Lec 8 Math 390 2/26/19

If $\text{rank}(X) = p+1 \Rightarrow \vec{b} = (X^\top X)^{-1} X^\top \vec{y} \in \mathbb{R}^{p+1}$, the OLS estimates

New prediction...

$$\hat{y}^* = g(\vec{x}^*) = \underset{1\times(p+1)}{\vec{x}^*} \underset{(p+1)\times 1}{\vec{b}}$$

What if $\vec{x}^*$ far from previous $\vec{x} \in \mathbb{D}$?

let $\text{Range}(X) = [x_{\cdot 1, min}, x_{\cdot 1, max}] \times \ldots \times [x_{\cdot p, min}, x_{\cdot p, max}]$

$\vec{x} \notin \text{Range}(X) \Rightarrow$ "extrapolation" Bad things happen when you do this.

In every model, you should always be aware of this!!

How well does the OLS do? Just like last time...

$$\hat{y} = X\vec{b} = \underbrace{X(X^\top X)X^\top}_{H}\vec{y} = Hy$$

called the "hat" matrix since it converts $y \to \hat{y}$

$\text{rank}[H]$? $= p+1$ why? $\vec{\hat{y}} = b_0 \vec{1}_n + b_1 \vec{x}_1 + \ldots + b_p \vec{x}_p$ i.e.

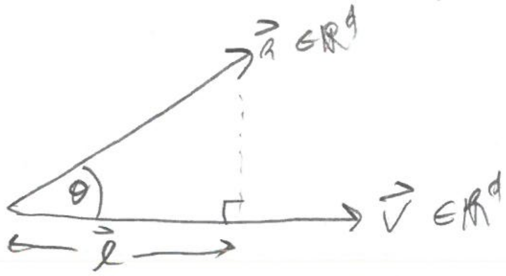
a lin. comb of $p+1$ vectors! That's rank!

$p+1$ = degrees of freedom $= \dim[\text{colsp}(X)]$ why is it called this?

$\vec{y} = \vec{\hat{y}} + \vec{e} \Rightarrow \vec{e} = \vec{y} - \vec{\hat{y}}$

$SSE = \sum e_i^2 = \vec{e}^\top \vec{e}$ $\quad MSE = \frac{1}{n-(p+1)} SSE$, $RMSE = \sqrt{MSE}$ $\quad R^2 = 1 - \frac{SSE}{SST} = \frac{s_y^2 - s_e^2}{s_y^2}$ All the same...

Let's do some linear algebra...



$\vec{l} = \text{proj}_{\vec{v}}(\vec{a})$ the "orthogonal projection of $\vec{a}$ onto $\vec{v}$"

perpendicular

By law of cosines, $\cos(\theta) = \frac{\vec{a}\cdot\vec{v}}{\|\vec{a}\|\|\vec{v}\|} = \frac{\|\vec{l}\|}{\|\vec{a}\|} \Rightarrow \|\vec{l}\| = \frac{\vec{a}\cdot\vec{v}}{\|\vec{v}\|}$

Now we need direction... multiply by length 1, correct direction $\frac{\vec{v}}{\|\vec{v}\|}$ does length of $\vec{v}$ matter at all?

$$\text{proj}_{\vec{v}}(\vec{a}) = \underbrace{\frac{\vec{a}\cdot\vec{v}}{\|\vec{v}\|^2}}_{\|\vec{l}\|}\vec{v} = \frac{(\vec{a}\cdot\vec{v})\,\vec{v}}{\|\vec{v}\|^2} = \frac{\vec{v}\vec{v}^T\vec{a}}{\|\vec{v}\|^2} = \underbrace{\frac{1}{\|\vec{v}\|^2}\overbrace{\vec{v}\vec{v}^T}^{\text{outer product}}}_{H \in \mathbb{R}^{d\times d}}\vec{a} = H\vec{a}$$

H is called a projection matrix. It projects onto $\text{colsp}(\vec{v})$

Let's project $\vec{v}$ onto itself $\text{proj}_{\vec{v}}(\vec{v}) = H\vec{v} = \frac{1}{\|\vec{v}\|^2}\vec{v}(\vec{v}^T\vec{v}) = \vec{v}$ ✓

Proj $\vec{a} - \text{proj}_{\vec{v}}(\vec{a})$ onto $\vec{v}$?
$H(\vec{a} - H\vec{a}) = H\vec{a} - HH\vec{a} = H\vec{a} - H\vec{a} = \vec{0}$

What if I project twice in a row?

$$\text{proj}_{\vec{v}}(\text{proj}_{\vec{v}}(\vec{a})) = HH\vec{a} = \frac{1}{\|\vec{v}\|^2}\vec{v}\vec{v}^T\frac{1}{\|\vec{v}\|^2}\vec{v}\vec{v}^T\vec{a} = \frac{1}{(\|\vec{v}\|^2)^2}\vec{v}\overbrace{(\vec{v}^T\vec{v})}^{\|\vec{v}\|^2}\vec{v}^T\vec{a}$$

$$= \frac{1}{\|\vec{v}\|^2}\vec{v}\vec{v}^T\vec{a} = H\vec{a} \Rightarrow HH = H \text{ "idempotency"}$$

Let's do $\vec{l} = \text{Proj}_V(\vec{a})$ where $V = [\vec{v}_1 | \vec{v}_2 | \ldots | \vec{v}_K] \in \mathbb{R}^{d\times K}$, full rank

We know $\vec{l} \in \text{Colsp}[V] \Rightarrow \vec{l} = V\vec{w}$ why?

where $\vec{w} = [w_1 \ldots w_K]$ vector of weights on each of the directions of V.

we need to solve for $\vec{w}$!