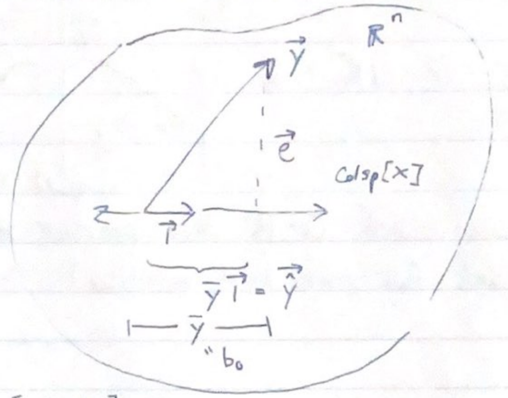
Let's examine the null model, $p = 0$ so that $X = [\vec{1}_n] \Rightarrow \vec{b} = b_0 = \bar{y}$

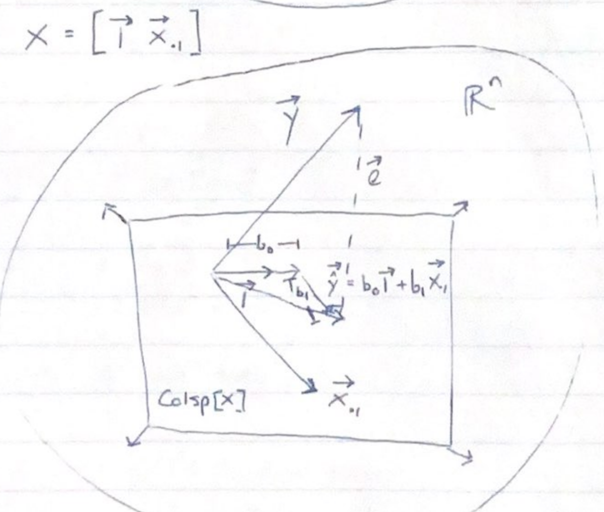
$$H = X\underbrace{\underbrace{(X^TX)}_{n}{}^{-1}}_{1/n}X^T = \frac{1}{n}\underbrace{\vec{1}}_{n\times 1}\underbrace{\vec{1}^T}_{1\times n} = \frac{1}{n}\begin{bmatrix} 1 & \cdots & 1 \\ \vdots & & \vdots \\ 1 & \cdots & 1 \end{bmatrix} = \begin{bmatrix} 1/n & \cdots & 1/n \\ \vdots & & \vdots \\ 1/n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

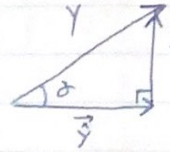
$$\hat{\vec{y}} = H\vec{y} = \begin{bmatrix} \bar{y} \\ \bar{y} \\ \vdots \\ \bar{y} \end{bmatrix} = \bar{y}\vec{1}_n$$



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

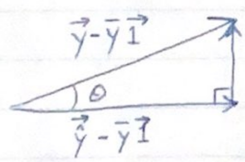
$\vec{x}_{\cdot 1}, \vec{1}$ define the plane





$\underset{\text{pythag}}{\Rightarrow} \|\hat{\vec{y}}\|^2 + \|\vec{e}\|^2 = \|\vec{y}\|^2$, $\cos^2(\theta) = \frac{\|\hat{\vec{y}}\|^2}{\|\vec{y}\|^2}$

is the following illustration accurate? Yes. $\text{proj}_{\text{colsp}[X]}(\vec{1}) = \vec{1}$



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

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

(same $\vec{e}$)

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

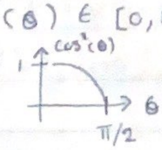
$= \hat{\vec{y}} - \bar{y}\vec{1}$ ✓

↓

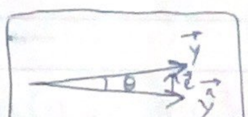
$\Rightarrow$ pythag. $\Rightarrow \|\vec{y} - \bar{y}\vec{1}\|^2 = \|\hat{\vec{y}} - \bar{y}\vec{1}\|^2 + \|\vec{e}\|^2$, $R^2 = \dfrac{SST - SSE}{SST} = \dfrac{SSR}{SST}$

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

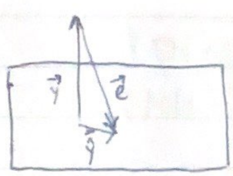
$= \cos^2(\theta) \in [0,1]$



Sum of squares total   Sum of squares regression   Sum of squares error

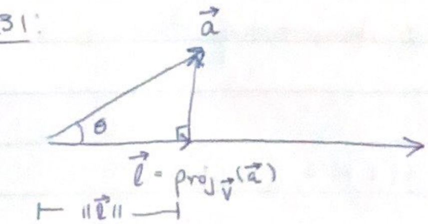


θ small $\Rightarrow R^2 \approx 1$



θ large $\Rightarrow R^2 \approx 0$

231:



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

def. cosine

$\Rightarrow \|\vec{l}\| = \dfrac{\vec{a}\cdot\vec{v}}{\|\vec{v}\|}$

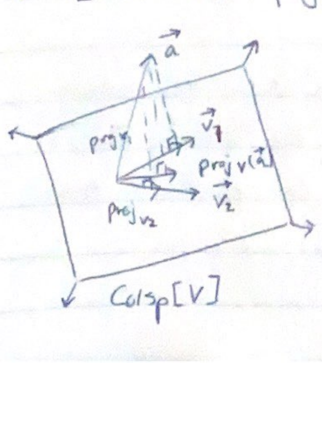
$$\Rightarrow \vec{l} = \|\vec{l}\| \cdot \frac{\vec{v}}{\|\vec{v}\|} = \frac{\vec{a}\cdot\vec{v}}{\|\vec{v}\|^2}\vec{v} = \frac{\vec{a}^T\vec{v}\vec{v}}{\|\vec{v}\|^2} = \frac{\vec{v}\vec{v}^T\vec{a}}{\|\vec{v}\|^2} = \underbrace{\frac{\vec{v}\vec{v}^T}{\|\vec{v}\|^2}}_{H}\vec{a} = H\vec{a} \checkmark$$

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

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

$$V = [\vec{v}_1 \mid \vec{v}_2] \quad \overbrace{\text{proj}_V(\vec{a})}^{H\vec{a}} \overset{?}{=} \underbrace{\overbrace{\text{proj}_{\vec{v}_1}(\vec{a})}^{H_1\vec{a}} + \overbrace{\text{proj}_{\vec{v}_2}(\vec{a})}^{H_2\vec{a}}}_{} = (H_1 + H_2)\vec{a}$$

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:

$$\text{proj}_V(\vec{a}) \cdot (\vec{a} - \text{proj}_V(\vec{a})) = 0$$

angle between $\vec{u},\vec{v}$

$\Rightarrow \ \text{proj}_{\vec{v}}(\vec{a})^T \vec{a} - \text{proj}_V(\vec{a})^T \, \text{proj}_V(\vec{a}) \qquad \|\vec{u}+\vec{v}\|^2 = \|\vec{u}\|^2 + \|\vec{v}\|^2 + 2\|\vec{u}\|\|\vec{v}\| \cos\theta$

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

$= \cancel{\vec{a}^T H_1 \vec{a}} + \vec{a}^T H_2 \vec{a} - \underbrace{\|H_1\vec{a}\|^2}_{\substack{(H_1\vec{a})^T(H_1 a) \\ \vec{a}^T H_1 H_1 \vec{a} \\ \vec{a}^T H_1 \vec{a}}} - \underbrace{\|H_2\vec{a}\|^2}_{\substack{(H_2\vec{a})^T(H_2\vec{a}) \\ \vec{a}^T H_2 H_2 \vec{a} \\ \vec{a}^T H_2 \vec{a}}} - 2\|H_1\vec{a}\|\|H_2\vec{a}\| \cos\theta \qquad \in [0,1]$

The only way to make this expression zero is if $\cos\theta = 0$ i.e. $\theta = 90°$. Thus, the full projection is a sum of the component projections if the components are orthogonal.

Let $V = [\vec{v}_1 | \vec{v}_2 | \dots | \vec{v}_d] \in \mathbb{R}^{n \times d}$, $\forall_{ij} \ \vec{v}_i \cdot \vec{v}_j = 0 \qquad V = \begin{bmatrix} \vec{v}_1 \cdots \vec{v}_d \end{bmatrix} \updownarrow n$

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

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

$$= \left(\frac{\vec{v}_1\vec{v}_1^T}{\|\vec{v}_1\|^2} + \dots + \frac{\vec{v}_d\vec{v}_d^T}{\|\vec{v}_d\|^2} \right)\vec{a} \qquad = (\vec{v}_1\vec{v}_1^T + \dots + \vec{v}_d\vec{v}_d^T)\vec{a} = (\underbrace{\sum_{i=1}^{d} \vec{v}_i\vec{v}_i^T}_{H})\vec{a}$$

if $\|\vec{v}_1\| = \|\vec{v}_2\| = \dots = \|\vec{v}_d\|$, i.e. all unit length

$\hookrightarrow$ $Q = [\vec{v}_1 | \dots | \vec{v}_d]$, which is an "orthonormal matrix"

$$\underbrace{Q^T Q}_{d\times d} = \begin{bmatrix} \longleftarrow \vec{v}_1^T \longrightarrow \\ \vdots \\ \longleftarrow \vec{v}_d^T \longrightarrow \end{bmatrix} \begin{bmatrix} \uparrow & & \uparrow \\ \vec{v}_1 & \cdots & \vec{v}_d \\ \downarrow & & \downarrow \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & & 0 \\ \vdots & & \ddots & & \\ & & & \ddots & \\ 0 & & & & 1 \end{bmatrix} = I_d$$

$$\underbrace{QQ^T}_{n\times n} = \begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_d \\ \downarrow & \downarrow & & \downarrow \end{bmatrix} \begin{bmatrix} \longleftarrow \vec{v}_1^T \longrightarrow \\ \longleftarrow \vec{v}_2^T \longrightarrow \\ \vdots \\ \longleftarrow \vec{v}_d^T \longrightarrow \end{bmatrix} = \vec{v}_1\vec{v}_1^T + \vec{v}_2\vec{v}_2^T + \dots + \vec{v}_d\vec{v}_d^T = H$$

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

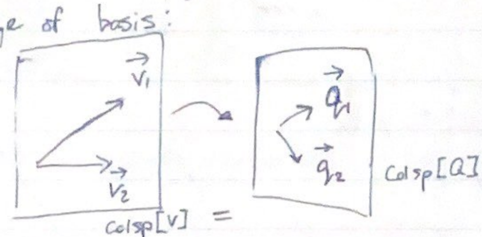
where the cols. of Q are the orthonormalized cols of $V = [\vec{v}_1 | \dots | \vec{v}_d]$. Further $\text{colsp}[Q] = \text{colsp}[V]$ since the col vectors in Q represent a change of basis of the col vecs of V.

$$= [A_1 A_2 \ \ A_d] \begin{bmatrix} B_1 \\ B_2 \\ \vdots \\ B_d \end{bmatrix} = A_1B_1 + A_2B_2 + \dots + A_dB_d$$

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

How can we convert matrix V to matrix Q? There is a computational algorithm called "Graham-Schmidt" and during the computation, you can collect a matrix that is the change of basis:

$$\underset{n\times d}{V} = \underset{n\times d}{Q}\,\underset{d\times d}{R} \Rightarrow VR^{-1} = Q$$



This is also called Q-R decomposition of a matrix. R will be upper triangular and full rank (and invertible).