

Recap:

① Projection matrix. Projection matrix \underline{P} for the $C(\underline{x})$ has the property that $\underline{P}\underline{v} = \underline{v}$ if $\underline{v} \in C(\underline{x})$ and $\underline{P}\underline{w} = \underline{0}$ if $\underline{w} \in$ orthogonal complement of the column space of \underline{x} .

Thm: If $\underline{M}, \underline{M}_0$ are projection matrices for $C(\underline{M})$ and $C(\underline{M}_0)$ with $C(\underline{M}_0) \subset C(\underline{M})$, then $C(\underline{M} - \underline{M}_0)$ is the orthogonal complement of $C(\underline{M}_0)$ w.r.t. $C(\underline{M})$.

If \underline{P} is the projection matrix onto the column space of $C(\underline{x})$ and \underline{x} is an $n \times p$ matrix then $(\underline{I} - \underline{P})$ is the projection matrix in the $N(\underline{x}^T)$.

Thm: $\text{rank}(\underline{P}) + \text{rank}(\underline{I} - \underline{P}) = n$

Pf: Any projection matrix satisfies $\underline{P}^2 = \underline{P}$
 $\Rightarrow \text{Rank}(\underline{P}) = \text{trace}(\underline{P}) = \text{sum of the diagonals of } \underline{P}$.

$$\text{trace}(\underline{P}) + \text{trace}(\underline{I} - \underline{P}) = n$$

$$\Rightarrow \text{rank}(\underline{P}) + \text{rank}(\underline{I} - \underline{P}) = n$$

Notice: $\underline{y} = \underline{P}\underline{y} + (\underline{I} - \underline{P})\underline{y}$

Here $\underline{P}\underline{y} \in C(\underline{x})$ and $(\underline{I} - \underline{P})\underline{y} \in N(\underline{x}^T)$

$$\text{Also, } \textcircled{0} (\underline{P}\underline{y})^T (\underline{I} - \underline{P})\underline{y} = \underline{y}^T \underline{P}^T (\underline{I} - \underline{P})\underline{y} \\ = \underline{y}^T (\underline{P}^T - \underline{P}^T \underline{P})\underline{y} = 0 \quad (\text{as } \underline{P} \text{ is projection matrix})$$

①

$\underline{P}\underline{y}$ and $(\underline{I}-\underline{P})\underline{y}$ are the two components residing in orthogonal spaces ($C(\underline{x})$ and $N(\underline{x}^T)$ respectively) such that their sum is \underline{y} .

Def: A generalized inverse of a $m \times n$ matrix \underline{A} is any $n \times m$ matrix \underline{G} satisfying $\underline{A}\underline{G}\underline{A} = \underline{A}$. The notation \underline{A}^- is used to denote the generalized inverse of \underline{A} .

Result: Let \underline{A} be an $m \times n$ matrix with rank k . If \underline{A} can be partitioned as below, with $\text{rank}(\underline{A}) = \text{rank}(\underline{C}) = k$,

$$\underline{A} = \begin{bmatrix} \underline{C}_{k \times k} & \underline{D}_{k \times (n-k)} \\ \underline{E}_{(m-k) \times k} & \underline{F}_{(m-k) \times (n-k)} \end{bmatrix}$$

then one generalized inverse of \underline{A} is given

$$\underline{A}^- = \begin{bmatrix} \underline{C}^{-1} & \underline{O}_{k \times (m-k)} \\ \underline{O}_{(n-k) \times k} & \underline{O}_{(n-k) \times (m-k)} \end{bmatrix}$$

Thm: For a given ~~matrix~~ $m \times n$ matrix \underline{A} with rank k , let \underline{P} and \underline{Q} be permutation matrices such that

$$\underline{P}\underline{A}\underline{Q} = \begin{bmatrix} \underline{C} & \underline{D} \\ \underline{E} & \underline{F} \end{bmatrix} \text{ with } \underline{C} \text{ nonsingular}$$

the generalized inverse of \underline{A} is given by

$$\underline{G} = \underline{S} \begin{bmatrix} \underline{C}^{-1} & \underline{O} \\ \underline{O} & \underline{O} \end{bmatrix} \underline{P}$$

Ex: $\underline{A} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & -1 \\ 1 & 0 & 1 & 2 \end{bmatrix}$. Find \underline{A}^{-}

clearly we have to find a 2×2 matrix \underline{C} with $\text{rank}(\underline{C}) = 2$ (since $\text{rank}(\underline{A}) = 2$)

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & -1 \\ 1 & 0 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & -1 \\ 1 & 0 & 1 & 2 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

$\underline{P} \qquad \underline{A} \qquad \underline{PA}$

$$\underline{PA} \underline{S} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & -1 & 0 & 1 \\ 1 & 2 & 1 & 0 \\ 1 & 1 & 1 & 0 \end{bmatrix}$$

\underline{S}

$$\underline{PA} \underline{S} = \begin{bmatrix} 0 & -1 & 0 & 1 \\ 1 & 2 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix} \qquad \underline{C} = \begin{bmatrix} 0 & -1 \\ 1 & 2 \end{bmatrix}$$

$$\underline{C}^{-1} = \begin{bmatrix} 2 & 1 \\ -1 & 0 \end{bmatrix}$$

$$\underline{A}^{-} = \underline{S} \begin{bmatrix} 2 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \underline{P}$$

~~Then~~ Remark: If \underline{A} is nonsingular then

\underline{A}^{-1} is a generalized inverse of \underline{A} .

$\underline{A}\underline{A}^{-1}\underline{A} = \underline{A}$. In fact, in this case it is the unique generalized inverse.

Let $(\underline{x}^T \underline{x})^-$ be the generalized inverse of $\underline{x}^T \underline{x}$.

Lemma: If \underline{G} and \underline{H} are generalized inverses of $\underline{x}^T \underline{x}$, then

$$\textcircled{1} \underline{x} \underline{G} \underline{x}^T \underline{x} = \underline{x} \underline{H} \underline{x}^T \underline{x} = \underline{x}$$

$$\textcircled{2} \underline{x} \underline{G} \underline{x}^T = \underline{x} \underline{H} \underline{x}^T$$

Note that $(\underline{x}^T \underline{x})^- \underline{x}^T$ is the generalized inverse of \underline{x} .

$$\underline{x} \left[(\underline{x}^T \underline{x})^- \underline{x}^T \right] \underline{x} = \underline{x} \quad (\text{use } \textcircled{1})$$

$\Rightarrow (\underline{x}^T \underline{x})^- \underline{x}^T$ is ~~the gen~~ a generalized inverse of \underline{x} .

Then: $\underline{P} = \underline{x} (\underline{x}^T \underline{x})^- \underline{x}^T$ is the projection matrix onto $C(\underline{x})$.

Hf: take $\underline{v} \in C(\underline{x}) \Rightarrow \underline{v} = \underline{x} \underline{a}$ for some \underline{a}

$$\underline{P} \underline{v} = \underline{x} (\underline{x}^T \underline{x})^- \underline{x}^T \underline{x} \underline{a} = \underline{x} \underline{a} \quad (\text{by } \textcircled{1}) \\ = \underline{v}$$

~~For~~ take $\underline{w} \in N(\underline{x}^T) \Rightarrow \underline{x}^T \underline{w} = \underline{0}$

$$\underline{P} \underline{w} = \underline{x} (\underline{x}^T \underline{x})^- \underline{x}^T \underline{w} = \underline{0}$$

clearly $\hat{y} = P y = X (X^T X)^{-1} X^T y$

when $(X^T X)$ has the full rank, then what is $X \hat{\beta}$?

$$\hat{\beta} = (X^T X)^{-1} X^T y \Rightarrow X \hat{\beta} = X (X^T X)^{-1} X^T y$$

$\text{rank}(X^T X) < p \Rightarrow$ there are infinitely many solutions to the NEs

$$X^T X \beta = X^T y$$

Thm: Let $A x = c$ be a system of equations and let G be a generalized inverse of A . Then \tilde{x} is a solution to the equation $A x = c$ if and only if there exists a vector z such that $\tilde{x} = G c + (I - G A) z$

In our case, $A = (X^T X)$, $c = X^T y$

⊙ class of all solutions of β

$$= \left\{ (X^T X)^{-1} X^T y + (I - (X^T X)^{-1} X^T X) z \mid \begin{array}{l} \text{for all } \\ (X^T X)^{-1} \\ \text{and for} \\ \text{all } z \end{array} \right\}$$

Result: If $\hat{\beta}_1$ and $\hat{\beta}_2$ are two solutions to the NEs,

$$\Rightarrow X^T X \hat{\beta}_1 = X^T y \quad \text{--- (1)}$$

$$X^T X \hat{\beta}_2 = X^T y \quad \text{--- (2)}$$

subtract ② from ① to get

$$\underline{X}^T \underline{X} (\hat{\underline{\beta}}_1 - \hat{\underline{\beta}}_2) = \underline{0}$$

$$(\hat{\underline{\beta}}_1 - \hat{\underline{\beta}}_2) \in \mathcal{N}(\underline{X}^T \underline{X}) = \mathcal{N}(\underline{X})$$

Thm: $\mathcal{N}(\underline{X}^T \underline{X}) = \mathcal{N}(\underline{X})$

pf: $\underline{c} \in \mathcal{N}(\underline{X}) \Rightarrow \underline{X} \underline{c} = \underline{0} \Rightarrow \underline{X}^T \underline{X} \underline{c} = \underline{0}$
 $\Rightarrow \underline{c} \in \mathcal{N}(\underline{X}^T \underline{X})$

take $\underline{c} \in \mathcal{N}(\underline{X}^T \underline{X}) \Rightarrow \underline{X}^T \underline{X} \underline{c} = \underline{0} \Rightarrow \underline{c}^T \underline{X}^T \underline{X} \underline{c} = 0$

$$\Rightarrow \|\underline{X} \underline{c}\|^2 = 0 \Rightarrow \underline{X} \underline{c} = \underline{0} \Rightarrow \underline{c} \in \mathcal{N}(\underline{X})$$

Thm: If $C(\underline{W}) \subset C(\underline{X})$ and \underline{P}_X and \underline{P}_W are the projection matrices onto $C(\underline{X})$ and $C(\underline{W})$ respectively, then $\underline{P}_X - \underline{P}_W$ is the projection onto $C((\underline{I} - \underline{P}_W) \underline{X})$.

try to see $C((\underline{I} - \underline{P}_W) \underline{X})$ as the orthogonal complement of $C(\underline{W})$.

take $\underline{z} \in C(\underline{W}) \cap C((\underline{I} - \underline{P}_W) \underline{X})$

$$\Rightarrow \exists \underline{h}_1, \underline{h}_2 \text{ s.t. } \underline{z} = \underline{W} \underline{h}_1 \text{ and } \underline{z} = (\underline{I} - \underline{P}_W) \underline{X} \underline{h}_2$$

$$\begin{aligned} \Rightarrow \underline{z}' \underline{z} &= \underline{h}_1' \underline{W}' (\underline{I} - \underline{P}_W) \underline{X} \underline{h}_2 = \underline{h}_1' (\underline{W}' - \underline{W}' \underline{P}_W) \underline{X} \underline{h}_2 \\ &= \underline{h}_1' (\underline{W} - \underline{P}_W \underline{W})' \underline{X} \underline{h}_2 = 0 \end{aligned}$$

$$\Rightarrow \|\underline{z}\|^2 = 0 \Rightarrow \underline{z} = \underline{0}$$

Ex: $y_i = \beta_0 + \beta_1 x_i + e_i \quad i=1, \dots, 4$

$x_1=1, x_2=2, x_3=3, x_4=4$

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \end{bmatrix}$$

Model 1

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \gamma + \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \end{bmatrix}$$

Model 2

$$\underline{X} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix}, \quad \underline{W} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \Rightarrow C(\underline{W}) \subset C(\underline{X})$$

Find $\underline{P}_X = \underline{X} (\underline{X}^T \underline{X})^{-1} \underline{X}^T$

$$= \frac{1}{10} \begin{bmatrix} 7 & 4 & 1 & -2 \\ 4 & 3 & 2 & 1 \\ 1 & 2 & 3 & 4 \\ -2 & 1 & 4 & 7 \end{bmatrix}$$

$$\underline{P}_W = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \frac{1}{4} (1 \ 1 \ 1 \ 1) = \frac{1}{4} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

$$= \frac{1}{4} \underline{J}_{4 \times 4}$$

with the reduced model (Model 2)

What is the ~~prop~~ predicted response.

$$\underline{P}_W \underline{y} = \frac{1}{4} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = \begin{pmatrix} \frac{y_1+y_2+y_3+y_4}{4} \\ \frac{y_1+y_2+y_3+y_4}{4} \\ \frac{y_1+y_2+y_3+y_4}{4} \\ \frac{y_1+y_2+y_3+y_4}{4} \end{pmatrix}$$

(7)

$$P_X - P_W = \frac{1}{100} \begin{bmatrix} 45 & 15 & -15 & -45 \\ 15 & 5 & -5 & -15 \\ -15 & -5 & 5 & -15 \\ -45 & -15 & -15 & 45 \end{bmatrix}$$

this is the projection matrix onto $C((\underline{I} - \underline{P}_W)\underline{x})$.

$$\underline{y} = \underline{P}_W \underline{y} + (\underline{P}_X - \underline{P}_W) \underline{y} + (\underline{I} - \underline{P}_X) \underline{y}$$