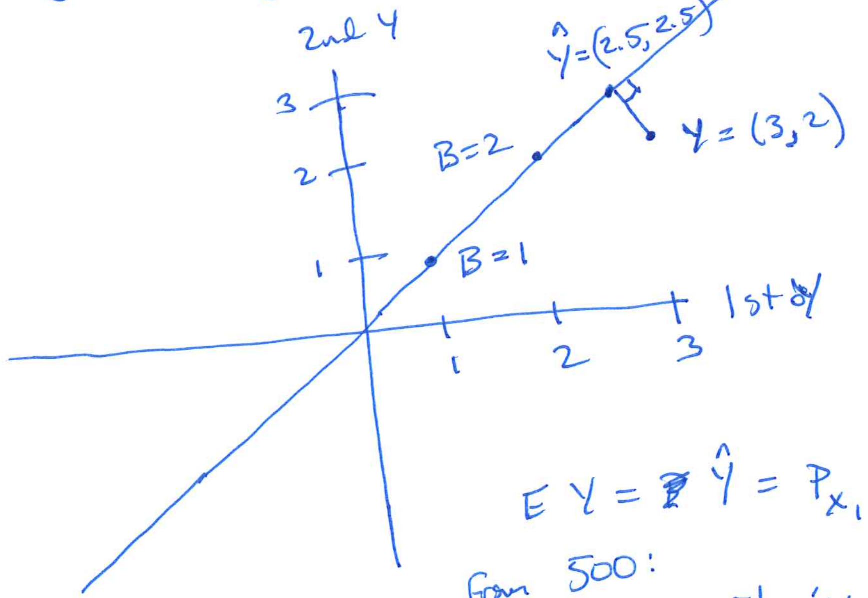
The geometry of regression

2 obs

X_1	Y
1	3
1	2



$$E\,Y = \hat{Y} = P_{X_1} Y$$

from 500:

$$\hat{\beta} = (X_1'X_1)^{-1} X_1' Y$$

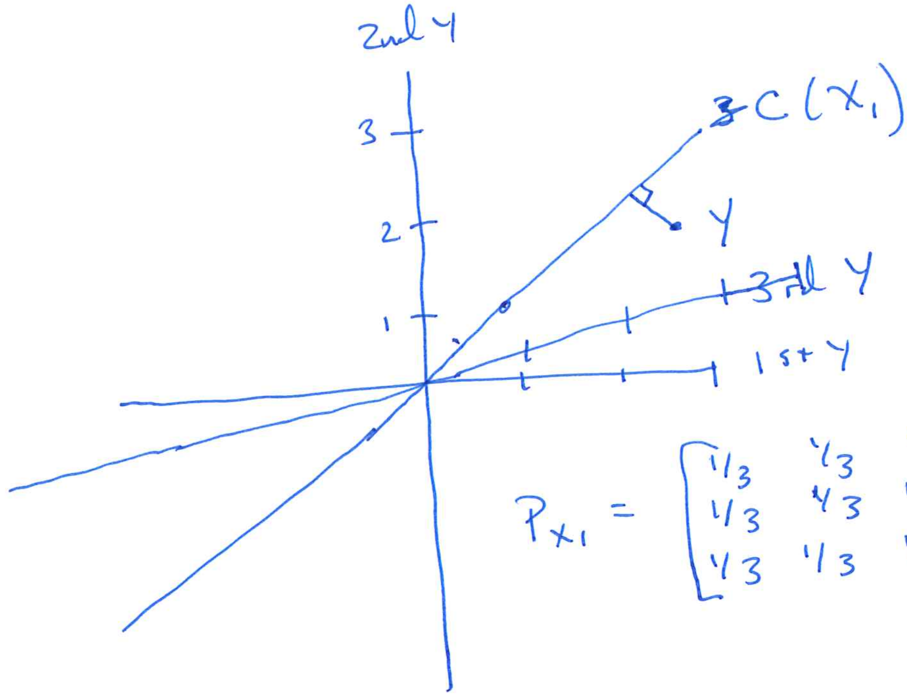
$$\hat{Y} = X\hat{\beta} = X_1 (X_1'X_1)^{-1} X_1'$$

$$P_{X_1} = X_1 (X'X)^{-1} X_1' = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}$$

$I - P_{X_1}$ = Projection operator into residual space (the vector $\perp$ to $C(X_1)$)

3 obs

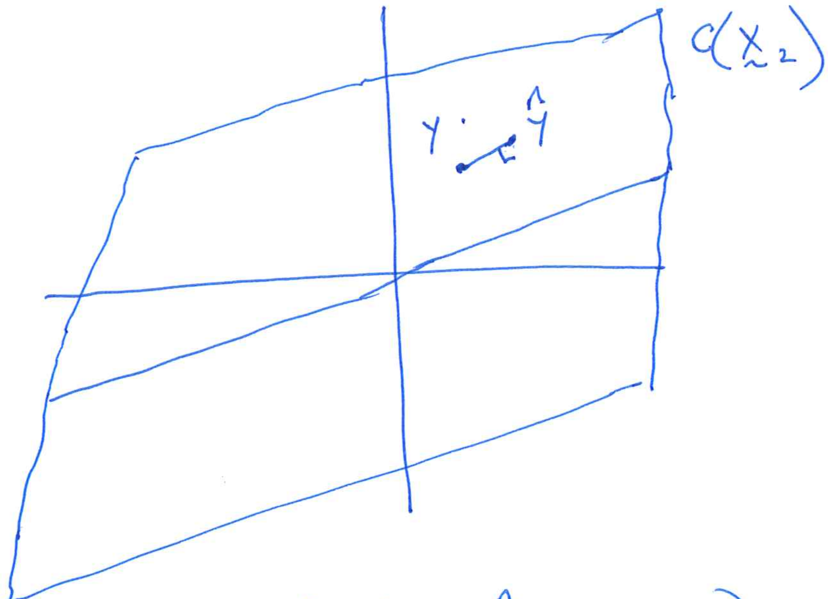
X_1	Y
1	3
1	2
1	4



$$P_{X_1} = \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{bmatrix}$$

3 obs, case 2

$\underset{\sim}{X}_2$		
X_1	X_2	Y
1	0	3
1	0	2
0	1	4



$$P_{X_2} = \begin{bmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 1/2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$C(X_2) = \{ \hat{Y}_1 = \hat{Y}_2 ,\ \hat{Y}_3 \text{ anything} \}$

a plane in 3D

1) The squared Euclidean distance between Y and $\hat{Y} = \Sigma (Y_i - \hat{Y}_i)^2$
$= SSE$!

2) The vectors $\hat{Y}$ and $Y - \hat{Y}$ are perpendicular
so crossproduct $= 0 \Rightarrow \hat{Y}$ and $Y - \hat{Y}$ are uncorrelated
if model is correct

3) 

$\Rightarrow$ Pythagorean Th'm

$\| \underset{\sim}{Y} \| = \| \hat{Y} \| + \| Y - \hat{Y} \|$

$SS_{total} = SS_{model} + SS_{error}$

4) a) $\underset{\sim}{\hat{Y}} = \underset{\sim}{P}_X \underset{\sim}{Y}$

b) $\underset{\sim}{P}_X \underset{\sim}{X} = \underset{\sim}{X}$, $\underset{\sim}{P}_X \underset{\sim}{X}' = \underset{\sim}{X}'$

c) $\underset{\sim}{P}_X$ is idempotent! $\quad \underset{\sim}{P}_X \underset{\sim}{P}_X = \underset{\sim}{P}_X$

d) $\underset{\sim}{P}_X$ is symmetric $\quad \underset{\sim}{P}_X' = \underset{\sim}{P}_X$

Non-full rank $\underset{\sim}{X}$:

$P_X = \underset{\sim}{X} (X'X)^{-} X'$

$P_X P_X =$
$\underset{\sim}{X} \underbrace{(X'X)^{-} X' X (X'X)^{-}}_{} X'$
$= X (X'X)^{-} X' = P_X$