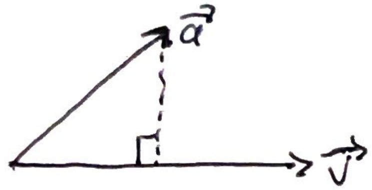
Math 342W Lecture 10



$\text{Proj}_{\vec{v}}(\vec{a}) = H\vec{a}, \ H = \frac{1}{\|\vec{v}\|^2}\vec{v}\vec{v}^T$

$d \times d$ Matrix, $\text{rank}[H] = 1$

$f(\vec{a}) = H\vec{a} \in \{c\vec{v}\} \rightarrow$ Lies on 1-d

$H^T = \left(\frac{1}{\|v\|^2}\vec{v}\vec{v}^T\right) = \frac{1}{\|v\|^2}(\vec{v}\vec{v}^T)^T = H \Rightarrow$ (symmetric)

$V = [\vec{v}_{\cdot 1} \mid \vec{v}_{\cdot 2} \mid \ldots \mid \vec{v}_{\cdot n}] \in \mathbb{R}^{d \times k}, \ k \le d, \ \text{rank}[V] = k$

↳ we are projecting $\vec{a}$ to the colspace of $\vec{V}$

$\vec{\ell} = \text{proj}_V(\vec{a}) = w_1\vec{v}_{\cdot 1} + w_2\vec{v}_{\cdot 2} + \ldots + w_k\vec{v}_{\cdot k} \quad (\exists\, w_1, \ldots, w_k)$

$= V\vec{w}$ (I)

(II) $\forall_j \ \vec{a} - \vec{\ell} \perp \vec{v}_{\cdot j} \Rightarrow \forall_j \ (\vec{a} - V\vec{w}) \cdot \vec{v}_{\cdot j} = 0$ ↝ Perpendicular

$$\Rightarrow \left.\begin{array}{l} \vec{v}_{\cdot 1}^T(\vec{a} - V\vec{w}) = 0 \\ \vec{v}_{\cdot 2}^T(\vec{a} - V\vec{w}) = 0 \\ \quad \vdots \qquad \vdots \\ \vec{v}_{\cdot k}^T(\vec{a} - V\vec{w}) = 0 \end{array}\right\} = V^T(\vec{a} - V\vec{w}) = \vec{0}_k$$

$\Rightarrow V^T\vec{a} - V^TV\vec{w} = \vec{0}_k$

$\Rightarrow V^T\vec{a} = V^TV\vec{w} \Rightarrow \vec{w} = \underline{(V^TV)^{-1}V^T\vec{a}}$

$V\vec{w} = V(V^TV)^{-1}V^T\vec{a} = \underbrace{H}_{d \times d} \Rightarrow \text{rank}[H] = k \le d = \text{proj}_V(\vec{a})$

↳ Orthogonal Projection Matrix

Two properties of Orthogonal Projection Matrices

① Symmetric

$$H^T = (V(V^TV)^{-1}V^T)^T = (V^T)^T((V^TV)^{-1})^T V^T$$

mini Lesson

$$\boxed{\begin{array}{l}\text{A is invertible and symmetric. Is } A^{-1} \text{ symmetric?}\\ A^{-1}A = I \Rightarrow A^TB = A^T(A^{-1})^T = (A^{-1}A)^T = I^T = I\\ \Rightarrow B = (A^T)^{-1} = (A^{-1})^T\end{array}}$$

$$= V((V^TV)^T)^{-1}V^T = (V(V^TV)^{-1}V^T) = H \checkmark \text{QED}$$

② Idempotent

$$HH = (V(V^TV)^{-1}V^T)(V(V^TV)^{-1}V^T) = V\overbrace{\cancel{(V^TV)^{-1}V^TV}}^{I}(V^TV)^{-1}V^T$$

$$= H \checkmark \text{ QED}$$

$$\vec{b}_{OLS} = (X^TX)^{-1}X^T\vec{y}$$

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

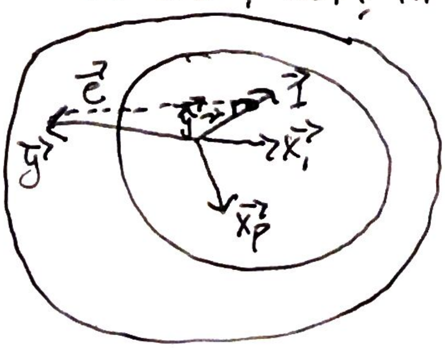
$$= H\vec{y} \Rightarrow \hat{\vec{y}} = \text{proj}_X(\vec{y})$$

$\mathcal{A}_{OLS}$: predictions are orthogonal projections of $\vec{y}$ onto colspace[X].

$X = [\vec{1} \mid \vec{x}_{\cdot 1} \mid \dots \vec{x}_{\cdot p}]$, rank[X] $= p+1$

Redraw This

$\mathbb{R}^n$



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

$\vec{e} \cdot \hat{\vec{y}} = 0$

Proof ⇒ $\vec{e} = \vec{y} - \hat{\vec{y}}$

$= \vec{y} - H\vec{y} = (I-H)\vec{y}$

$= ((I-H)\vec{y}) \cdot (H\vec{y}) = ((I-H)\vec{y})^T H\vec{y}$

$= \vec{y}^T (I-H)^T H\vec{y} = \vec{y}^T (I^T - H^T) H\vec{y}$

$= \vec{y}^T (I-H) H\vec{y} = \vec{y}^T (H - (HH))\vec{y} = \vec{y}^T (H-H)\vec{y} = 0$

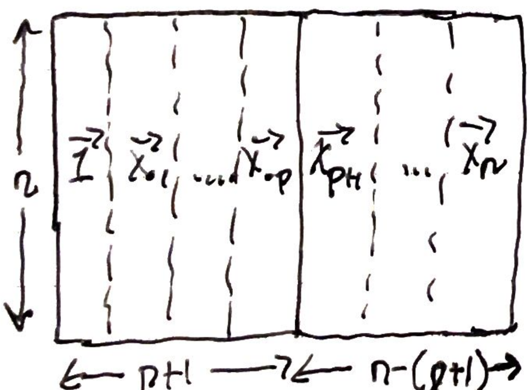
Figure I

Is $I-H$ an orthogonal projection matrix

1) $(I-H)^T = I^T - H^T = I - H$ ✓

2) $(I-H)(I-H) = II - HI - IH + HH = I - H - \cancel{H + HH} = I - H$ ✓

$(I-H)$ ⇒ projection matrix onto the residual space



Degrees of Freedom of Data

Degrees of Freedom of Residuals

$F = [X \,\vdots\, X_\perp]$

$\text{rank}[F] = n$

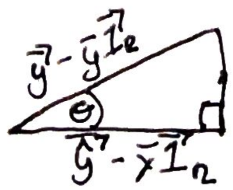
$H = X(X^T X)^{-1} X^T$

$(I-H) = X_\perp (X_\perp^T X_\perp)^{-1} X_\perp^T$

There is no unique $X_\perp$

$\|\hat{\vec{y}}\|^2 + \|\vec{e}\|^2 = \|\vec{y}\|^2$ (Figure I)

$\cos(\theta) = \dfrac{\|\hat{\vec{y}}\|}{\|\vec{y}\|}$, $\cos^2(\theta) = \dfrac{\|\hat{\vec{y}}\|^2}{\|\vec{y}\|^2}$



→ $a^2 + b^2 = c^2$

SSR + SSE = SST

Null Model g_0

$$X = [\vec{1}_n] \Rightarrow \hat{\vec{y}} = H\vec{y} = X(X^TX)^{-1}X^T\vec{y} = \vec{1}_n(\vec{1}_n^T\vec{1}_n)^{-1}\vec{1}_n^T\vec{y}$$

$$= \vec{1}_n(\vec{1}_n^T\vec{1}_n)^{-1}\Sigma y_i = \vec{1}_n \frac{1}{n}\Sigma y_i = \bar{y}\vec{1}_n$$

$$V = [\vec{v}_1 \mid \vec{v}_2] \Rightarrow \text{proj}_V(\vec{a}) \overset{?}{=} \text{proj}_{\vec{v}_1}(\vec{a}) + \text{proj}_{\vec{v}_2}(\vec{a})$$

$$= H_1\vec{a} + H_2\vec{a} = (H_1 + H_2)\vec{a} \Rightarrow H \overset{?}{=} H_1 + H_2$$

$$0 \overset{?}{=} \text{proj}_V(\vec{a})^T(\vec{a} - \text{proj}_V(\vec{a}))$$

$$= \text{proj}_V(\vec{a})^T\vec{a} - \|\text{proj}_V(\vec{a})\|^2 = (H_1\vec{a} + H_2\vec{a})^T\vec{a} - \|H_1\vec{a} + H_2\vec{a}\|^2$$

$$= ((H_1\vec{a})^T + (H_2\vec{a})^T)) - \|H_1\vec{a}\|^2 - \|H_2\vec{a}\|^2 + 2\|H_1\vec{a}\|\|H_2\vec{a}\|\cos(H_1\vec{a}\,H_2\vec{a})$$

$$= (\vec{a}^TH_1 + \vec{a}^TH_2)\vec{a} - (H_1\vec{a})^T(H_1\vec{a}) - (H_2\vec{a})^T(H_2\vec{a})$$
$$+ 2\|H_1\vec{a}\|$$

ETC

$=$