EECS 245, Winter 2026

LEC 15 Projections Revisited; Regression using Linear Algebra

Agenda

- Brief recap: invertibility
- Recap: Projecting one vector onto another
- Big idea: projecting one vector onto the span of several other vectors

↑ this is how we get back to regression!

Announcements

- HW 7 due Friday
- HW 6 solutions up
- New video since last lecture, linked next to Lecture 14:

Proof that A^2 invertible $\Rightarrow$ A invertible

watch!

① If $A\begin{bmatrix}3\\1\\-1\end{bmatrix}=\begin{bmatrix}9\\4\\3\end{bmatrix}$ and

$A\begin{bmatrix}4\\0\\2\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}$, is A invertible?

no! A's cols are not linearly independent

② If AB is invertible, must A and B be invertible?

no!

A, B don't have to be square

$A=\begin{bmatrix}1&0&0\\0&1&0\end{bmatrix}_{2\times3}$ $B=\begin{bmatrix}1&0\\0&1\\0&0\end{bmatrix}_{3\times2}$

$AB=\begin{bmatrix}1&0\\0&1\end{bmatrix}$

③ If $A = BC$, and A, B, C are invertible, what is B^{-1}?

$$AC^{-1} = BCC^{-1}$$

$$AC^{-1} = B$$

$$(AC^{-1})^{-1} = B^{-1}$$

$$CA^{-1} = B^{-1}$$

$$A = BC$$

$$AA^{-1} = BCA^{-1}$$

$$I = BCA^{-1}$$

$$B^{-1}$$

same answer!

$$ABX = I$$

$$ABB^{-1}A^{-1} = I$$

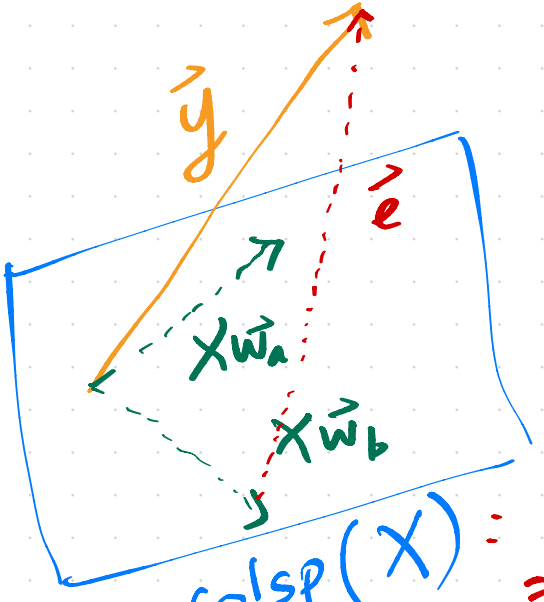
$$(AB)^{-1} = B^{-1}A^{-1}$$

$$X = \begin{bmatrix} | & | & & | \\ \vec{x}^{(1)} & \vec{x}^{(2)} & \cdots & \vec{x}^{(d)} \\ | & | & & | \end{bmatrix}_{n \times d}$$

$\vec{y} \in \mathbb{R}^n$

$X\vec{w}$, $w \in \mathbb{R}^d$,

is a linear combination of X's cols!

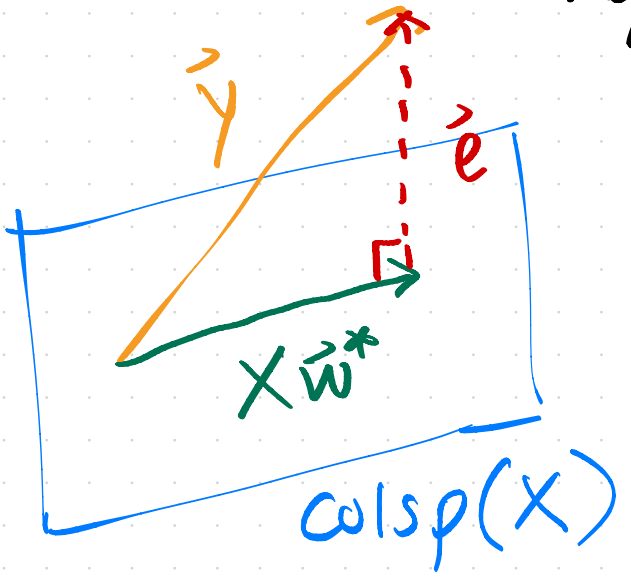


$\text{colsp}(X)$: r-dimensional subspace of $\mathbb{R}^n$

= the set of all possible values of $X\vec{w}$, where $\vec{w} \in \mathbb{R}^d$

Goal: Minimize $\|\vec{e}\|^2 = \|\vec{y} - X\vec{w}\|^2$ ~ function of $\vec{w}$ only!

How? Pick $\vec{w}$ so that $\vec{y} - X\vec{w}$ is orthogonal to every vector in colsp(X)



most important diagram of the semester!

$X_{n \times d} \vec{w}_{d \times 1}$

How do we actually find $\vec{w}^*$?

$$X = \begin{bmatrix} 1 & 1 & & 1 \\ \vec{x}^{(1)} & \vec{x}^{(2)} & \cdots & \vec{x}^{(d)} \\ 1 & 1 & & 1 \end{bmatrix}_{n \times d}$$

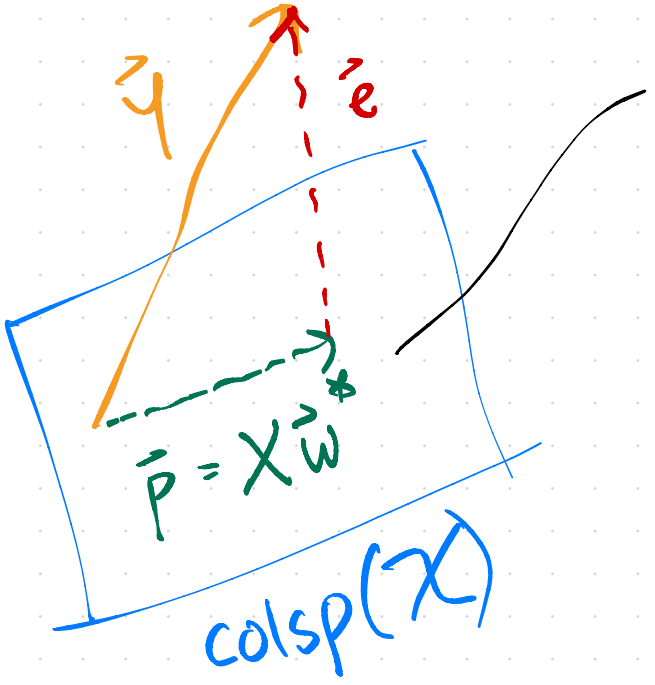
Need $\vec{y} - X\vec{w}$ to be orthogonal to all d of these columns

$$\vec{x}^{(1)} \cdot (\vec{y} - X\vec{w}) = 0$$
$$\vec{x}^{(2)} \cdot (\vec{y} - X\vec{w}) = 0$$
$$\vdots$$

$$X^T = \begin{bmatrix} — \vec{x}^{(1)T} — \\ — \vec{x}^{(2)T} — \\ \vdots \\ — \vec{x}^{(d)T} — \end{bmatrix}_{d \times n} \qquad \begin{bmatrix} \quad \end{bmatrix}$$

$\vec{x}^{(1)} \cdot (\vec{y} - X\vec{w})$

$$X^T(\vec{y} - X\vec{w}) = \begin{bmatrix} \vec{x}^{(1)T}(\vec{y} - X\vec{w}) \\ \vec{x}^{(2)T}(\vec{y} - X\vec{w}) \\ \vdots \\ \vec{x}^{(d)T}(\vec{y} - X\vec{w}) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \vec{0}_d$$



all we have to do is find the $\vec{w}^*$ that satisfies

$$X^T(\vec{y} - X\vec{w}) = \vec{0}$$

error vector

that'll give us the shortest possible error vector!

$$X^T(\vec{y} - X\vec{w}) = \vec{0}$$

$$X^T\vec{y} - X^TX\vec{w} = \vec{0}$$

$$X^TX\vec{w} = X^T\vec{y}$$

"normal equation(s)"

Case 1: If X^TX is invertible...

i.e. if X's cols are linearly independent

$$(X^TX)^{-1}X^TX\vec{w} = (X^TX)^{-1}X^T\vec{y}$$

$(X^TX)^{-1}X^TX = I$

$$\vec{w}^* = (X^TX)^{-1}X^T\vec{y}$$

unique solution

most important equation

Important:

rank(X) = rank(X^TX) !

$X: n \times d$

$X^TX: d \times d$

Case 2: X^TX isn't invertible,

i.e. X's columns are not linearly independent

Then, there are infinitely many "best" $\vec{w}^*$'s!

BUT THEY ALL CORRESPOND TO THE SAME PROJECTION, $\vec{p}$!

$\vec{p} = X\vec{w}^*$ → see HW 8, Chapter 6.4

example

$$X = \begin{bmatrix} 1 & 0 \\ 2 & 1 \\ 1 & 1 \\ 0 & -1 \end{bmatrix} \qquad \vec{y} = \begin{bmatrix} 1 \\ 0 \\ 4 \\ 5 \end{bmatrix}$$

Q: Which linear combination of X's columns is closest to $\vec{y}$?

A: $X\vec{w}^*$, where $\vec{w}^* = (X^TX)^{-1}X^T\vec{y}$

$$X^TX = \begin{bmatrix} 6 & 3 \\ 3 & 3 \end{bmatrix} \Rightarrow (X^TX)^{-1} = \frac{1}{9}\begin{bmatrix} 3 & -3 \\ -3 & 6 \end{bmatrix}$$

orthogonality

$\vec{p}$ orthogonal to $\vec{y} - \vec{p}$

$$\vec{w}^* = \begin{bmatrix} 2 \\ -8/3 \end{bmatrix}$$

of all linear combinations of X's cols, the vector closest to $\vec{y}$ is

$$2\begin{bmatrix} 1 \\ 2 \\ 1 \\ 0 \end{bmatrix} + \left(-\frac{8}{3}\right)\begin{bmatrix} 0 \\ 1 \\ 1 \\ -1 \end{bmatrix} = \vec{p}$$

first col of X