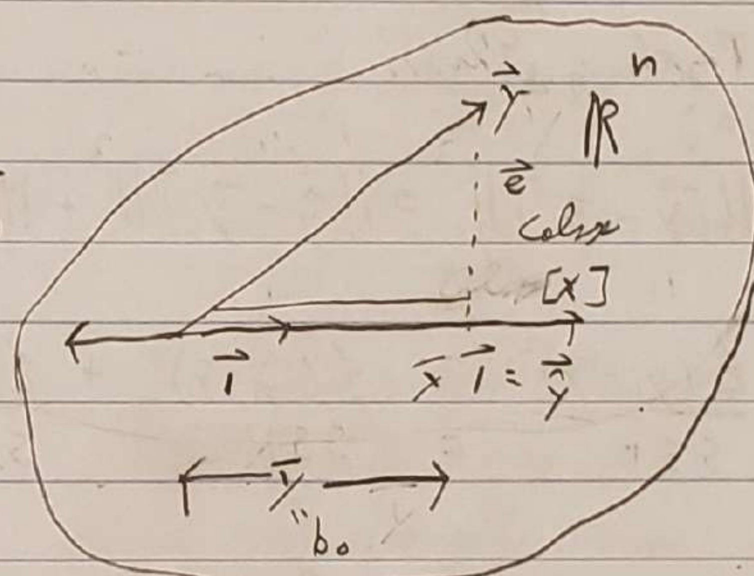
3/3/21

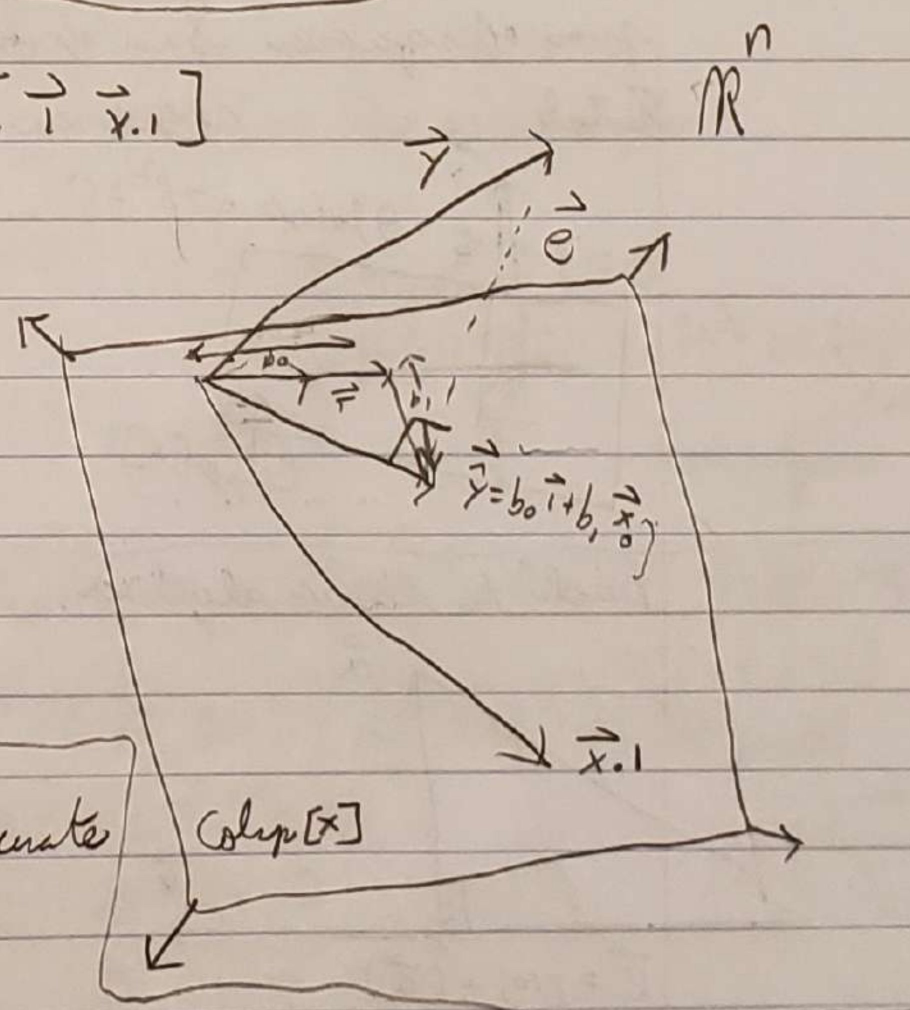
Let's examine the null model, $p=0$ so that $X = [\vec{1}_n] \Rightarrow \vec{b} = b_0 = \bar{y}$

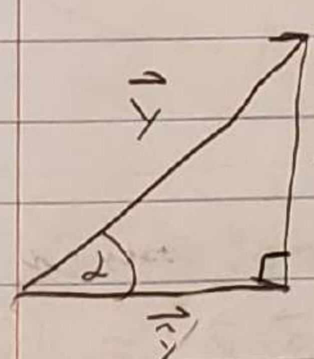
$$H = X \underbrace{\underbrace{(X^T X)}_{n}{}^{-1}}_{\frac{1}{n}} X^T = \frac{1}{n} \underset{n\times 1}{\vec{1}} \, \underset{1\times n}{\vec{1}^T} = \frac{1}{n}\begin{bmatrix} 1 & \cdots & 1 \\ \vdots & & \vdots \\ 1 & \cdots & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{n} & \cdots & \frac{1}{n} \\ \vdots & & \vdots \\ \frac{1}{n} & \cdots & \frac{1}{n} \end{bmatrix}$$

$$\hat{\vec{y}} = H \vec{y} = \begin{bmatrix} \bar{y} \\ \bar{y} \\ \vdots \\ \bar{y} \end{bmatrix} = \bar{y} \, \vec{1}_n$$



Consider $p=1$ so that $X = [\vec{1} \; \vec{x}_{.1}]$

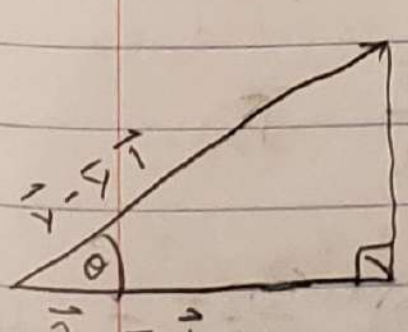




$\vec{e} \Rightarrow ||\hat{\vec{y}}||^2 + ||\vec{e}||^2 = ||\vec{y}||^2$

$$\cos^2(\alpha) = \frac{||\hat{\vec{y}}||^2}{||\vec{y}||^2}$$

Is the following illustration accurate



$\vec{e}$

(1) $H(\vec{y} - \bar{y}\vec{1})$

$$\vec{e} := \vec{y} - \hat{\vec{y}} = \vec{y} - \bar{y}\vec{1} + \bar{y}\vec{1} - \hat{\vec{y}} = (\vec{y} - \bar{y}\vec{1}) - (\hat{\vec{y}} - \bar{y}\vec{1})$$

$\text{proj}_{colsp[X]}(\vec{1}) = \vec{1}$

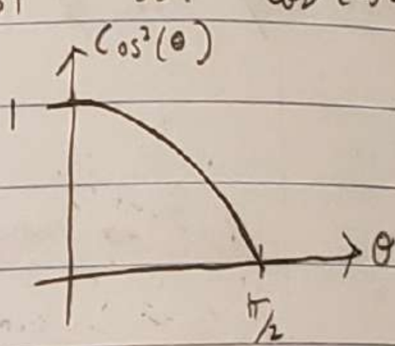
$$\text{proj}_{colsp[X]}(\vec{y} - \bar{y}\vec{1}) = H(\vec{y} - \bar{y}\vec{1}) = H\vec{y} - \bar{y}H\vec{1}$$

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

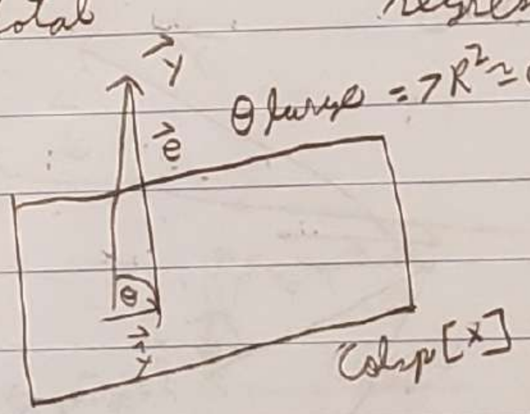
Pythag. Thm.

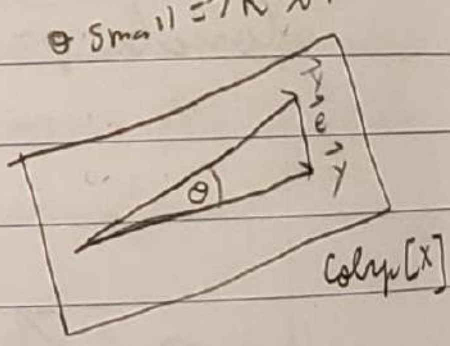
$$\|\vec{y} - \bar{y}\vec{1}\|^2 = \|\hat{\vec{y}} - \bar{y}\vec{1}\|^2 + \|\vec{e}\|^2, \quad R^2 = \frac{SST - SSE}{SST} = \frac{SSR}{SST} = \cos^2(\theta) \in [0,1]$$

$$\underbrace{\sum (y_i - \bar{y})^2}_{SST} = \underbrace{\sum (\hat{y}_i - \bar{y})^2}_{SSR} + \underbrace{\sum e_i^2}_{SSE}$$

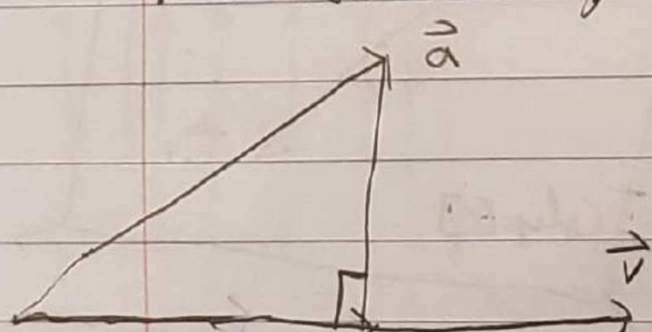


sum of squares total   Sum of squares regression   Sum of squares errors





Back to linear algebra, ...



$\vec{e} = \text{proj}_{\vec{v}}(\vec{a})$

By law of cosines,

$$\cos(\theta) = \frac{\vec{a}\cdot\vec{v}}{\|\vec{a}\|\|\vec{v}\|} \underset{\text{def cosine}}{\overset{\uparrow}{=}} \frac{\|\vec{e}\|}{\|\vec{a}\|}$$

$$\Rightarrow \|\vec{e}\| = \frac{\vec{a}\cdot\vec{v}}{\|\vec{v}\|}$$

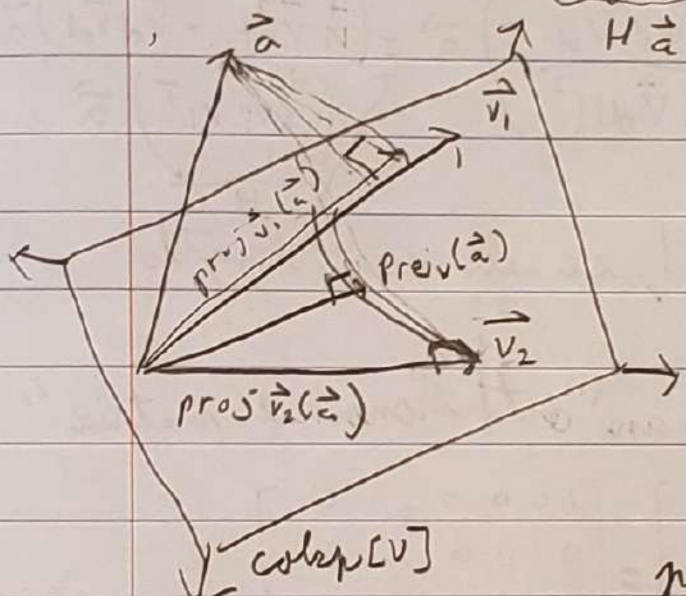
$\leftarrow \|\vec{e}\| \rightarrow$

$$\Rightarrow \vec{e} = \|\vec{e}\| \cdot \frac{\vec{v}}{\|\vec{v}\|} = \frac{\vec{a}\cdot\vec{v}}{\|\vec{v}\|^2}\vec{v} = \frac{\vec{a}^T\vec{v}\,\vec{v}}{\|\vec{v}\|^2} = \frac{\vec{v}\,\vec{v}^T\vec{a}}{\|\vec{v}\|^2} = \underbrace{\frac{\vec{v}\,\vec{v}^T}{\|\vec{v}\|^2}}_{H}\vec{a} = H\vec{a}$$

$$H = \frac{1}{\|\vec v\|^2}\vec v\,\vec v^T = \left[\frac{v_1}{\|\vec v\|^2}\vec v \;\middle|\; \frac{v_2}{\|\vec v\|^2}\vec v \;\middle|\; \cdots \;\middle|\; \frac{v_n}{\|\vec v\|^2}\vec v\right], \quad \text{rank}[H]=1$$

$$HH = \left(\frac{1}{\|\vec v\|^2}\vec v\,\vec v^T\right)\left(\frac{1}{\|\vec v\|^2}\vec v\,\vec v^T\right) = \frac{1}{\|\vec v\|^{\cancel{4}\,2}}\vec v\,\underbrace{\vec v^T\vec v}_{\cancel{\|\vec v\|^2}}\,\vec v^T = \frac{1}{\|\vec v\|^2}\vec v\,\vec v^T = H$$

$V = [\vec v_1 \mid \vec v_2]$ $\quad \underbrace{proj_V(\vec a)}_{H\vec a} \overset{?}{=} \underbrace{\overbrace{proj_{\vec v_1}(\vec a)}^{H_1\vec a} + \overbrace{proj_{\vec v_2}(\vec a)}^{H_2\vec a}} = (H_1+H_2)\vec a$ sometimes......



will always project onto Colsp[V] but it may not be the correct length (it can over/under count), the correct length gives you the right angle

$$proj_V(\vec a)\cdot(\vec a - proj_V(\vec a)) = 0$$

$\|\vec u+\vec v\|^2 = \|\vec u\|^2 + \|\vec v\|^2 + 2\|\vec u\|\|\vec v\|\underbrace{\cos(\theta)}_{\text{angle between } \vec u, \vec v}$

$$\Rightarrow proj_V(\vec a)^T\vec a - proj_V(\vec a)^T proj_V(\vec a)$$

$$= (H_1\vec a + H_2\vec a)^T\vec a - (H_1\vec a + H_2\vec a)^T(H_1\vec a + H_2\vec a) = (\vec a^T H_1 + \vec a^T H_2)\vec a - \|H_1\vec a + H_2\vec a\|^2$$

$$= \vec a^T H_1\vec a + \vec a^T H_2\vec a - \cancel{\|H_1\vec a\|^2} - \cancel{\|H_2\vec a\|^2} - 2\overbrace{\|H_1\vec a\|}^{>0}\,\overbrace{\|H_2\vec a\|}^{>0}\,\overbrace{\underbrace{\cos(\theta)}_{\text{angle between } \vec v_1, \vec v_2}}^{\in[0,1]}$$

$(H_1\vec a)^T(H_1\vec a) \quad (H_2\vec a)^T(H_2\vec a)$

$\vec a^T H_1 H_1\vec a \quad \vec a^T H_2 H_2\vec a$

$\vec a^T H_1\vec a \quad \vec a^T H_2\vec a$

The only way to make this expression zero is if $\cos(\theta)=0$ ie θ = a right angle, Thus, the full projection is a sum of the component projections if the components are orthogonal

→ How can we convert matrix V to matrix Q? There is a computational algorithm called "Gram-Schmidt" and during the computation, you can collect a matrix that is change of basis $\underset{n\times d}{V} = \underset{n\times d}{Q}\,\underset{d\times d}{R} \Rightarrow V R^{-1}$ — called Q-R decomposition of a matrix. R will be upper tri / full rank

$\mathrm{colsp}[V] = \mathrm{colsp}[Q]$

Let $V = [\vec{v}_1 \mid \vec{v}_2 \mid \ldots \mid \vec{v}_d]$, $\forall i,j\ \vec{v}_i \cdot \vec{v}_j = 0$

$V = [\vec{v}_1 \ldots \vec{v}]$, n

$$\Rightarrow \mathrm{proj}_{\mathrm{colsp}[V]}(\vec{a}) = \mathrm{proj}_{\vec{v}_1}(\vec{a}) + \ldots + \mathrm{proj}_{\vec{v}_d}(\vec{a})$$

$$= \frac{\vec{v}_1 \vec{v}_1^T}{\|\vec{v}_1\|^2}\vec{a} + \ldots + \frac{\vec{v}_d \vec{v}_d^T}{\|\vec{v}_d\|^2}\vec{a}$$

$$= \left(\frac{\vec{v}_1 \vec{v}_1^T}{\|\vec{v}_1\|^2} + \ldots + \frac{\vec{v}_d \vec{v}_d^T}{\|\vec{v}_d\|^2}\right)\vec{a} = (\vec{v}_1\vec{v}_1^T + \ldots + \vec{v}_d\vec{v}_d^T)\vec{a}$$

$$\underbrace{\left(\sum_{i=1}^{d} v_i v_i^T\right)}_{H}\vec{a}$$

If $\|\vec{v}_1\| = \|\vec{v}_2\| = \ldots = \|\vec{v}_d\| = 1$, ie all unit length

→ $Q = [\vec{v}_1 \mid \ldots \mid \vec{v}_d]$, which is an "orthonormal matrix"

$$\underset{d\times d}{Q^T Q} = \begin{bmatrix} \leftarrow \vec{v}_1^T \rightarrow \\ \leftarrow \vec{v}_d^T \rightarrow \end{bmatrix}\begin{bmatrix} \uparrow & & \uparrow \\ \vec{v}_1 & \cdots & \vec{v}_d \\ \downarrow & & \downarrow \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & \ldots & 0 \\ 0 & 1 & 0 & 0 & \ldots & 0 \\ 0 & 0 & 1 & 0 & \ldots & 0 \\ \vdots & & & & \ddots & \\ 0 & & & & & 1 \end{bmatrix} = I_d$$

$$\underset{n\times n}{Q Q^T} = \begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_d \\ \downarrow & \downarrow & & \downarrow \end{bmatrix}\begin{bmatrix} \longleftarrow \vec{v}_1^T \longrightarrow \\ \longleftarrow \vec{v}_2^T \longrightarrow \\ \vdots \\ \longleftarrow \vec{v}_d^T \longrightarrow \end{bmatrix} = \vec{v}_1\vec{v}_1^T + \vec{v}_2\vec{v}_2^T + \ldots + \vec{v}_d\vec{v}_d^T = H$$

$$= [A_1\ A_2\ \ldots\ A_d]\begin{bmatrix} B_1 \\ B_2 \\ \vdots \\ B_d \end{bmatrix} = A_1 B_1 + A_2 B_2 + \ldots + A_d B_d$$

$$\Rightarrow Q Q^T = V(V^T V)^{-1} V^T = H$$

where the columns of Q are the orthonormalized columns of $V = [v_1 | \ldots | v_d]$. Further $\mathrm{colsp}[Q] = \mathrm{colsp}[V]$ since the column vectors in Q represent a change of basis of the column vectors of V

$$\mathrm{proj}_{\mathrm{colsp}[Q]}(\vec{a}) = Q(\underbrace{Q^T Q}_{I})^{-1} Q^T = Q Q^T.$$