4) A square + symmetric $\Rightarrow$
$\exists$ at least one symmetric A^-

P_X is sometimes called "the hat matrix" and written H

Properties of projection

$P_X' = P_X$ (P_X is symmetric)

and $P_X P_X = P_X$ (P_X is idempotent)

This makes sense because $P_X y \in C(X)$ and thus

$$P_X(P_X y) = P_X y \quad \forall y$$

Notice

$$\underset{\sim}{e} = Y - \hat{Y} = (I - P_X)Y$$

vector of residuals

and we need to think about this matrix $I - P_X$, what it's doing and more about the geometry of least squares

A bit more matrix/linear algebra background

- Two vectors $\underset{\sim}{u}$ and $\underset{\sim}{v}$ are perpindicular provided $\underset{\sim}{u}'\underset{\sim}{v} = 0$ often people write $\underset{\sim}{u} \perp \underset{\sim}{v}$

day 3

- There is a subspace of $\mathbb{R}^n$ that consists of all vectors perpindicular to all elements of $C(X)$ i.e.

$$C(X)^\perp = \{\underset{\sim}{u} \mid \underset{\sim}{u}'\underset{\sim}{v} = 0 \ \forall \underset{\sim}{v} \in C(X)\}$$

$$= \{\underset{\sim}{u} \mid \underset{\sim}{u}'\underset{\sim}{x}_i = 0 \ \forall \underset{\sim}{x}_i \text{ a column of } X\}$$

- $I - P_X$ is an orthogonal projection matrix (it's symmetric and idempotent)
- $C(I - P_X) = C(X)^\perp$
- $\text{rank}(X) = \text{rank}(P_X) = $ dimension of $C(X)$

$\quad = \text{trace}(P_X)$

Thm 2.13D page 50 of Rencher (also top of page 414 Christensen) and note

$\text{rank}(I - P_X) = \text{dimension of } C(X)^{\perp}$
$\parallel$
$\text{trace}(I - P_X)$ — again see page 50 of Rencher

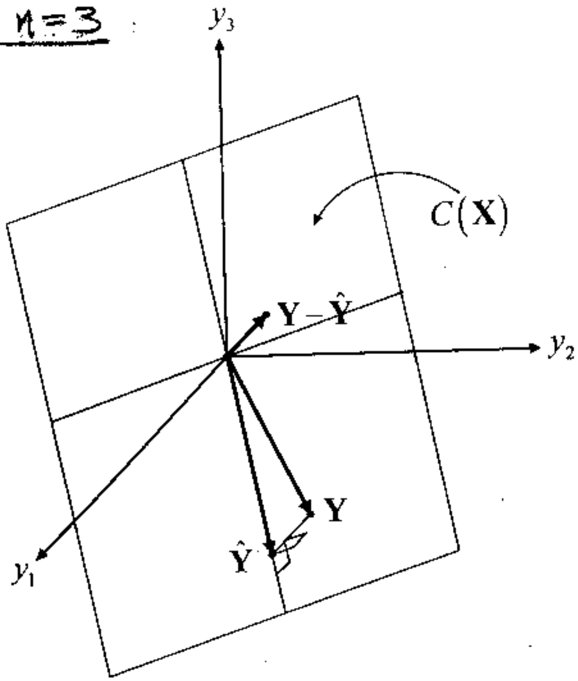
$\text{trace } I = \text{trace}(P_X) + \text{trace}(I - P_X)$
$\parallel$
n

Thm 2.11A page 40 Rencher

rank (X) rank $(I - P_X)$

So $\text{rank}(I - P_X) = \dim C(X)^{\perp} = n - \text{rank}(X)$

Geometry $n = 3$



$$
\begin{aligned}
Y'Y &= \left((P_X + (I - P_X))Y\right)'Y \\
&= Y'P_X'Y + Y'(I - P_X)'Y \\
&= Y'P_XY + Y'(I - P_X)Y \\
&= Y'P_X'P_XY + Y'(I - P_X)'(I - P_X)Y \\
&= (P_XY)'P_XY + ((I - P_X)Y)'(I - P_X)Y
\end{aligned}
$$

that is, the squared length of Y is the sum of squared lengths of P_XY and $(I - P_X)Y$ – This is an ANOVA identity/ a Pythgorean Thm

(Ordinary) Least Squares Estimation of $\underset{\sim}{\beta}$ in Full Rank Cases

When $\text{rank}(X) = k$ X has k linearly independent columns and every $\underset{\sim}{w} \in C(X)$ has a unique representation as a linear combination of columns of X –

$$X\underset{\sim}{b}_1 = X\underset{\sim}{b}_2 \Rightarrow X(\underset{\sim}{b}_1 - \underset{\sim}{b}_2) = \underset{\sim}{0}$$

$$\Rightarrow \underset{\sim}{b}_1 - \underset{\sim}{b}_2 = \underset{\sim}{0}$$

$$\underset{\sim}{b}_1 = b_2$$

(otherwise the columns of X are linearly dependent)

So in this case $\exists$ a unique $\underset{\sim}{b}$ that solves

$$X\underset{\sim}{b} = P_X Y \leftarrow \hat{Y} = \widehat{X\underset{\sim}{\beta}}$$

we can call this $\underset{\sim}{b}$ the ordinary least squares estimate of $\underset{\sim}{\beta}$ — notice that

$$X\underset{\sim}{b}_{OLS} = P_X Y = X(X'X)^- X'Y$$

$$(X'X)\underset{\sim}{b}_{OLS} = X'X(X'X)^- X'Y$$

$k \times k$ with the same rank as X, namely k, so $X'X$ is non-singular, $(X'X)^- = (X'X)^{-1}$

(Ordinary) Least Squares Estimation of $\underset{\sim}{c}'\underset{\sim}{\beta}$

day 4

When X is of full rank the obvious OLS estimate of $\underset{\sim}{c}'\underset{\sim}{\beta}$ is

$$\widehat{\underset{\sim}{c}'\underset{\sim}{\beta}}_{OLS} = \underset{\sim}{c}'\underset{\sim}{b}_{OLS}$$

When X is not of full rank $\underset{\sim}{c}'\underset{\sim}{\beta}$ can be unambigous (even though $\underset{\sim}{\beta}$ is ambiguous)

so

$$(X'X)\underset{\sim}{b}_{OLS} = X'Y$$ so called "normal equations"

$$\underset{\sim}{b}_{OLS} = (X'X)^{-1}X'Y$$

When X is not of full rank, there are multiple $\underset{\sim}{b}$'s that will solve

$$X\underset{\sim}{b} = P_X Y$$

and multiple $\underset{\sim}{\beta}$'s that could be used to represent $EY \in C(X)$ . . . ????
— so there is no sensible unique "least squares estimator of $\underset{\sim}{\beta}$"

Example b) 1 way ANOVA model

Version 1 $y_{ij} = \mu_i + \epsilon_{ij}$

will typically give full rank X and meaningful $\underset{\sim}{b}_{OLS}$

and "clearly" we should call $\bar{y}_{i\cdot}$ the OLS estimate of μ_i

Version 2 $y_{ij} = \mu + \tau_i + \epsilon_{ij}$

This gives a less than full rank X

— clearly $\mu_1 = \mu + \tau_1$

$$\underbrace{(1\ 1\ 0\ 0)}_{\underset{\sim}{c}'}\underbrace{\begin{pmatrix}\mu\\ \tau_1\\ \tau_2\\ \tau_3\end{pmatrix}}_{\underset{\sim}{\beta}} = \mu + \tau_1$$

?? What is "special" about $\underset{\sim}{c}$ in this example version 2 that makes $\underset{\sim}{c}'\underset{\sim}{\beta}$ unambiguous even though $\underset{\sim}{\beta}$ is ambiguous?

<u>Result</u>: The following are equivalent regarding $\underset{\sim}{c} \in \mathbb{R}^k$

1) $\exists\ \underset{\sim}{a} \in \mathbb{R}^n$ such that $\underset{\sim}{a}' X \underset{\sim}{\beta} = \underset{\sim}{c}'\underset{\sim}{\beta}\ \forall \underset{\sim}{\beta}$

2) $\underset{\sim}{c} \in C(X')$ ($\underset{\sim}{c}'$ is a l.c. of <u>rows</u> of X)

3) $X\underset{\sim}{\beta}_1 = X\underset{\sim}{\beta}_2 \Rightarrow \underset{\sim}{c}'\underset{\sim}{\beta}_1 = \underset{\sim}{c}'\underset{\sim}{\beta}_2$

(This last one is a condition about lack of ambiguity for $\underset{\sim}{c}'\underset{\sim}{\beta}$)

condition 1) is

$$\underset{\sim}{a}' X \underset{\sim}{\beta} = \underset{\sim}{a}' EY = E\underset{\sim}{a}' Y$$
$$\underset{\sim}{a}' X \underset{\sim}{\beta} = \underset{\sim}{c}'\underset{\sim}{\beta}$$

i.e. $\underset{\sim}{a}' Y$ can be used to estimate $\underset{\sim}{c}'\underset{\sim}{\beta}$ in an unbiased fashion

<u>Pf:</u> Suppose 2) holds

$\underset{\sim}{c} \in C(X') \Rightarrow \exists\ \underset{\sim}{a}$ s.t. $\underset{\sim}{c}' = \underset{\sim}{a} X$

so for this $\underset{\sim}{a}$, $\underset{\sim}{c}'\underset{\sim}{\beta} = \underset{\sim}{a} X \underset{\sim}{\beta}\ \ \forall \underset{\sim}{\beta}$

and 1) holds

Now suppose 1) holds i.e. $\exists\ \underset{\sim}{a}$ s.t. $\underset{\sim}{a}' X \underset{\sim}{\beta} = \underset{\sim}{c}'\underset{\sim}{\beta}\ \forall \beta$. Suppose $X\underset{\sim}{\beta}_1 = X\underset{\sim}{\beta}_2$

then for this $\underset{\sim}{a}$, $\underset{\sim}{a}' X \underset{\sim}{\beta}_1 = \underset{\sim}{a}' X \underset{\sim}{\beta}_2$

$$\underset{\sim}{a}' X \underset{\sim}{\beta}_1 = \underset{\sim}{c}'\underset{\sim}{\beta}_1, \qquad \underset{\sim}{a}' X \underset{\sim}{\beta}_2 = \underset{\sim}{c}'\underset{\sim}{\beta}_2$$

and 3) holds

Now suppose that 3) holds –

it is "obviously" equivalent to write

$$X(\underset{\sim}{\beta}_1 - \underset{\sim}{\beta}_2) = \underset{\sim}{0} \Rightarrow \underset{\sim}{c}'(\underset{\sim}{\beta}_1 - \underset{\sim}{\beta}_2) = 0\ \ \forall \underset{\sim}{\beta}_1, \underset{\sim}{\beta}_2$$

i.e. its equivalent to write

$$X\underset{\sim}{d} = \underset{\sim}{0} \Rightarrow \underset{\sim}{c}'\underset{\sim}{d} = 0\ \ \forall \underset{\sim}{d}$$

claim that 3) $\Rightarrow$ 2) is the claim that

$\{\underset{\sim}{c} \mid$ holds$\} \subset C(X')$

Suppose that $\underset{\sim}{c}$ is such that holds

Write $\underset{\sim}{c} = P_{X'}\underset{\sim}{c} + (I - P_{X'})\underset{\sim}{c}$

and let $\underset{\sim}{d}^* = (I - P_{X'})\underset{\sim}{c}$ – I'll argue $\underset{\sim}{d}^* = \underset{\sim}{0}$

clearly $\underset{\sim}{d}^* \in C(X')^{\perp}$ so it must be that

$$\underset{\sim}{c}'\underset{\sim}{d}^* = 0$$

from $\boxed{}$ condition but

$$\underset{\sim}{c}'\underset{\sim}{d}^* = \underset{\sim}{c}'(I - P_{X'})\underset{\sim}{c}$$
$$= \underset{\sim}{c}'\underset{\sim}{c} - \underset{\sim}{c}'P_{X'}\underset{\sim}{c}$$

So $\underset{\sim}{c}'\underset{\sim}{c} = \underset{\sim}{c}'P_{X'}\underset{\sim}{c}$. But

$$\underset{\sim}{c}'\underset{\sim}{c} = \underset{\sim}{c}'\left(P_{X'} + (I - P_{X'})\right)\underset{\sim}{c}$$
$$= \underset{\sim}{c}'P_{X'}\underset{\sim}{c} + \underset{\sim}{c}'(I - P_{X'})\underset{\sim}{c}$$

$$= \underset{\sim}{c}'P_{X'}\underset{\sim}{c} + \underbrace{\underset{\sim}{c}'(I - P_{X'})'(I - P_{X'})\underset{\sim}{c}}_{\underset{\sim}{d}^{*\prime}\underset{\sim}{d}}$$

$\underset{\sim}{d}^{*\prime}\underset{\sim}{d} = 0$ so $\underset{\sim}{d}^* = \underset{\sim}{0}$ and $\underset{\sim}{c} = P_{X'}\underset{\sim}{c}$ and $\underset{\sim}{c} \in C(X')$

days