221

LAST TIME

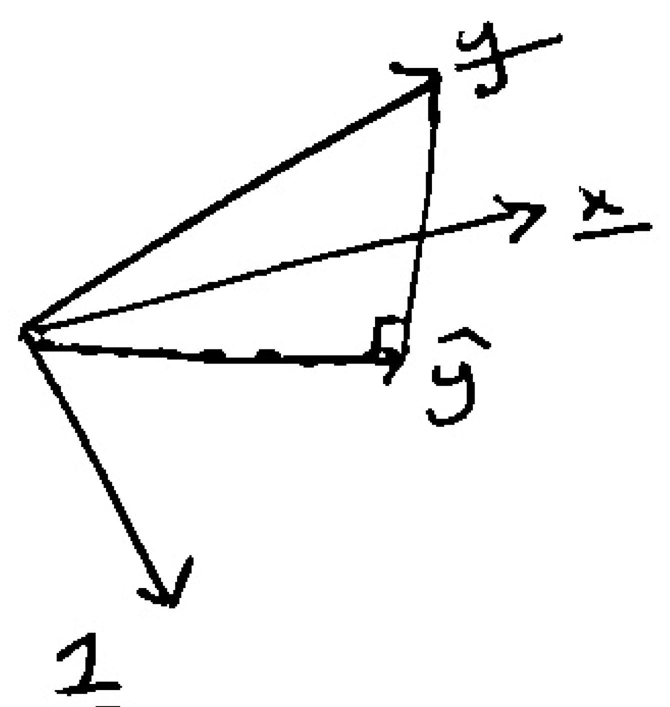
$$\hat{\underline{y}} = \underline{\underline{X}}\hat{\beta}_{OLS}$$

$\hat{\underline{y}}$ is the orthogonal projection of $\underline{y}$ onto the column space of $\underline{\underline{X}}$

In other words $\hat{\varepsilon} = y - \hat{y} \perp \text{col}(\underline{\underline{X}})$

"is orthogonal to"

PICTURE: (every observation, i.e. row of y or row of $\underline{\underline{X}}$ is a dimension in our picture).



How can we write $\hat{y}$ in multivariate terms?

$$\hat{y} = \underline{\underline{X}}\hat{\underline{\beta}} = \boxed{\underline{\underline{X}}(\underline{\underline{X}}^T\underline{\underline{X}})^{-1}\underline{\underline{X}}^T} y = \underline{\underline{H}}y$$

HAT MATRIX "$\underline{\underline{H}}$" {because when you multiply $\underline{\underline{H}}$ by y, it puts a hat on y.}

PROPERTIES OF $\underline{\underline{H}}$:

$\underline{\underline{H}}$ projects y onto the $\text{col}(\underline{\underline{X}})$

(1) symmetric: $\underline{\underline{H}} = \underline{\underline{H}}^T$

(2) idempotent: $\underline{\underline{H}}^2 = \underline{\underline{H}}$

idempotent $\Rightarrow$ proj. matrix

idempotent & symmetric $\Rightarrow$ orthogonal projection matrix.

Exercise: Show H is idempotent.

Notice that we can break apart a vector y into two components:

$$y = \underline{\underline{H}}y + y - \underline{\underline{H}}y$$
$$= \underbrace{\underline{\underline{H}}y}_{\hat{y}} + \underbrace{(\underline{\underline{I}} - \underline{\underline{H}})y}_{y-\hat{y}}$$

$\hat{y}$ — component of y projected onto $\text{col}(\underline{X})$

$y - \hat{y} = \hat{\underline{\varepsilon}}$ — residual

Exercise: Suppose $\underline{\underline{H}}$ idempotent. Show $\hat{\underline{\varepsilon}} \perp \hat{y}$ only if $\underline{\underline{H}}$ is symmetric.

SOLUTION: $\hat{\underline{\varepsilon}} \perp \hat{y}$ means the inner product between these vectors is zero, i.e. $\sum \hat{\varepsilon}_i \hat{y}_i = 0 = \hat{\underline{\varepsilon}}^T \hat{y}$

$$\hat{\underline{\varepsilon}}^T \hat{y} = \left((\underline{\underline{I}} - \underline{\underline{H}})y\right)^T \underline{\underline{H}}y$$
$$= (y^T - y^T \underline{\underline{H}}^T)\underline{\underline{H}}y$$
$$= y^T \underline{\underline{H}}y - y^T \underline{\underline{H}}^T \underline{\underline{H}}y$$

& this $= 0 \iff \underline{\underline{H}}^T \underline{\underline{H}} = \underline{\underline{H}}$ for nonzero y.