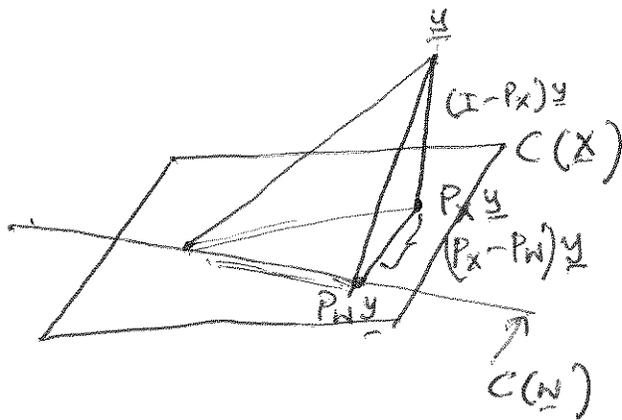


Recap: Projection matrix. How to construct the projection matrix for the $C(\underline{X})$. We have seen $P_{\underline{X}} = \underline{X}(\underline{X}^T \underline{X})^{-1} \underline{X}^T$ is the projection matrix for $C(\underline{X})$.

If $C(\underline{W}) \subset C(\underline{X})$ then $C(\underline{W})$ & $C((\underline{I} - P_{\underline{W}})\underline{X})$ are orthogonal complements.



let $n=3$

$$\underline{y} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

$C(\underline{X})$ is two dimensional

Two models are equivalent (or reparametrization of each other) if column spaces of the design matrices are the same.

Def: Two linear models $\underline{y} = \underline{X}\beta + \underline{e}$, where \underline{X} is an $n \times p$ matrix and $\underline{y} = \underline{W}\gamma + \underline{\tilde{e}}$ where \underline{W} is an $n \times t$ matrix are equivalent iff $C(\underline{X}) = C(\underline{W})$.

If $C(\underline{X}) = C(\underline{W})$

① $P_{\underline{X}} = P_{\underline{W}}$ ② $\hat{\underline{y}} = P_{\underline{X}} \underline{y} = P_{\underline{W}} \underline{y}$

③ $\hat{\underline{e}} = (\underline{I} - P_{\underline{X}}) \underline{y} = (\underline{I} - P_{\underline{W}}) \underline{y}$

Example: 1 way ANOVA

Model 1: $y_{ij} = \mu + \alpha_i + \epsilon_{ij}$, $i=1,2,3; j=1, \dots, n$

Model 2: $y_{ij} = \theta_j + \epsilon_{ij}$, $i=1,2,3; j=1, \dots, n$.

Write

Model 1 as

$$\underline{y} = \underline{X} \underline{\beta} + \underline{e}$$

$$\underline{X} = \begin{bmatrix} \underline{1}_n & \underline{1}_n & \underline{0}_n & \underline{0}_n \\ \underline{1}_n & \underline{0}_n & \underline{1}_n & \underline{0}_n \\ \underline{1}_n & \underline{0}_n & \underline{0}_n & \underline{1}_n \end{bmatrix}$$

$$\underline{\beta} = \begin{pmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix}$$

$$\begin{pmatrix} y_{11} \\ y_{12} \\ \vdots \\ y_{1n} \\ y_{21} \\ \vdots \\ y_{2n} \\ y_{31} \\ \vdots \\ y_{3n} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & 0 & 1 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & 0 & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix} \begin{pmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} + \underline{e}$$

Model 2:

$$\begin{pmatrix} y_{11} \\ \vdots \\ y_{1n} \\ y_{21} \\ \vdots \\ y_{2n} \\ y_{31} \\ \vdots \\ y_{3n} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ \vdots & \vdots & \vdots \\ \vdots & 0 & 0 \\ \vdots & \vdots & \vdots \\ \vdots & 0 & 0 \\ \vdots & \vdots & \vdots \\ \vdots & 0 & 0 \\ \vdots & \vdots & \vdots \\ \vdots & 0 & 0 \\ \vdots & \vdots & \vdots \end{pmatrix} \begin{pmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \end{pmatrix} + \underline{e}$$

$$\underline{y} = \underline{W} \underline{\gamma} + \underline{e} \Rightarrow \underline{W} = \begin{bmatrix} \underline{1}_n & \underline{0}_n & \underline{0}_n \\ \underline{0}_n & \underline{1}_n & \underline{0}_n \\ \underline{0}_n & \underline{0}_n & \underline{1}_n \end{bmatrix}, \quad \underline{\gamma} = \begin{pmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \end{pmatrix}$$

$C(\underline{W}) = C(\underline{X})$ Hence these two models are equivalent.

Ex: $i=1, 2; j=1, \dots, n$.

~~Model 1:~~ Model 1: $y_{ij} = \beta_0^{(i)} + \beta_1^{(i)} x_{ij} + e_{ij}$

Model 2: $d_{ij} = \begin{cases} 0 & \text{if } i=1 \text{ (group 1)} \\ 1 & \text{if } i=2 \text{ (group 2)} \end{cases}$

$y_{ij} = \delta_0 + \delta_1 x_{ij} + \delta_2 d_{ij} + \delta_3 d_{ij} x_{ij} + \tilde{e}_{ij}$

For group 1, $y_{ij} = \delta_0 + \delta_1 x_{ij} + \tilde{e}_{ij}$

For group 2, $y_{ij} = (\delta_0 + \delta_2) + (\delta_1 + \delta_3) x_{ij} + \tilde{e}_{ij}$

~~If~~ If \underline{X} is the predictor matrix (design matrix) from the first model,

$$\underline{X} = \begin{bmatrix} 1 & x_{11} & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & x_{1n} & 0 & 0 \\ 0 & 0 & 1 & x_{21} \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 1 & x_{2n} \end{bmatrix} \quad \underline{\beta} = \begin{pmatrix} \beta_0^{(1)} \\ \beta_1^{(1)} \\ \beta_0^{(2)} \\ \beta_1^{(2)} \end{pmatrix}$$

If \underline{W} is the predictor matrix from Model 2,

$$\underline{W} = \begin{bmatrix} 1 & x_{11} & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & 0 & 0 \\ \vdots & \vdots & 1 & x_{21} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_{2n} & 1 & x_{2n} \end{bmatrix}, \quad \underline{\delta} = \begin{pmatrix} \delta_0 \\ \delta_1 \\ \delta_2 \\ \delta_3 \end{pmatrix}$$

first column of \underline{W} = 1st column of \underline{X} + 3rd column of \underline{X}
 2nd column of \underline{W} = 2nd column of \underline{X} + 4th column of \underline{X}

$\Rightarrow C(\underline{W}) = C(\underline{X}) \Rightarrow$ These two models are reparametrizations of each other.

Recall NEs $\underline{X}^T \underline{X} \underline{\beta} = \underline{X}^T \underline{y}$.

① $K(\underline{X}) = p$ (full column rank)

$\Rightarrow \hat{\underline{\beta}}$ is unique. \Rightarrow Any function of $\underline{\beta}$ can be estimated

② When $K(\underline{X}) < p \Rightarrow$ Multiple solutions to the normal equations.

\Rightarrow We cannot estimate all functions of $\underline{\beta}$?

Q: Which functions can be or can't be estimated?

Identifiability: The parametrization $\underline{\beta}$ is identifiable

if for any $\underline{\beta}_1$ and $\underline{\beta}_2$, $\underline{X} \underline{\beta}_1 = \underline{X} \underline{\beta}_2$ implies

$$\underline{\beta}_1 = \underline{\beta}_2.$$

① Knowing $E[\underline{y}] = \underline{X} \underline{\beta}$ means knowing $\underline{\beta}$.

Also, a difference in the parameter values \Rightarrow difference in the means.

② Look at the one-way ANOVA.

$$y_{ij} = \mu + \alpha_i + \epsilon_{ij}, \quad i = 1, 2, \dots, K; \quad j = 1, \dots, n$$

$$\underline{\beta} = \begin{bmatrix} \mu \\ \alpha_1 \\ \vdots \\ \alpha_K \end{bmatrix} \quad \text{take } \underline{\beta}_1 = \begin{bmatrix} \mu + c \\ \alpha_1 - c \\ \vdots \\ \alpha_K - c \end{bmatrix}$$

clearly $\underline{\beta}$ and $\underline{\beta}_1$ lead to the same values of $E[y_{ij}] \Rightarrow$ Not identifiable.

Ex: Consider a linear regression $C(\underline{x}) = \underline{\beta}$
 If $\underline{x}\underline{\beta}_1 = \underline{x}\underline{\beta}_2$. Since $C(\underline{x}) = \underline{\beta} \Rightarrow \underline{x}^T \underline{x}$ is invertible.

$$\underline{\beta}_1 = (\underline{x}^T \underline{x})^{-1} \underline{x}^T \underline{x} \underline{\beta}_1 = (\underline{x}^T \underline{x})^{-1} \underline{x}^T \underline{x} \underline{\beta}_2 = \underline{\beta}_2$$

Def:- A vector-valued function $g(\underline{\beta})$ is identifiable if $\underline{x}\underline{\beta}_1 = \underline{x}\underline{\beta}_2$ implies $g(\underline{\beta}_1) = g(\underline{\beta}_2)$.

In specific focus on the linear function $g(\underline{\beta}) = \underline{\lambda}^T \underline{\beta}$. Thus $\underline{\lambda}^T \underline{\beta}$ is identifiable if

$$\underline{x}\underline{\beta}_1 = \underline{x}\underline{\beta}_2 \text{ implies } \underline{\lambda}^T \underline{\beta}_1 = \underline{\lambda}^T \underline{\beta}_2.$$

Thm: A function $g(\underline{\beta})$ is identifiable iff $g(\underline{\beta})$ is a function $\underline{\lambda}^T \underline{\beta}$.

Which $\underline{\lambda}^T \underline{\beta}$ are reasonable to estimate?

Identifiable functions.

Def: An estimator $t(\underline{y})$ is an unbiased estimator for the scalar $\underline{\lambda}^T \underline{\beta}$ iff $E[t(\underline{y})] = \underline{\lambda}^T \underline{\beta}$ for all $\underline{\beta}$.

An estimator $t(\underline{y})$ is a linear estimator in \underline{y} iff $t(\underline{y}) = c + \underline{a}^T \underline{y}$ for constants c, a_1, \dots, a_n .

$$\underline{a} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$$

A function $\underline{\lambda}^T \underline{\beta}$ is linearly estimable iff there exists a linear unbiased estimator for $\underline{\lambda}^T \underline{\beta}$.

$\lambda^T \beta$ is linearly estimable

$$\Rightarrow \exists \text{ some } \underline{a}^T \underline{y} \text{ s.t. } E[\underline{a}^T \underline{y}] = \underline{\lambda}^T \underline{\beta} \quad \text{--- (1)}$$

~~(1)~~ let's assume $E[\underline{e}] = \underline{0}$

$$E[\underline{y}] = \underline{X} \underline{\beta} \Rightarrow E[\underline{a}^T \underline{y}] = \underline{a}^T \underline{X} \underline{\beta}$$

according to (1),

$$\underline{a}^T \underline{X} \underline{\beta} = \underline{\lambda}^T \underline{\beta} \quad \text{--- } \underline{X}^T \underline{a} \neq \underline{\beta}$$

$$\Rightarrow \underline{a}^T \underline{X} = \underline{\lambda}^T \Rightarrow \underline{\lambda} = \underline{X}^T \underline{a} \Rightarrow \underline{\lambda} \in C(\underline{X}^T)$$

EX: Consider one way ANOVA,

$$i=1, 2, 3, \quad j=1, \dots, n_i, \quad n_1=3, n_2=2, n_3=1$$

$$\begin{bmatrix} y_{11} \\ y_{12} \\ y_{13} \\ y_{21} \\ y_{22} \\ y_{31} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} + \begin{pmatrix} e_{11} \\ \vdots \\ e_{31} \end{pmatrix}$$

⊙ Whether the following functions are linearly estimable

1. α_1
2. $\mu + \alpha_1$
3. $\alpha_1 - \alpha_3$
4. $\alpha_1 + \alpha_2 - 2\alpha_3$

⊙ For 2. $E[y_{11}] = E[y_{12}] = E[y_{13}] = \mu + \alpha_1$

Also, $E\left[\frac{y_{11} + y_{12} + y_{13}}{3}\right] = \mu + \alpha_1$

3. $E[y_{31}] = \mu + \alpha_3 \Rightarrow E\left[\frac{y_{11} + y_{12} + y_{13}}{3} - y_{31}\right] = \alpha_1 - \alpha_3$

back (1)

$$4. E\left[\frac{y_{11} + y_{12} + y_{13}}{3}\right] = \mu + \alpha_1$$

$$E\left[\frac{y_{21} + y_{22}}{2}\right] = \mu + \alpha_2$$

$$E[y_{31}] = \mu + \alpha_3$$

$$\Rightarrow E\left[\frac{y_{11} + y_{12} + y_{13}}{3} + \frac{y_{21} + y_{22}}{2} - 2y_{31}\right] = \alpha_1 + \alpha_2 - 2\alpha_3$$

$$1. X^T = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Let's find basis for the column space of \underline{X}^T .

$$\begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

Since $\kappa(\underline{X}^T) = \kappa(\underline{X}) = 3$, if the above three vectors are linearly independent then they form the basis of $C(\underline{X}^T)$.

$$\text{take } \lambda_1 \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} + \lambda_3 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\Rightarrow \lambda_3 = 0, \lambda_2 = 0, \lambda_1 = 0$$

$\Rightarrow \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}$ are linearly independent.

Hence they form a basis for $C(\underline{X}^T)$.

$$\alpha_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \begin{pmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix}$$

check if $\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \in C(\underline{X}^T)$. To have that

$$\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \lambda_1 \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} + \lambda_3 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

$$\Rightarrow \lambda_3 = 0, \lambda_2 = 0, \text{ ①}$$

$$\lambda_1 = 1, \text{ --- ②}$$

$$\lambda_1 + \lambda_2 + \lambda_3 = 0 \text{ --- ③}$$

① & ② can't be satisfied along with $\lambda_3 = 0, \lambda_2 = 0$.

$\Rightarrow \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \notin C(\underline{X}^T) \Rightarrow \alpha_1$ is not linearly estimable.

$\mathcal{N}(\underline{X})$ is orthogonal complement \perp of the $C(\underline{X}^T)$

Thus $\underline{\lambda} \in C(\underline{X}^T) \Leftrightarrow \underline{\lambda} \perp \mathcal{N}(\underline{X})$.

In our case $\dim(\mathcal{N}(\underline{X})) = 1$

\Rightarrow there is only one basis vector in the null space.

$$\underline{X} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}$$

$$\underline{v} \in \mathcal{N}(\underline{X}) \text{ iff } \underline{X}\underline{v} = \underline{0}$$

$$\begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\begin{aligned} \Rightarrow v_1 + v_2 &= 0 & \Rightarrow v_2 &= -v_1 \\ v_1 + v_3 &= 0 & v_4 &= -v_1 \\ v_1 + v_4 &= 0 & v_3 &= -v_1 \end{aligned}$$

$$\begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ -1 \\ -1 \end{pmatrix} v_1 \Rightarrow \begin{pmatrix} 1 \\ -1 \\ -1 \\ -1 \end{pmatrix} \text{ is a basis for}$$

$\mathcal{N}(\underline{X})$.

$$\alpha_1 = (0 \ 1 \ 0 \ 0) \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix}$$

For α_1 to be estimable $\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \perp \mathcal{N}(\underline{X})$

but inner product of $\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ -1 \\ -1 \\ -1 \end{pmatrix}$ is -1 .

Since the inner product is not zero

$$\Rightarrow \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \notin \mathcal{N}(\underline{X}).$$

back (4)