Warm Up:

1. Why does $SSE = e^Te$?

2. Why is H a projection matrix?

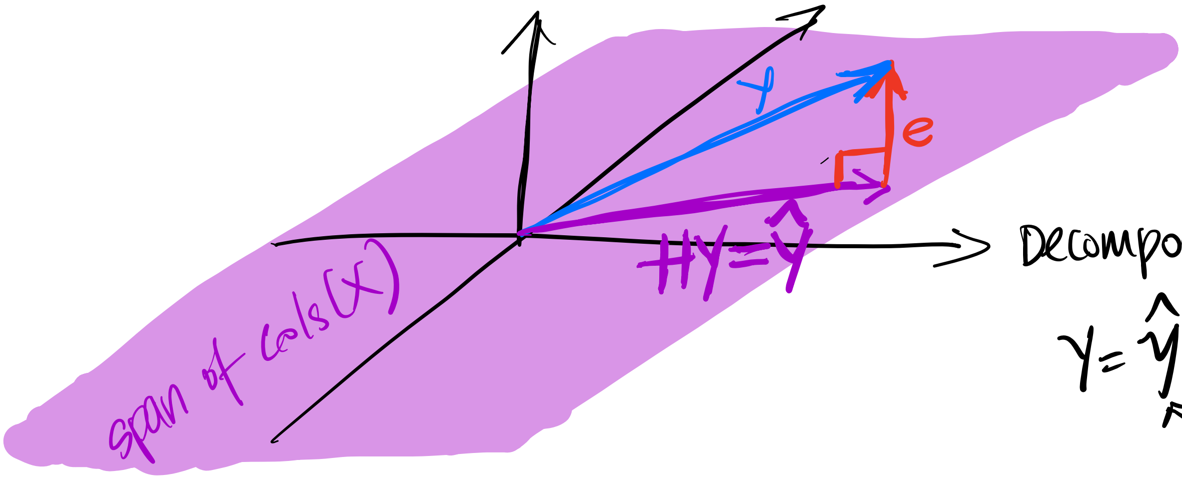
3. What does the linear map corresponding to $(I-H)$ do?

---

1. Multiply $e^Te = [e_1 \cdots e_n]\begin{bmatrix} e_1 \\ \vdots \\ e_n \end{bmatrix} = SSE$

2. H is: $\left.\begin{array}{l} \text{-idempotent} \\ \text{-symmetric} \end{array}\right]$ Prove each.

Idea: H projects Y onto the span of columns of X.



Decomposition:

$$y = \hat{y} + e$$

orthogonal

We will prove $\hat{y} \perp e$ later today.

(3) Linear map of $(I-H)$:

$$L_{(I-H)}(v) = (I-H)\,v$$

if we apply L to $v = y$, it maps $y \mapsto e$.

$$L_{(I-H)}(y) = (I-H)\,y = e.$$

Why is (2) useful?

Consider estimating σ^2....

Recall:

$$Y \sim N(X\beta, \sigma^2 I_n)$$

c.f. $y_i \sim N(\beta_0 + \beta_1 x_i, \sigma^2)$

$$Y \sim N(X\beta, \sigma^2 I_n)$$

$$Y - X\beta \sim N(0, \sigma^2 I_n)$$

$$\frac{Y - X\beta}{\sigma} \sim N(0, I_n)$$

By ④ of MVN properties...

$= 0 \dots$ why?

$$\left(\frac{Y-X\beta}{\sigma}\right)^T (I-H) \left(\frac{Y-X\beta}{\sigma}\right) = \frac{Y^T(I-H)Y + \boxed{\beta^T X^T (I-H) X\beta}}{\sigma^2}$$

$$= \frac{Y^T(I-H)Y}{\sigma^2} = \frac{SSE}{\sigma^2}$$

why!

$$(I-H)X = IX - HX$$

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

$$= X - X = 0$$

$$(I-H)(I-H) = I - HI - IH + H^2$$

$$= I - 2H + H = I - H$$

$$\frac{Y^T(I-H)Y}{\sigma^2} \sim \chi^2_{df}\ ;\ df = \text{rank}(I-H)$$

$$\text{rank}(I-H) = ???$$

Since $I-H$ is idempotent, rank = trace.

$$U^TU = I$$
$$UU^T = I$$

$$\text{rank}(I-H) = tr(I-H)$$
$$= tr(I - U\Lambda U^T)$$
$$= tr(UU^T - U\Lambda U^T)$$
$$= tr(\boxed{\underline{U}(I-\Lambda)U^T})$$

tr is cyclic

$$= tr((I-\Lambda)U^TU)$$
$$= tr(I-\Lambda)$$

p of these

$$= \mathrm{tr}\left(\begin{bmatrix} 1 & & & 0 \\ & 1 & & \\ & & \ddots & \\ 0 & & & 1 \end{bmatrix} - \begin{bmatrix} 1 & & & & & \\ & 1 & & & & \\ & & \ddots & & & \\ & & & 1 & & \\ & & & & 0 & \\ & & & & & \ddots \; 0 \end{bmatrix} \right)$$

$$= \mathrm{tr}\left(\begin{bmatrix} 0 & & & & & \\ & 0 & & & & \\ & & \ddots \, 0 & & & \\ & & & 1 & & \\ & & & & 1 \ddots & \\ & & & & & 1 \end{bmatrix} \right) = n-p$$

n-p of these

$$\therefore \quad \underbrace{\frac{Y^T(I-H)Y}{\sigma^2}}_{} \sim \underbrace{\chi^2_{n-p}}_{}$$

Claim:

$$\frac{Y^T(I-H)Y}{\sigma^2} = \frac{SSE}{\sigma^2}$$

$$SSE = e^Te = \left((I-H)Y\right)^T\left((I-H)Y\right)$$

$$= Y^T(I-H)(I-H)Y = Y^T(I-H)Y$$

Goal: Find a "good" estimator of σ^2:

$$\frac{SSE}{\sigma^2} \sim \chi^2_{n-p}$$

$$E\left(\frac{SSE}{\sigma^2}\right) = n-p$$

$$\Rightarrow E\left(\frac{SSE}{n-p}\right) = \sigma^2$$

$$\uparrow$$

$$\hat{\sigma}^2 = MSE$$

$$E(\hat{\sigma}^2) = \sigma^2$$

$\therefore$ $\hat{\sigma}^2$ is <u>unbiased</u> for σ^2 ☺

What next:

in notes from last time

① Proving that rank determines df?

② Showing that $\hat{Y}$ is independent of e?

Claim: the fitted values & the residuals are independent.

i.e. $\hat{Y} = X\hat{\beta} \perp\!\!\!\perp e = Y - X\hat{\beta} = (I-H)Y$

Pf:

Recall the eigendecomposition of H:

$$H = U \Lambda U^T = U \left[\begin{array}{c|c} I_p & 0 \\ \hline 0 & 0 \end{array}\right] U^T$$

let

$$\boxed{U = [\underset{n\times p}{U_1} \quad \underset{n\times(n-p)}{U_2}]}$$

so that

$$H = [u_1\ u_2]\begin{bmatrix} I_p & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} u_1^T \\ u_2^T \end{bmatrix} = u_1 u_1^T$$

Do something similar for

$$I - H = [u_1, u_2]\begin{bmatrix} 0 & 0 \\ 0 & I_{n-p} \end{bmatrix}\begin{bmatrix} u_1^T \\ u_2^T \end{bmatrix}$$

$$= u_2 u_2^T$$

$$Y \sim N(XB, \sigma^2 I_n)$$

$$U^T Y \sim N(U^T XB, \sigma^2 I_n)$$

$$U^T Y = \begin{bmatrix} u_1^T Y \\ u_2^T Y \end{bmatrix} \sim N\left(\begin{pmatrix} u_1^T XB \\ u_2^T XB \end{pmatrix}, \sigma^2 \begin{bmatrix} I_p & 0 \\ 0 & I_{n-p} \end{bmatrix}\right)$$

What does this say about $u_1^T Y$ & $u_2^T Y$?

$u_1^T Y \perp\!\!\!\perp u_2^T Y$ b/c

$Cov(u_1^T Y, u_2^T Y) = 0$ & joint normal.

So I know:

$$u_1^T Y \perp\!\!\!\perp u_2^T Y \Rightarrow$$

$$u_1 u_1^T Y \perp\!\!\!\perp u_2^T Y \Rightarrow$$

$$u_1 u_1^T Y \perp\!\!\!\perp u_2 u_2^T Y \Rightarrow$$

$$HY \perp\!\!\!\perp (I-H)Y$$

$$\hat{Y} \perp\!\!\!\perp e$$

$$H \cdot (I-H) = H - H^2 = H - H = 0$$

$\Rightarrow$ they live in orthogonal subspaces!

$\hat{Y} \perp\!\!\!\perp e \Rightarrow$ allows us to do inference!

Matrix Form of:

$$SSTO = SSR + SSE$$

① $SSE = e^T e = Y^T (I-H) Y$ from before.

② $SSR = \sum_{i=1}^{n} (\hat{Y}_i - \bar{Y})^2$

$$\hat{Y} = HY$$

$$\bar{Y} = \begin{bmatrix} \bar{Y} \\ \bar{Y} \\ \vdots \\ \bar{Y} \end{bmatrix} = \frac{1}{n} \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & & & \\ 1 & 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} \frac{1}{n}\sum_i y_i \\ \bar{Y} \\ \vdots \\ \bar{Y} \end{bmatrix}$$

$$= \frac{1}{n} J_n Y$$

matrix of all 1s.

$$\hat{Y} - \bar{Y} = HY - \frac{1}{n}J_n Y = (H - \frac{1}{n}J_n)Y$$

$$\parallel$$

$$\begin{bmatrix} \hat{Y}_1 - \bar{Y} \\ \hat{Y}_2 - \bar{Y} \\ \vdots \\ \hat{Y}_n - \bar{Y} \end{bmatrix}$$

$$\begin{bmatrix} \hat{Y}_1 - \bar{Y} & \hat{Y}_2 - \bar{Y} & \cdots & \hat{Y}_n - \bar{Y} \end{bmatrix} \begin{bmatrix} \hat{Y}_1 - \bar{Y} \\ \vdots \\ \hat{Y}_n - \bar{Y} \end{bmatrix} = \sum_{i=1}^{n} (\hat{Y}_i - \bar{Y})^2 = SSR$$

$$= (\hat{Y} - \bar{Y})^T (\hat{Y} - \bar{Y}) =$$

$$\left[(H - \frac{1}{n}J_n)Y \right]^T \left[(H - \frac{1}{n}J_n)Y \right]$$

$$= \left[Y^T (H - \frac{1}{n}J_n)^T (H - \frac{1}{n}J_n) Y \right] = Y^T (H - \frac{1}{n}J_n) Y$$

Is $H - \frac{1}{n}J_n$ idempotent?

$$(H - \frac{1}{n}J_n)(H - \frac{1}{n}J_n) = H^2 - \frac{HJ_n}{n} - \frac{J_n H}{n} + \frac{1}{n^2}J_n^2$$

$$= H - \frac{HJ_n}{n} - \frac{J_n H}{n} + \frac{J_n}{n}$$

Why does $HJ_n = J_n H$?

symmetry!!

When is $HJ_n = J_n$?

Notice $J_n = \mathbb{1}_n \mathbb{1}_n^T$ where $\mathbb{1}_n = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$.

we need $\mathbb{1}_n$ to be in the column space of X for $HJ_n = J_n$.

$$X = \begin{bmatrix} 1 & x_{11} & x_{21} & \cdots & x_{(p-1)1} \\ 1 & x_{12} & & & \\ \vdots & \vdots & & & \\ 1 & \vdots & & & x_{(p-1)n} \end{bmatrix}$$

it is!... if we have an intercept!!

If we have an intercept,

$$HJ_n = \boxed{H\mathbb{1}_n\mathbb{1}_n^T = \mathbb{1}_n\mathbb{1}_n^T = J_n.}$$

detailed args next page

$$= H - \frac{J_n}{n} - \frac{J_n}{n} + \frac{J_n}{n} = H - \frac{1}{n}J_n$$

Details:

$H J_n = H \mathbb{1}_n \mathbb{1}_n^T = \mathbb{1}_n \mathbb{1}_n^T = J_n$

H is a projection matrix onto the span of X.

Since $\mathbb{1}_n$ is already in the span of X, we know

$H \mathbb{1}_n = \mathbb{1}_n$.

③ $SSTO = \sum_{i=1}^{n} (y_i - \bar{y})^2$

$$Y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}, \quad \bar{Y} = \begin{bmatrix} \bar{y} \\ \vdots \\ \bar{y} \end{bmatrix} = J_n Y$$

$$= \frac{\mathbb{1}_n \mathbb{1}_n^T Y}{n}$$

$$SSTO = (Y - \bar{Y})^T (Y - \bar{Y})$$

$$= (Y^T (I - \tfrac{1}{n} J_n)^T (I - \tfrac{1}{n} J_n) Y)$$

$$= Y^T (I - \tfrac{1}{n} J_n) Y$$

ANOVA

Decomposition:

$$SSTO = SSE + SSR$$

$$Y^T\left(I-\frac{1}{n}J_n\right)Y = Y^T(I-H)Y + Y^T\left(H-\frac{1}{n}J_n\right)Y$$