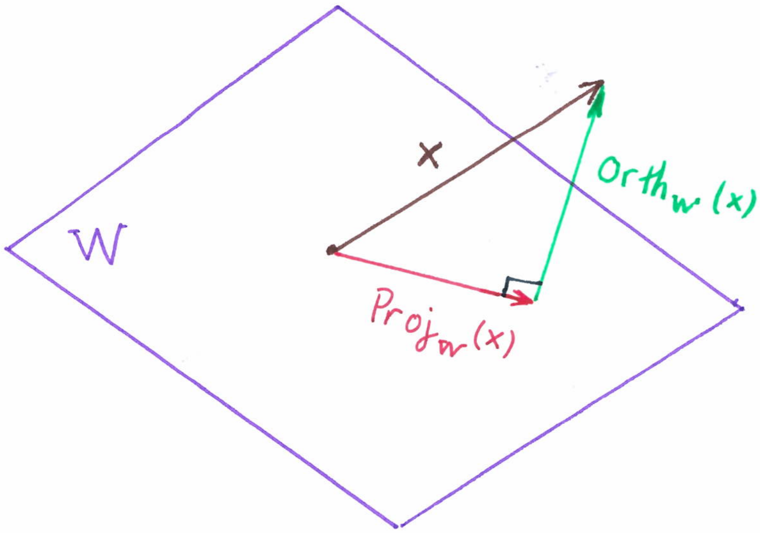
Now by construction $\text{Proj}_W(x) \in W$ whereas $\text{Orth}_W(x) \in W^\perp$

Th$^{\underline{m}}$/ Let $x \in \mathbb{R}^n$ then if W is subspace of $\mathbb{R}^n$ then $\text{Proj}_W(x)$ is the vector (point) closest to x in W. That is, if $w \in W$ then $\|w - x\|$ is minimized for the point $w = \text{Proj}_W(x)$.

<u>Proof:</u> let $w \in W$ and $x \in \mathbb{R}^n$ then $x = \text{Proj}_W(x) + \text{Orth}_W(x)$ and $w + \text{Proj}_W(x) \in W$ whereas $\text{Orth}_W(x) \in W^\perp$ and

$$\begin{aligned}
\|w - x\|^2 &= \|w - \text{Proj}_W(x) - \text{Orth}_W(x)\|^2 \\
&= \underbrace{\|w - \text{Proj}_W(x)\|^2}_{\text{this is smallest when } w = \text{Proj}_W(x)} + \|\text{Orth}_W(x)\|^2
\end{aligned}$$

Thus $\text{Proj}_W(x)$ is the closest point on W to x and we note $\|\text{Orth}_W(x)\|$ is the distance from x to W. //



E1 $W = \text{span}\{\frac{1}{\sqrt{2}}(1,1,0,0), \frac{1}{\sqrt{2}}(0,0,1,-1)\}$ then $W^{\perp} = \text{span}\{\frac{1}{\sqrt{2}}(1,-1,0,0), \frac{1}{\sqrt{2}}(0,0,1,1)\}$. Calculate

$$\text{Proj}_W(a,b,c,d) = \left(\frac{a+b}{2}\right)(1,1,0,0) + \left(\frac{c-d}{2}\right)(0,0,1,-1)$$
$$= \left(\frac{a+b}{2}, \frac{a+b}{2}, \frac{c-d}{2}, \frac{d-c}{2}\right)$$

Then, for example,

$$\text{Proj}_W(1,1,0,0) = (1,1,0,0) \quad \text{as expected.}$$
$$\text{Proj}_W(1,2,3,4) = \left(\frac{1+2}{2}, \frac{1+2}{2}, \frac{3-4}{2}, \frac{4-3}{2}\right) = \left(\frac{3}{2}, \frac{3}{2}, \frac{-1}{2}, \frac{1}{2}\right)$$

E2 $W = \text{span}\{e_3, e_5\} \subseteq \mathbb{R}^5$ then

$$\text{Proj}_W(x) = (x \cdot e_3)e_3 + (x \cdot e_5)e_5 = (0,0,x_3,0,x_5)$$
$$\text{Orth}_W(x) = x - \text{Proj}_W(x) = (x_1, x_2, 0, x_4, 0)$$

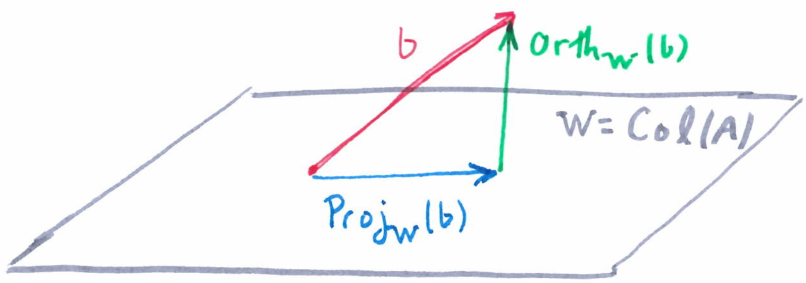
$$\text{Proj}_W(1,2,3,4,5) = \underbrace{(0,0,3,0,5)}_{\text{closest pt. to } W}$$

$$\text{Orth}_W(1,2,3,4,5) = \underbrace{(1,2,0,4,0)}_{\text{displacement from } W \text{ to } x}$$

distance $\sqrt{21}$ from W to $(1,2,3,4,5)$.

LEAST SQUARES APPROXIMATION FOR INCONSISTENT SYSTEMS

Let's return to the problem of solving $Ax = b$. If a solution $x = x_0$ exists then $Ax_0 = b$ implies that b is a linear combination of columns of A, that means $b \in Col(A)$. Therefore, if $b \notin Col(A)$ then $Ax = b$ is an <u>inconsistent</u> system. We cannot solve $Ax = b$, however we can aim to minimize $\|Ax - b\|$ to find the best approximation to $Ax = b$.



Notice $Proj_{Col(A)}(b) \in Col(A)$ thus $Au = Proj_{Col(A)}(b)$ is a consistent linear system. Moreover,

Th^m/ If $Ax = b$ is inconsistent then the solution of $Au = Proj_{Col(A)}(b)$ minimizes $\|Ax - b\|^2$

<u>Proof</u>: Suppose $Ax = b$ is inconsistent. Then $b \notin Col(A)$ and so $Orth_{Col(A)}(b) \neq 0$. Observe that, setting $W = Col(A)$,

$$\|Ax - b\|^2 = \|Ax - Proj_W(b) - Orth_W(b)\|^2$$
$$= \underbrace{\|Ax - Proj_W(b)\|^2}_{\text{this is zero when } x \text{ solves } Ax = Proj_{Col(A)}(b)} + \|Orth_W(b)\|^2$$

Th$^{\underline{m}}$ (Normal Equations)

> If $Ax = b$ is inconsistent then the solution(s) of the <u>normal equation</u> $A^T A x = A^T b$ minimizes the error given by $\|Ax - b\|$.

<u>Proof</u>: Suppose that $x \in \mathbb{R}^n$ and $A \in \mathbb{R}^{m \times n}$ then

$$
\begin{aligned}
Ax = \mathrm{Proj}_{\mathrm{Col}(A)}(b) &\iff b - Ax = b - \mathrm{Proj}_{\mathrm{Col}(A)}(b) \\
&\iff b - Ax = \mathrm{Orth}_{\mathrm{Col}(A)}(b) \\
&\iff b - Ax \in \mathrm{Col}(A)^{\perp} \\
&\iff b - Ax \in \mathrm{Null}(A^T) \\
&\iff A^T(b - Ax) = 0 \\
&\iff \underline{A^T A x = A^T b}.
\end{aligned}
$$

The minimizing assertion follows since $\mathrm{Proj}_{\mathrm{Col}(A)}(b)$ is the closest vector in $\mathrm{Col}(A)$ to $b \notin \mathrm{Col}(A)$.

<u>Remark</u>: we'll look at my typed notes for examples. Also, if $A = QR$ then the normal eq$^{\underline{n}}$s are especially nice...