Recap: A function of parameters $\underline{\lambda}^T \underline{\beta}$ is linearly estimable if $\exists$ $t(\underline{y}) = c + \underline{a}^T \underline{y}$ s.t. $E[t(\underline{y})] = \underline{\lambda}^T \underline{\beta}$ $\forall$ $\underline{\beta}$.

$\Leftrightarrow$ $\underline{\lambda} \in C(\underline{X}^T)$

Ex!

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

whether 1. α_1 2. $\alpha_1 - \alpha_3$ are linearly estimable.

$$\alpha_1 = (0, 1, 0, 0) \begin{pmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix}$$ we checked if $\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \notin C(\underline{X}^T)$

$$\alpha_1 - \alpha_3 = (0, 1, 0, -1) \begin{pmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix}$$ is s.t. $\begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \end{pmatrix} \in C(\underline{X}^T)$

$$E[y_{11}] = \mu + \alpha_1 \quad ; \quad E[y_{31}] = \mu + \alpha_3$$

$$E[y_{11} - y_{31}] = \alpha_1 - \alpha_3 .$$

for linear estimability $\underline{\lambda} \in C(\underline{X}^T)$

What is the connection between $C(\underline{X}^T)$ and $\mathcal{N}(\underline{X})$.

We have seen that they are orthogonal complement to each other. $\Rightarrow$ $\underline{\lambda} \perp \mathcal{N}(\underline{X})$.

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

we know $\underline{v} \in \mathcal{N}(\underline{X})$

$\Rightarrow \underline{X}\,\underline{v} = \underline{0}$

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

$v_1 + v_2 = 0$

$v_1 + v_3 = 0$

$v_1 + v_4 = 0$

$\Rightarrow v_2 = v_3 = v_4 = -v_1$

$$\begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{pmatrix} = v_1 \begin{pmatrix} 1 \\ -1 \\ -1 \\ -1 \end{pmatrix}$$

$$\mathcal{N}(\underline{X}) = \left\{ v_1 \begin{pmatrix} 1 \\ -1 \\ -1 \\ -1 \end{pmatrix} : v_1 \in \mathbb{R} \right\}$$

$\underline{\lambda}^T \underline{\beta}$ is linearly estimable $\Leftrightarrow \underline{\lambda} \perp \mathcal{N}(\underline{X})$.

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

~~does~~ is $\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \not\perp \mathcal{N}(\underline{X})$? $\left\langle \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, v_1 \begin{pmatrix} 1 \\ -1 \\ -1 \\ -1 \end{pmatrix} \right\rangle = -v_1$

$\Rightarrow \alpha_1$ is not linearly estimable

$$\alpha_1 - \alpha_3 = (0, 1, 0, -1) \begin{pmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix}$$, is $\begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \end{pmatrix} \perp \mathcal{N}(\underline{X})$?

$\Rightarrow \alpha_1 - \alpha_3$ is linearly estimable.

If you are told to check if $\underline{\lambda}^T\underline{\beta}$ is linearly estimable:

Method 1: Check if $\underline{\lambda} \in C(\underline{X}^T)$.

Method 2: If you can immediately find some linear combination $\underline{a}^T\underline{y}$ of $\underline{y}$ s.t.

$E[\underline{a}^T\underline{y}] = \underline{\lambda}^T\underline{\beta}$. (e.g. $E[y_{11} - y_{31}] = \alpha_1 - \alpha_3$)

Method 3: check if $\underline{\lambda} \perp N(\underline{X})$

In fact in method 1, you might work with basis elements of $C(\underline{X}^T)$

In method 3, you can also find the basis of $N(\underline{X})$. If $\{\underline{h}_1, \underline{h}_2, \dots, \underline{h}_k\}$ is the set of basis vectors for the $N(\underline{X})$, all you have to do is to check if $\underline{\lambda} \perp \underline{h}_1, \dots, \underline{\lambda} \perp \underline{h}_k$.

any vector $\underline{h} \in N(\underline{X})$ can be written as

$$\underline{h} = c_1\underline{h}_1 + c_2\underline{h}_2 + \dots + c_k\underline{h}_k$$

$$\langle \underline{\lambda}, \underline{h} \rangle = \langle \underline{\lambda}, c_1\underline{h}_1 + \dots + c_k\underline{h}_k \rangle = \sum_{s=1}^{k} c_s \underbrace{\langle \underline{\lambda}, \underline{h}_s \rangle}_{= 0} = 0$$

if $\underline{\lambda} \perp \underline{h}_s \Rightarrow \underline{\lambda} \perp \underline{h}$.

Qn: How to construct a linear unbiased estimator of a linearly estimable fn. $\underline{\lambda}^T \underline{\beta}$?

Recall that $\underline{X}\hat{\underline{\beta}} = \underline{P}\underline{y}$, [where P is the projection matrix]

and $\hat{\underline{\beta}}$ is a solution to the NEs $\underline{X}^T\underline{X}\hat{\underline{\beta}} = \underline{X}^T\underline{y}$.

<u>claim</u>: If $\underline{\lambda}^T\underline{\beta}$ is linearly estimable, then $\underline{\lambda}^T\hat{\underline{\beta}}$ is the same for all solutions $\hat{\underline{\beta}}$ to the NEs and $E[\underline{\lambda}^T\hat{\underline{\beta}}] = \underline{\lambda}^T\underline{\beta}$.

<u>Pf</u>:- Note that $\underline{\lambda}^T\underline{\beta}$ is estimable $\Rightarrow \underline{\lambda} \in C(\underline{X}^T)$

$\Rightarrow \exists\ \underline{a}$ s.t. $\underline{\lambda} = \underline{X}^T\underline{a}$.

$\Rightarrow \underline{\lambda}^T\hat{\underline{\beta}} = (\underline{X}^T\underline{a})^T\hat{\underline{\beta}} = \underline{a}^T\underline{X}\hat{\underline{\beta}} = \underline{a}^T\underline{P}\underline{y}$ [for all solution $\hat{\underline{\beta}}$]

thus $\underline{\lambda}^T\hat{\underline{\beta}}$ is fixed even you use different solutions $\hat{\underline{\beta}}$ to the NEs.

$$E[\underline{\lambda}^T\hat{\underline{\beta}}] = E[\underline{a}^T\underline{P}\underline{y}] = \underline{a}^T\underline{P}\,E[\underline{y}] = \underline{a}^T\underline{P}\underline{X}\underline{\beta}$$

$$= \underline{a}^T\underline{X}\underline{\beta} = \underline{\lambda}^T\underline{\beta}.$$

$$\boxed{\begin{aligned}\underline{y} &= X\underline{\beta} + \underline{e}\\ E[\underline{y}] &= X\underline{\beta} + E[\underline{e}]\\ &= X\underline{\beta}\end{aligned}}$$

~~If you get estim~~

$$\underline{y} = \underline{X}\begin{pmatrix}\beta_0\\ \beta_1\\ \beta_2\end{pmatrix} + \underline{e}$$

you found estimates of $\begin{pmatrix}\beta_0\\ \beta_1\\ \beta_2\end{pmatrix}$ given by $\begin{pmatrix}\hat{\beta}_0\\ \hat{\beta}_1\\ \hat{\beta}_2\end{pmatrix}$

What is the estimate of $\beta_1 - \beta_2$?

You will write $\hat{\beta}_1 - \hat{\beta}_2$.

Now think of the NEs again,

$\underline{X}^T \underline{X} \underline{\beta} = \underline{X}^T \underline{y}$. When $r(\underline{X}) < p$ or $\dim(CC(\underline{X})) < p$ then the NEs have multiple solutions.

If we want to create a unique solution from the NEs we have to add some constraints on the parameter space.

Qn: ① which constraints to add?

② How many constraints to add?

Ex: Consider One way Anova.

$y_{ij} = \mu + \alpha_i + e_{ij}$, $i = 1, 2, 3$, $j = 1, 2$

$$\begin{bmatrix} y_{11} \\ y_{12} \\ y_{21} \\ y_{22} \\ y_{31} \\ y_{32} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} + \begin{bmatrix} e_{11} \\ e_{12} \\ e_{21} \\ e_{22} \\ e_{31} \\ e_{32} \end{bmatrix}$$

$\text{rank}(\underline{X}) = 3$ $\quad \underline{X}_{6\times 4}$ matrix, clearly $\text{rank}(\underline{X}) < 4$

$\underline{X}^T \underline{X} \underline{\beta} = \underline{X}^T \underline{y}$ has multiple solutions

$\underline{X}^T \underline{X} \underline{\beta} = \underline{X}^T \underline{y}$

$$\Rightarrow \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \begin{pmatrix} y_{11} \\ y_{12} \\ y_{21} \\ y_{22} \\ y_{31} \\ y_{32} \end{pmatrix}$$

$$\Rightarrow \begin{bmatrix} 6 & 2 & 2 & 2 \\ 2 & 2 & 0 & 0 \\ 2 & 0 & 2 & 0 \\ 2 & 0 & 0 & 2 \end{bmatrix} \begin{pmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = \begin{bmatrix} \sum_{i=1}^{3} \sum_{j=1}^{2} y_{ij} \\ \sum_{j=1}^{2} y_{1j} \\ \sum_{j=1}^{2} y_{2j} \\ \sum_{j=1}^{2} y_{3j} \end{bmatrix} \begin{matrix} \rightarrow \text{Sum of all } y\text{'s} \\ \rightarrow \text{Sum of } y\text{'s in the 1st gp.} \\ \rightarrow \text{Sum of } y\text{'s in the 2nd gp.} \\ \rightarrow \text{Sum of } y\text{'s in the 3rd gp.} \end{matrix}$$

$\sum_{i=1}^{3} \sum_{j=1}^{2} y_{ij} = y_{..}$, $\sum_{j=1}^{2} y_{1j} = y_{1.}$, $\sum_{j=1}^{2} y_{2j} = y_{2.}$,

$\sum_{j=1}^{2} y_{3j} = y_{3.}$

$$\begin{aligned} 6\mu + 2\alpha_1 + 2\alpha_2 + 2\alpha_3 &= y_{..} \\ 2\mu + 2\alpha_1 \qquad\qquad\qquad &= y_{1.} \\ 2\mu + \qquad + 2\alpha_2 \qquad\quad &= y_{2.} \\ 2\mu + \qquad\qquad\quad + 2\alpha_3 &= y_{3.} \end{aligned}$$

3 independent equations and 4 parameters

$\Rightarrow$ the NEs do not have a unique soln.

Lets put the constraint

$$\alpha_1 + \alpha_2 + \alpha_3 = 0$$

$$2\mu + 2\alpha_1 = y_{1.}$$

$$2\mu + 2\alpha_2 = y_{2.}$$

$$2\mu + 2\alpha_3 = y_{3.}$$

$$\alpha_1 + \alpha_2 + \alpha_3 = 0$$

$$\Rightarrow \mu = \frac{y_{..}}{6}, \quad \alpha_1 = \frac{y_{1.}}{2} - \frac{y_{..}}{6}, \quad \alpha_2 = \frac{y_{2.}}{2} - \frac{y_{..}}{6},$$

$$\alpha_3 = \frac{y_{3.}}{2} - \frac{y_{..}}{6}$$

Imposing restrictions

~~We use~~ To get unique solutions of the parameters we use a set of constraints $\underline{C}\,\underline{\beta} = \underline{0}$, where $\underline{C}$ is a $s \times p$ matrix with $s = p - r$ (where rank $(\underline{X}) = r$) and $C([\underline{X}^T : \underline{C}^T]) = \mathbb{R}^p$.

$$\begin{aligned} \underline{X}^T \underline{X}\, \underline{\beta} &= \underline{X}^T \underline{y} \\ \underline{C}\, \underline{\beta} &= \underline{0} \end{aligned} \quad \Rightarrow \quad \begin{pmatrix} \underline{X}^T \underline{X} \\ \underline{C} \end{pmatrix} \underline{\beta} = \begin{pmatrix} \underline{X}^T \underline{y} \\ \underline{0} \end{pmatrix}$$

We ~~want~~ have added this extra equations so as to make $\begin{pmatrix} \underline{X}^T \underline{X} \\ \underline{C} \end{pmatrix}$ full column rank.

Qn: How to choose $\underline{C}$?

~~Now $C(\underline{X})$ in this case is not~~ ~~$\mathbb{R}^p$ although $\underline{X}$ is a $n\times p$ matrix.~~

Note that in this case $\underline{X}$ does not have the full column rank (that is why NEs are not solvable uniquely).

and so not all linear combination $\underline{\lambda}^T\underline{\beta}$ is estimable linearly.

there exists many $\underline{\lambda} \notin C(\underline{X}^T)$.

$\underline{C}^T$ cannot be orthogonal to the $\mathcal{N}(\underline{X})$.

$\Leftrightarrow$ the components of $\underline{C}\,\underline{\beta}$ cannot be estimable linearly.

In other words, identify $\mathcal{N}(\underline{X})$ and find basis vectors from $\mathcal{N}(\underline{X})$ or identify vectors which do not become orthogonal to $\mathcal{N}(\underline{X})$. Those vectors will help us constructing $\underline{C}\,\underline{\beta}$.

Think of the one way anova example.

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

basis for $C(\underline{X}^T) = \left\{ \begin{pmatrix}1\\1\\0\\0\end{pmatrix}, \begin{pmatrix}1\\0\\1\\0\end{pmatrix}, \begin{pmatrix}1\\0\\0\\1\end{pmatrix} \right\}$.

$\alpha_1 + \alpha_2 + \alpha_3 = 0$

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

$$\begin{pmatrix}0\\1\\1\\1\end{pmatrix} \in C(\underline{X}^T)$$

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

$\Rightarrow \lambda_1 = 1, \quad \lambda_2 = 1, \quad \lambda_3 = 1, \qquad \lambda_1 + \lambda_2 + \lambda_3 = 0$

they are not consistent, hence $\begin{pmatrix}0\\1\\1\\1\end{pmatrix} \notin C(\underline{X}^T)$.

$\underline{v} \in N(\underline{X}) \Rightarrow \underline{X}\,\underline{v} = \underline{0}$

$$\begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{bmatrix} = \begin{bmatrix} 0\\0\\0\\0\\0\\0 \end{bmatrix} \Rightarrow \begin{array}{l} v_1 + v_2 = 0 \\ v_1 + v_3 = 0 \\ v_1 + v_4 = 0 \end{array}$$

$v_2 = v_3 = v_4 = -v_1$

$$\mathcal{N}(\underline{X}) = \left\{ v_1 \begin{pmatrix} 1 \\ -1 \\ -1 \\ -1 \end{pmatrix} : v_1 \in \mathbb{R} \right\}$$

$$\begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix} \not\in \mathcal{N}(\underline{X})$$

the constraint $\alpha_1 + \alpha_2 + \alpha_3 = 0$ is the right constraint.

What is $\dim(\mathcal{N}(\underline{X})) = 1$

thus only one constraint is enough.

① How many constraints are required?

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

② Which constraints can we use?

If $\dim(\mathcal{N}(\underline{X})) = H$,

we will use constraints

$\underline{\lambda}_1^T \underline{\beta} = \underline{0}, \ldots, \underline{\lambda}_H^T \underline{\beta} = \underline{0}$

s.t. $\underline{\lambda}_1, \ldots, \underline{\lambda}_H \not\in \mathcal{N}(\underline{X})$

and $\underline{\lambda}_1, \ldots, \underline{\lambda}_H$ are linearly independent.

Jacob ①

Recap:

In the last class we have seen how to add constraints in the NEs.

$\underline{X}^T\underline{X}\,\underline{\beta} = \underline{X}^T\underline{y}$ are the NEs and when $\dim(\mathcal{C}(\underline{X}^T)) < p$ then we can add some constraints $\underline{C}\,\underline{\beta} = \underline{0}$ so that

$\begin{pmatrix} \underline{X}^T\underline{X} \\ \underline{C} \end{pmatrix}$ has the full column rank.

① $\dim(\mathcal{N}(\underline{X})) = r$ then find a few vectors $\underline{c}_1, \ldots, \underline{c}_r$ s.t. they are not orthogonal to the $\mathcal{N}(\underline{X}) \Leftrightarrow$ they are not in the column space of $\underline{X}^T$ and they are linearly independent.

Ex: Two way Anova.

$y_{ij} = \mu + \alpha_i + \beta_j + e_{ij}$, $\quad i = 1, 2;\ j = 1, 2, 3.$

$$\begin{pmatrix} y_{11} \\ y_{12} \\ y_{13} \\ y_{21} \\ y_{22} \\ y_{23} \end{pmatrix} = \underbrace{\begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}}_{\underline{X}} \begin{pmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} + \begin{pmatrix} e_{11} \\ e_{12} \\ e_{13} \\ e_{21} \\ e_{22} \\ e_{23} \end{pmatrix}$$

$$\underline{X}^T\underline{X}\,\underline{\beta} = \underline{X}^T\underline{y}$$

$$\begin{bmatrix} 6 & 3 & 3 & 2 & 2 & 2 \\ 3 & 3 & 0 & 1 & 1 & 1 \\ 3 & 0 & 3 & 1 & 1 & 1 \\ 2 & 1 & 1 & 2 & 0 & 0 \\ 2 & 1 & 1 & 0 & 2 & 0 \\ 2 & 1 & 1 & 0 & 0 & 2 \end{bmatrix} \begin{pmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} y_{11} \\ y_{12} \\ y_{13} \\ y_{21} \\ y_{22} \\ y_{23} \end{pmatrix}$$

$$\sum_{i=1}^{2}\sum_{j=1}^{3} y_{ij} = y_{..}\,, \quad y_{i.} = \sum_{j=1}^{3} y_{ij}\,, \quad y_{.j} = \sum_{i=1}^{2} y_{ij}$$

$$\begin{bmatrix} 6 & 3 & 3 & 2 & 2 & 2 \\ 3 & 3 & 0 & 1 & 1 & 1 \\ 3 & 0 & 3 & 1 & 1 & 1 \\ 2 & 1 & 1 & 2 & 0 & 0 \\ 2 & 1 & 1 & 0 & 2 & 0 \\ 2 & 1 & 1 & 0 & 0 & 2 \end{bmatrix} \begin{pmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} = \begin{bmatrix} y_{..} \\ y_{1.} \\ y_{2.} \\ y_{.1} \\ y_{.2} \\ y_{.3} \end{bmatrix}$$

$$6\mu + 3\alpha_1 + 3\alpha_2 + 2\beta_1 + 2\beta_2 + 2\beta_3 = y_{..}$$
$$3\mu + 3\alpha_1 + \qquad + \beta_1 + \beta_2 + \beta_3 = y_{1.}$$
$$3\mu + \qquad + 3\alpha_2 + \beta_1 + \beta_2 + \beta_3 = y_{2.}$$
$$2\mu + \alpha_1 + \alpha_2 + 2\beta_1 = y_{.1}$$
$$2\mu + \alpha_1 + \alpha_2 + 2\beta_2 = y_{.2}$$
$$2\mu + \alpha_1 + \alpha_2 + 2\beta_3 = y_{.3}$$

there are 4 equations. To estimate them uniquely we need 2 more equations. Two ~~more~~ constraints needed to be added.

Try to find $\mathcal{N}(\underline{X})$.

$\underline{\gamma} \in \mathcal{N}(\underline{X}) \Rightarrow \underline{X}\,\underline{\gamma} = \underline{0}$

$$\begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \\ \gamma_4 \\ \gamma_5 \\ \gamma_6 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$\gamma_1 + \gamma_2 + \gamma_4 = 0$ — ①

$\gamma_1 + \gamma_2 + \gamma_5 = 0$ — ②

$\gamma_1 + \gamma_2 + \gamma_6 = 0$ — ③

$\gamma_1 + \gamma_3 + \gamma_4 = 0$ — ④

$\gamma_1 + \gamma_3 + \gamma_5 = 0$ — ⑤

$\gamma_1 + \gamma_3 + \gamma_6 = 0$ — ⑥

$\gamma_4 = \gamma_5 = \gamma_6$ [from ①, ②, ③]

using ④ & ①

$\gamma_2 = \gamma_3$

~~② & ⑤ gives~~

~~$\gamma_1 = -(\gamma_2 + \gamma_4)$~~

$\gamma_1 = -\gamma_2 - \gamma_4$

$$\begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \\ \gamma_4 \\ \gamma_5 \\ \gamma_6 \end{pmatrix} = \begin{pmatrix} -\gamma_2 - \gamma_4 \\ \gamma_2 \\ \gamma_2 \\ \gamma_4 \\ \gamma_4 \\ \gamma_4 \end{pmatrix}$$

$$\mathcal{N}(\underline{X}) = \left\{ \begin{pmatrix} 1 \\ -1 \\ -1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ -1 \\ -1 \\ -1 \end{pmatrix} \right\}$$

$\dim(\mathcal{N}(\underline{X})) = 2$

$\Rightarrow$ we need two constraints.

$c_1^T \beta = 0$ and $\underline{c}_2^T \underline{\beta} = 0$

are the two constraints then $\underline{c}_1, \underline{c}_2 \not\in \mathcal{N}(\underline{X})$

$$\underline{c}_1^T = (0, 1, 1, 0, 0, 0) \qquad \underline{c}_1 \not\perp \begin{pmatrix} 1 \\ -1 \\ -1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\underline{c}_1 \not\perp \mathcal{N}(\underline{X})$$

$$\underline{c}_2^T = (0, 0, 0, 1, 1, 1) \quad , \quad \underline{c}_2 \not\perp \begin{pmatrix} 1 \\ 0 \\ 0 \\ -1 \\ -1 \\ -1 \end{pmatrix}$$

$$\Rightarrow \underline{c}_2 \not\perp \mathcal{N}(\underline{X})$$

hence,
$$c_1^T \begin{pmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} = 0 \quad \Rightarrow \quad \alpha_1 + \alpha_2 = 0$$

$$c_2^T \begin{pmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} = 0 \quad \Rightarrow \quad \beta_1 + \beta_2 + \beta_3 = 0 .$$

$$\underline{C}\,\underline{\beta} = \underline{0} \quad \Leftrightarrow \quad \begin{pmatrix} 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Till now we have used any distributional assumption on the error of the linear model $E[e_{ij}] = 0$. Now we will make distributional assumptions on the error.

Def: Consider an $n \times n$ matrix $\underline{A}$, the

$$\text{trace}(\underline{A}) = \sum_{i=1}^{n} a_{ii} = \text{sum of the diagonal elements}$$

Result: (i) $tr(\underline{A}\,\underline{B}) = tr(\underline{B}\,\underline{A})$

(ii) $tr(\underline{A}^T \underline{A}) = \sum_{i,j} a_{ij}^2$ = sum of square of all elements of $\underline{A}$.

Def: The determinant of an $n \times n$ matrix $\underline{A}$ is a scalar given by

$$|\underline{A}| = \sum_{i=1}^{n} a_{i1} (-1)^{i+1} |M_{i1}|$$

M_{ij} is the determinant of the (i,j)th minor of the matrix.

$$\underline{A} = \begin{pmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & & & & \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nn} \end{pmatrix}$$

$$M_{ij} = \begin{pmatrix} a_{11} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1n} \\ \vdots & & & & & \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & & & & & \vdots \\ a_{n1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{nn} \end{pmatrix}_{(n-1)\times(n-1)}$$

$$\underline{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

$$|\underline{A}| = a_{11}(-1)^{1+1} a_{22} + a_{12}(-1)^{1+2} a_{21} + a_{21}(-1)^{1+2} a_{12} + a_{22}(-1)^{2+2} a_{11}$$

$$|\underline{A}| = a_{11}a_{22} - a_{12}a_{21}$$

$$\underline{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

$$|\underline{A}| = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} + a_{12}(-1)^{1+2} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}(-1)^{1+3} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

$$= a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{31}a_{23}) + a_{13}(a_{21}a_{32} - a_{22}a_{31})$$

$\underline{D}$ is a diagonal matrix $\underline{D} = \begin{pmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{pmatrix}$

$$|\underline{D}| = \prod_{i=1}^{n} d_i .$$

Results: $|\underline{A}\,\underline{B}| = |\underline{B}\,\underline{A}|$, $|\underline{A}^{-1}| = \dfrac{1}{|\underline{A}|}$

and $|c \underset{n\times n}{\underline{A}}| = c^n |\underline{A}|$

the following are equivalent:

(1) $\underline{A}\underline{x} = \underline{0}$ implies $\underline{x} = \underline{0}$ (i.e. rank$(\underline{A}) = n$)

(2) $|\underline{A}| \neq 0$ (3) $\underline{A}$ is invertible.

Def: Let $\underline{A}$ be an $n \times n$ matrix and let $\underline{A}$ be symmetric. Consider an equation

$$\underline{A}\,\underline{q} = \lambda \underline{q}.$$

Here λ is called an eigenvalue of $\underline{A}$ and $\underline{q}$ is called the associated eigenvector.

There are n eigenvalues for an $n \times n$ matrix, they can repeat.

For an $n \times n$ matrix

(1) there are n eigenvalues (though they can repeat)

(2) eigenvectors associated with eigenvalues are orthogonal to each other.

(3) eigenvectors can be normalized i.e. $\underline{q}$ can be chosen s.t. $\underline{q}^T\underline{q} = 1$.

eigenvalues of $\underline{A}$ are solutions to the equations $|\underline{A} - \lambda \underline{I}| = 0$

$$\underline{A} = \begin{bmatrix} 5 & 8 \\ 2 & 3 \end{bmatrix} \Rightarrow A - \lambda \underline{I} = \begin{bmatrix} 5-\lambda & 8 \\ 2 & 3-\lambda \end{bmatrix}$$

$$|A - \lambda \underline{I}| = \begin{vmatrix} 5-\lambda & 8 \\ 2 & 3-\lambda \end{vmatrix}$$

$$= (5-\lambda)(3-\lambda) - 16$$

two eigenvalues can be found by solving

$$(5-\lambda)(3-\lambda) - 16 = 0 \Rightarrow \lambda = \frac{8 \pm \sqrt{64+4}}{2}$$

$$\underline{A} = \begin{bmatrix} 10 & 3 & 2 \\ 3 & 9 & 3 \\ 2 & 3 & 10 \end{bmatrix}$$

$$|A - \lambda \underline{I}| = 0 \Rightarrow \begin{vmatrix} 10-\lambda & 3 & 2 \\ 3 & 9-\lambda & 3 \\ 2 & 3 & 10-\lambda \end{vmatrix} = 0$$

$\Rightarrow$ $\lambda_1 = 15$, $\lambda_2 = 8$, $\lambda_3 = 6$

eigenvectors are obtained by writing

$A \underline{v} = 15 \underline{v}$ [$\underline{v}$ here will give eigenvector corresponding to the eigenvalue 15]

If $\underline{v}_1$ is a soln.

$A \underline{v}_1 = 15 \underline{v}_1 \Rightarrow \underline{A} (c \underline{v}_1) = 15 (c \underline{v}_1)$

then $c \underline{v}_1$ is also a soln. for any c

Rather you get a space of solutions and you can use any of the solution as an eigenvector corresponding to that eigenvalue

$$\underline{q}_1 = \begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}, \quad \underline{q}_2 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ -\frac{1}{\sqrt{2}} \end{bmatrix}, \quad \underline{q}_3 = \begin{bmatrix} \frac{1}{\sqrt{6}} \\ -\frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \end{bmatrix}$$

$$\underline{q}_1^T \underline{q}_1 = 1 \qquad \underline{q}_2^T \underline{q}_2 = 1 \qquad \underline{q}_3^T \underline{q}_3 = 1$$

What can be seen is if the eigenvectors are normalized, then

$$\underline{A} = \underline{Q} \, \underline{\Lambda} \, \underline{Q}^T \quad \text{where} \quad \underline{Q} = [\underline{q}_1 : \underline{q}_2 : \underline{q}_3]$$

$$\underline{\Lambda} = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$$

This is called the spectral decomposition of a ~~matrix~~ symmetric matrix.

If $\underline{A}$ is an $n \times n$ matrix with eigenvalues $\lambda_1, \dots, \lambda_n$ and the corresponding eigenvectors $\underline{q}_1, \dots, \underline{q}_n$ then

$$\underline{A} = \underline{Q} \, \Lambda \, \underline{Q}^T \quad \text{where} \quad \underline{Q} = [\underline{q}_1 : \cdots : \underline{q}_n]$$

$$\underline{\Lambda} = \begin{pmatrix} \lambda_1 & & \underline{0} \\ & \ddots & \\ \underline{0} & & \lambda_n \end{pmatrix}$$

$$\underline{A} = \underline{Q}\,\Lambda\,\underline{Q}^T = [\underline{q}_1 : \cdots : \underline{q}_n] \begin{bmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{bmatrix} \begin{bmatrix} \underline{q}_1^T \\ \vdots \\ \underline{q}_n^T\end{bmatrix}$$

$$= \sum_{i=1}^{n} \lambda_i\, \underline{q}_i\, \underline{q}_i^T$$

$$[\underline{q}_1 : \cdots : \underline{q}_n] \begin{bmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{bmatrix} = [\underline{q}_1 \lambda_1 : \cdots : \underline{q}_n \lambda_n]$$

(1) $|\underline{A}| = \prod_{i=1}^{n} \lambda_i$ (2) $\mathrm{tr}(\underline{A}) = \sum_{i=1}^{n} \lambda_i$

(2) $\mathrm{rank}(\underline{A})$ = the number of nonzero eigenvalues.

Since each $\underline{q}_j^T \underline{q}_j = 1$ and $\underline{q}_j^T \underline{q}_{j'} = 0$ for $j \neq j'$

$\Rightarrow \underline{Q}\,\underline{Q}^T = \underline{I}_n = \underline{Q}^T \underline{Q}$ . .

$\Rightarrow \underline{Q}^T = \underline{Q}^{-1}$

when all eigenvalues are nonzero

$\underline{A}^{-1} = \underline{Q}\,\Lambda^{-1}\underline{Q}^T$ as $(\underline{Q}\,\underline{\Lambda}^{-1}\underline{Q}^T)(\underline{Q}\,\Lambda\,\underline{Q}^T)$

$= \underline{Q}\,\underline{Q}^T = \underline{I}_n$

$\Rightarrow \underline{A}^{-1} = \underline{Q}\,\underline{\Lambda}^{-1}\underline{Q}^T$

<u>Thm</u>: If $\underline{A}$ is a symmetric matrix then $\exists$ an orthonormal basis of $C(\underline{A})$ consisting of eigenvectors corresponding to the nonzero eigenvalues.

eigenvectors corresponding zero eigenvalues are the basis for $\mathcal{N}(\underline{A}^T) = \mathcal{N}(\underline{A})$

When $\underline{A}^{-1}$ doesn't exist $\Rightarrow$ some eigenvalues are zero. In that case

$$\underline{A} = \underline{Q} \begin{bmatrix} \underline{\Lambda}_r & \underline{0} \\ \underline{0} & \underline{0} \end{bmatrix} \underline{Q}^T$$

where $\underline{\Lambda}_r$ consisting in a diagonal matrix of the r nonzero eigenvalues in the diagonal.

$$\underline{A} = \Big[\underbrace{\underline{Q}_1}_{n \times r} : \underbrace{\underline{Q}_2}_{n \times (n-r)}\Big] \begin{bmatrix} \underline{\Lambda}_r & \underline{0} \\ \underline{0} & \underline{0} \end{bmatrix} \begin{bmatrix} \underline{Q}_1^T \\ \underline{Q}_2^T \end{bmatrix}$$

then a generalized is given by $A^- = \underline{Q}_1 \Lambda_r^{-1} \underline{Q}_1^T$