9.1. Finding Least-Squares Estimates if $\mathrm{r}(\mathbf{X}) < p$

If $\mathbf{X}_{n\times p}$ has rank $r < p$, there is not a unique solution $\hat{\boldsymbol{\beta}}$ to the normal equations. We have 3 ways to find *a* solution $\hat{\boldsymbol{\beta}}$ and *the* orthogonal projection $\hat{\mathbf{Y}}$:

1. *Use a generalized inverse* $(\mathbf{X}'\mathbf{X})^{-}$.
2. *Reduce the model to one of full rank.*
3. *Impose identifiability constraints.*

9.2. METHOD 1: Use a generalized inverse $(\mathbf{X}'\mathbf{X})^-$.

Definition: For $\mathbf{A}_{m\times n}$, a *generalized inverse (g-inv)* of $\mathbf{A}$ is an $n \times m$ matrix $\mathbf{A}^-$ satisfying

$$\mathbf{A}\mathbf{A}^-\mathbf{A} = \mathbf{A}.$$

Definition: $\mathbf{A}^-$ is a *reflexive g-inv* if

$$\mathbf{A}\mathbf{A}^-\mathbf{A} = \mathbf{A}$$

and, in addition,

$$\mathbf{A}^-\mathbf{A}\mathbf{A}^- = \mathbf{A}^-.$$

Lemma 9.2.1a: $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^-\mathbf{X}'\mathbf{Y}$ is a solution to the normal equations $\mathbf{X}'\mathbf{Y} = \mathbf{X}'\mathbf{X}\boldsymbol{\beta}$.

Proof: Rewrite the normal equations as

$$\begin{aligned}
\mathbf{X}'\mathbf{Y} &= \mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} \\
&= \overbrace{\mathbf{X}'\mathbf{X}(\mathbf{X}'\mathbf{X})^-\mathbf{X}'\mathbf{X}}^{=\mathbf{X}'\mathbf{X}}\hat{\boldsymbol{\beta}} \\
&= \mathbf{X}'\mathbf{X}(\mathbf{X}'\mathbf{X})^-\underbrace{\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}}} \\
&= \mathbf{X}'\mathbf{X}(\mathbf{X}'\mathbf{X})^-\mathbf{X}'\mathbf{Y} \quad \text{(from 1st line)}
\end{aligned}$$

Therefore, $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^-\mathbf{X}'\mathbf{Y}$ is a solution to the normal equations.

Lemma 9.2.1b: The orthogonal projection $\hat{\mathbf{Y}}$ of $\mathbf{Y}$ onto $\mathcal{R}(\mathbf{X})$ is given by $\mathbf{PY}$, where

$$\mathbf{P} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'$$

for any generalized inverse $(\mathbf{X}'\mathbf{X})^{-}$.

Proof: We need to build up to establishing this result.

Lemma 9.2.2: If $\mathbf{PX}'\mathbf{X} = \mathbf{QX}'\mathbf{X}$ then $\mathbf{PX}' = \mathbf{QX}'$.

Proof: For any matrix $\mathbf{A}$, if $\mathbf{AA}' = \mathbf{0}$, then $\mathbf{A} = \mathbf{0}$ (easy to check).

$$(\mathbf{PX}' - \mathbf{QX}')(\mathbf{PX}' - \mathbf{QX}')' = (\mathbf{PX}' - \mathbf{QX}')\mathbf{X}(\mathbf{P} - \mathbf{Q})'$$

$$= (\mathbf{PX}'\mathbf{X} - \mathbf{QX}'\mathbf{X})(\mathbf{P} - \mathbf{Q})' = \mathbf{0}(\mathbf{P} - \mathbf{Q})' = \mathbf{0},$$

And so $\mathbf{PX}' - \mathbf{QX}' = \mathbf{0}$.

Theorem (Searle 7.1, 1987) If $\mathbf{G}$ is a g-inv of $\mathbf{X}'\mathbf{X}$, then

1. $\mathbf{G}'$ is also a *g-inv* of $\mathbf{X}'\mathbf{X}$

   $$(\mathbf{X}'\mathbf{X}\mathbf{G}'\mathbf{X}'\mathbf{X})' = \mathbf{X}'\mathbf{X}\mathbf{G}\mathbf{X}'\mathbf{X} = \mathbf{X}'\mathbf{X}$$

   Now take the transpose of both sides.

2. $\mathbf{G}\mathbf{X}'\mathbf{X}\mathbf{G}'$ is a symmetric reflexive g-inv of $\mathbf{X}'\mathbf{X}$

   Symmetry is clear. To show $\mathbf{G}\mathbf{X}'\mathbf{X}\mathbf{G}'$ is a g-inv of $\mathbf{X}'\mathbf{X}$:

   $\mathbf{X}'\mathbf{X}(\mathbf{G}\mathbf{X}'\mathbf{X}\mathbf{G}')\mathbf{X}'\mathbf{X} = (\mathbf{X}'\mathbf{X}\mathbf{G}\mathbf{X}'\mathbf{X})\mathbf{G}'\mathbf{X}'\mathbf{X} = (\mathbf{X}'\mathbf{X})\mathbf{G}'\mathbf{X}'\mathbf{X}$

   which equals $\mathbf{X}'\mathbf{X}$ using (1.) above.

   To show $\mathbf{X}'\mathbf{X}$ is a g-inv of $\mathbf{G}\mathbf{X}'\mathbf{X}\mathbf{G}'$:

   $$\mathbf{G}\mathbf{X}'\mathbf{X}\mathbf{G}'(\mathbf{X}'\mathbf{X})\mathbf{G}\mathbf{X}'\mathbf{X}\mathbf{G}' =$$

   $$\mathbf{G}\mathbf{X}'\mathbf{X}\mathbf{G}'(\mathbf{X}'\mathbf{X}\mathbf{G}\mathbf{X}'\mathbf{X})\mathbf{G}'$$
   $$= \mathbf{G}\mathbf{X}'\mathbf{X}\mathbf{G}'(\mathbf{X}'\mathbf{X})\mathbf{G}' = \mathbf{G}(\mathbf{X}'\mathbf{X}\mathbf{G}'\mathbf{X}'\mathbf{X})\mathbf{G}'$$

   which equals $\mathbf{G}\mathbf{X}'\mathbf{X}\mathbf{G}'$ using (1.) above again.

3. (a) $\mathbf{X}'\mathbf{X}\mathbf{G}\mathbf{X}' = \mathbf{X}'$ and (b) $\mathbf{X}\mathbf{G}\mathbf{X}'\mathbf{X} = \mathbf{X}$

   $\mathbf{X}'\mathbf{X}\mathbf{G}\mathbf{X}'\mathbf{X} = \mathbf{I}\mathbf{X}'\mathbf{X}$, so $\mathbf{X}'\mathbf{X}\mathbf{G}\mathbf{X}' = \mathbf{I}\mathbf{X}'$ by Lemma 9.2.2, proving (a). To prove (b), combine (a) and (1.) above: $\mathbf{X}'\mathbf{X}\mathbf{G}'\mathbf{X}' = \mathbf{X}'$, now take transposes.

4. $\mathbf{X}\mathbf{G}\mathbf{X}' = \mathbf{X}\mathbf{H}\mathbf{X}'$ for any other *g-inv* $\mathbf{H}$ (KEY RESULT!)

   By (3.), $\mathbf{X}\mathbf{G}\mathbf{X}'\mathbf{X} = \mathbf{X} = \mathbf{X}\mathbf{H}\mathbf{X}'\mathbf{X}$. Applying Lemma 9.2.2 gives the result.

5. $\mathbf{X}\mathbf{G}\mathbf{X}'$ is symmetric

   $(\mathbf{X}\mathbf{G}\mathbf{X}')' = \mathbf{X}\mathbf{G}'\mathbf{X}'$. From (1.), $\mathbf{G}'$ is also a g-inv. Therefore, by (4.), $\mathbf{X}\mathbf{G}'\mathbf{X}' = \mathbf{X}\mathbf{G}\mathbf{X}'$.

9.3. Computing a generalized inverse of $\mathbf{X}'\mathbf{X}$.

Let $\mathbf{X} = (\mathbf{X_1}, \mathbf{X_2})$, where $\mathbf{X}_1$ consists of r linearly independent columns from $\mathbf{X}$. Then a generalized inverse of $\mathbf{X}'\mathbf{X}$ is

$$(\mathbf{X}'\mathbf{X})^{-} = \begin{pmatrix} (\mathbf{X}_1'\mathbf{X}_1)^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}.$$

This result is a special case of the following Lemma.

Lemma 9.3.1: Let the matrix $\mathbf{W}_{p\times p}$ have rank r and be partitioned as

$$\mathbf{W} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix},$$

where $\mathbf{A}$ has full rank r. Then a generalized inverse of $\mathbf{W}$ is

$$\mathbf{W}^- = \begin{pmatrix} \mathbf{A}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}.$$

Proof:

$$\begin{aligned} \mathbf{W}\mathbf{W}^-\mathbf{W} &= \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} \begin{pmatrix} \mathbf{A}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{C}\mathbf{A}^{-1} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{C}\mathbf{A}^{-1}\mathbf{B} \end{pmatrix} \end{aligned}$$

This equals $\mathbf{W}$ if $\mathbf{D} = \mathbf{C}\mathbf{A}^{-1}\mathbf{B}$. Why does this have to be true? Since $\mathbf{A}$ has rank r (the same rank as $\mathbf{W}$), this means that any column of $\begin{pmatrix} \mathbf{B} \\ \mathbf{D} \end{pmatrix}$ can be written as a linear combination of the columns of $\begin{pmatrix} \mathbf{A} \\ \mathbf{C} \end{pmatrix}$. I.E., $\begin{pmatrix} \mathbf{B} \\ \mathbf{D} \end{pmatrix} = \begin{pmatrix} \mathbf{A} \\ \mathbf{C} \end{pmatrix}\mathbf{F}$, for some matrix $\mathbf{F}$. But then $\mathbf{B} = \mathbf{A}\mathbf{F}$, so $\mathbf{F} = \mathbf{A}^{-1}\mathbf{B}$. Since $\mathbf{D} = \mathbf{C}\mathbf{F}$, $\mathbf{D} = \mathbf{C}\mathbf{A}^{-1}\mathbf{B}$.

Also note that for the projection matrix onto $\mathcal{R}(\mathbf{X})$, we can always use

$$\mathbf{P} = \mathbf{X}_1(\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1'.$$

We know that such a $\mathbf{P}$ is a projection onto $\mathcal{R}(\mathbf{X}_1)$, but $\mathcal{R}(\mathbf{X}_1) = \mathcal{R}(\mathbf{X})$.

Example: One-way ANOVA with 2 groups

$$\begin{pmatrix} Y_{11} \\ \vdots \\ Y_{1n_1} \\ Y_{21} \\ \vdots \\ Y_{2n_2} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ \vdots & \vdots & \vdots \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ \vdots & \vdots & \vdots \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} \mu \\ \alpha_1 \\ \alpha_2 \end{pmatrix} + \begin{pmatrix} \varepsilon_{11} \\ \vdots \\ \varepsilon_{1n_1} \\ \varepsilon_{21} \\ \vdots \\ \varepsilon_{2n_2} \end{pmatrix}$$

(What is the rank of $\mathbf{X}$? 2)

$$\mathbf{X}'\mathbf{X} = \begin{pmatrix} n & n_1 & n_2 \\ n_1 & n_1 & 0 \\ n_2 & 0 & n_2 \end{pmatrix}.$$

Let $\mathbf{X}_1$ be the first two columns of $\mathbf{X}$. Then

$$(\mathbf{X}_1'\mathbf{X}_1)^{-1} = \begin{pmatrix} n & n_1 \\ n_1 & n_1 \end{pmatrix}^{-1} = \frac{1}{nn_1 - n_1^2} \begin{pmatrix} n_1 & -n_1 \\ -n_1 & n \end{pmatrix}$$

$$= \begin{pmatrix} \frac{1}{n_2} & \frac{-1}{n_2} \\ \frac{-1}{n_2} & \frac{1}{n_1} + \frac{1}{n_2} \end{pmatrix}$$

and a generalized inverse of $\mathbf{X}'\mathbf{X}$ is

$$(\mathbf{X}'\mathbf{X})^{-} = \begin{pmatrix} n_2^{-1} & -n_2^{-1} & 0 \\ -n_2^{-1} & n_1^{-1} + n_2^{-1} & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

What is the solution $\hat{\boldsymbol{\beta}}$ to the normal equations that corresponds to this generalized inverse?

$$\begin{aligned}\hat{\boldsymbol{\beta}} &= \begin{pmatrix} n_2^{-1} & -n_2^{-1} & 0 \\ -n_2^{-1} & n_1^{-1} + n_2^{-1} & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \sum_j Y_{1j} + \sum_j Y_{2j} \\ \sum_j Y_{1j} \\ \sum_j Y_{2j} \end{pmatrix} \\ &= \begin{pmatrix} \bar{Y}_2 \\ \bar{Y}_1 - \bar{Y}_2 \\ 0 \end{pmatrix}\end{aligned}$$

Compute $\hat{\mathbf{Y}}$ as $\mathbf{X}\hat{\boldsymbol{\beta}}$:

$$\begin{pmatrix} 1 & 1 & 0 \\ \vdots & \vdots & \vdots \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ \vdots & \vdots & \vdots \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} \bar{Y}_2 \\ \bar{Y}_1 - \bar{Y}_2 \\ 0 \end{pmatrix} = \begin{pmatrix} \bar{Y}_{1\cdot} \\ \vdots \\ \bar{Y}_{1\cdot} \\ \bar{Y}_{2\cdot} \\ \vdots \\ \bar{Y}_{2\cdot} \end{pmatrix}$$

9.4. Properties of $\mathbf{P}$

Lemma 9.5.1: Let $\text{rank}(\mathbf{X}) = r < p$ and $\mathbf{P} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'$ where $(\mathbf{X}'\mathbf{X})^{-}$ is a generalized inverse of $\mathbf{X}'\mathbf{X}$. Then

(i) $\mathbf{P}$ and $\mathbf{I} - \mathbf{P}$ are projection matrices.

(ii) $\text{rank}(\mathbf{I} - \mathbf{P}) = \text{tr}(\mathbf{I} - \mathbf{P}) = n - r$.

(iii) $\mathbf{P}\mathbf{X} = \mathbf{X}$.

Note that by the Theorem part (4.), all generalized inverses give the same $\mathbf{P}$.

Proof:

(i) $\mathbf{P}$ is symmetric by the Theorem part (5.).

$\mathbf{P}$ is idempotent: Let $\mathbf{G} = (\mathbf{X}'\mathbf{X})^{-}$ so that $\mathbf{P} = \mathbf{X}\mathbf{G}\mathbf{X}'$. Then $\mathbf{P}^2 = (\mathbf{X}\mathbf{G}\mathbf{X}')(\mathbf{X}\mathbf{G}\mathbf{X}') = (\mathbf{X}\mathbf{G})(\mathbf{X}'\mathbf{X}\mathbf{G}\mathbf{X}') = \mathbf{X}\mathbf{G}\mathbf{X}'$ by the Theorem part (3a.).

So $\mathbf{P}$ is symmetric and idempotent and thus a projection. It then follows easily that $\mathbf{I} - \mathbf{P}$ is as well.

(ii) We know rank = trace for projection matrices.

(iii) $\mathbf{P}\mathbf{X} = \mathbf{X}\mathbf{G}\mathbf{X}'\mathbf{X} = \mathbf{X}$ by the Theorem part (3b.).

9.5. METHOD 2: Reduce the model to one of full rank.

Let $\mathbf{X}_1$ consist of r linearly independent columns from $\mathbf{X}$ and let $\mathbf{X}_2$ consist of the remaining columns. Then $\mathbf{X_2} = \mathbf{X_1 F}$ for some $\mathbf{F}$ because the columns of $\mathbf{X}_1$ span the column space of $\mathbf{X}$. Write

$$\mathbf{X} = (\mathbf{X_1}, \mathbf{X_2}) = (\mathbf{X_1}, \mathbf{X_1 F}) = \mathbf{X_1}(\mathbf{I_{r\times r}}, \mathbf{F}),$$

This is a special case of the factorization

$$\mathbf{X} = \mathbf{KL},$$

where $\text{rank}(\mathbf{K}_{n\times r}) = r$ and $\text{rank}(\mathbf{L}_{r\times p}) = r$. Using this general notation,

$$E[\mathbf{Y}] = \mathbf{X}\boldsymbol{\beta} = \mathbf{KL}\boldsymbol{\beta} = \mathbf{K}\boldsymbol{\alpha}.$$

Therefore, we have reparametrized the linear model into one of full rank.

Because $\mathbf{K}$ has full rank, the least squares estimate of $\boldsymbol{\alpha}$ is

$$\hat{\boldsymbol{\alpha}} = (\mathbf{K'K})^{-\mathbf{1}}\mathbf{K'Y}$$

and

$$\hat{\mathbf{Y}} = \mathbf{K}\hat{\boldsymbol{\alpha}} = \mathbf{K}(\mathbf{K'K})^{-\mathbf{1}}\mathbf{K'Y}.$$

Therefore, $\mathbf{P} = \mathbf{K}(\mathbf{K'K})^{-\mathbf{1}}\mathbf{K'}$. If $\mathbf{K} = \mathbf{X}_1$, this gives us our previous expression $\mathbf{P} = \mathbf{X_1}(\mathbf{X_1'X_1})^{-\mathbf{1}}\mathbf{X_1'}$.

In-Class Exercise: One-way ANOVA with 2 groups – revisited

Let $\mathbf{X}_1$ consist of the first 2 columns of $\mathbf{X}$. Apply the method on the previous page. What is $\mathbf{F}$? What is $\boldsymbol{\alpha}$? What is $\hat{\boldsymbol{\alpha}}$? What is $\hat{\mathbf{Y}}$?

See the end of this document for the solution.

9.6. METHOD 3: Impose identifiability constraints

Impose $s \equiv p - r$ linear constraints on $\boldsymbol{\beta}$ of the form $\mathbf{H}_{s\times p}\boldsymbol{\beta} = \mathbf{0}$ to make $\boldsymbol{\beta}$ uniquely determined ("identifiable"), i.e., $\hat{\boldsymbol{\beta}}$ is unique: for any $\hat{\mathbf{Y}} \in \mathcal{R}(\mathbf{X})$, there is a unique $\hat{\boldsymbol{\beta}}$ satisfying

$$\mathbf{X}\hat{\boldsymbol{\beta}} = \hat{\mathbf{Y}} \ \underline{\text{and}} \ \mathbf{H}\hat{\boldsymbol{\beta}} = \mathbf{0}.$$

Let's drop the "hats" for aestetics. We can re-write this as

$$\begin{pmatrix} \mathbf{Y} \\ \mathbf{0} \end{pmatrix} = \begin{pmatrix} \mathbf{X} \\ \mathbf{H} \end{pmatrix} \boldsymbol{\beta} \equiv \mathbf{G}\boldsymbol{\beta}.$$

Under what conditions have we accomplished our goal? When is there a unique solution to $\begin{pmatrix} \mathbf{Y} \\ \mathbf{0} \end{pmatrix} = \mathbf{G}\boldsymbol{\beta}$?

Lemma 9.6.1: A unique solution exists if and only if $\mathbf{G}$ has rank p and the rows of $\mathbf{H}$ are linearly independent of the rows of $\mathbf{X}$.

Proof: A solution $\boldsymbol{\beta}$ will exist for all $\mathbf{Y} \in \mathcal{R}(\mathbf{X})$ if and only if

$$\begin{pmatrix} \mathbf{Y} \\ \mathbf{0} \end{pmatrix} \in \mathcal{R}(\mathbf{G}),$$

if and only if, for all $\mathbf{u} \in \Re^{n+s}$,

$$\mathbf{G}'\mathbf{u} = \mathbf{0} \Longrightarrow \begin{pmatrix} \mathbf{Y} \\ \mathbf{0} \end{pmatrix}' \mathbf{u} = 0$$

$$(\text{ since } \begin{pmatrix} \mathbf{Y} \\ \mathbf{0} \end{pmatrix} = \mathbf{GF}, (\mathbf{GF})'\mathbf{u} = \mathbf{F}'\mathbf{G}'\mathbf{u} = \mathbf{F}'\mathbf{0} = \mathbf{0}).$$

This is equivalent to

$$(\mathbf{X}', \mathbf{H}') \begin{pmatrix} \mathbf{u}_X \\ \mathbf{u}_H \end{pmatrix} = \mathbf{0} \Longrightarrow (\mathbf{Y}', \mathbf{0}') \begin{pmatrix} \mathbf{u}_X \\ \mathbf{u}_H \end{pmatrix} = 0,$$

i.e.,

$$\begin{aligned} \mathbf{X}'\mathbf{u}_X + \mathbf{H}'\mathbf{u}_H = \mathbf{0} &\Longrightarrow \mathbf{Y}'\mathbf{u}_X = 0, \text{ for all } \mathbf{Y} \in \mathcal{R}(\mathbf{X}) \\ &\Longrightarrow (\mathbf{Xa})'\mathbf{u}_X = 0, \text{ for all } \mathbf{a} \\ &\Longrightarrow \mathbf{X}'\mathbf{u}_X = \mathbf{0}, \end{aligned}$$

if and only if

$$\mathbf{X}'\mathbf{u}_X + \mathbf{H}'\mathbf{u}_H = \mathbf{0} \Longrightarrow \mathbf{X}'\mathbf{u}_X = \mathbf{0} \text{ and } \mathbf{H}'\mathbf{u}_H = \mathbf{0},$$

This just says that the rows of $\mathbf{X}$ are linearly independent of the rows of $\mathbf{H}$.

The solution will be unique if and only if the columns of $\mathbf{G}$ are linearly independent, i.e., $\text{rank}(\mathbf{G}) = p$.

Corollary: A unique solution exists if and only if $\mathbf{G}$ has rank p and $\mathbf{H}$ has rank $s = p - r$.

To use this method to estimate $\boldsymbol{\beta}$, we solve $\hat{\mathbf{Y}} = \mathbf{X}\hat{\boldsymbol{\beta}}$ and $\mathbf{H}\hat{\boldsymbol{\beta}} = \mathbf{0}$, i.e., we solve the augmented normal equations:

$$\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{X}'\mathbf{Y} \text{ and } \mathbf{H}'\mathbf{H}\hat{\boldsymbol{\beta}} = \mathbf{0}$$

These equalities together give:

$$(\mathbf{X}'\mathbf{X} + \mathbf{H}'\mathbf{H})\hat{\boldsymbol{\beta}} = (\mathbf{G}'\mathbf{G})\hat{\boldsymbol{\beta}} = \mathbf{X}'\mathbf{Y},$$

The solution is given by:

$$\hat{\boldsymbol{\beta}} = (\mathbf{G}'\mathbf{G})^{-1}\mathbf{X}'\mathbf{Y},$$

and $\hat{\mathbf{Y}} = \mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{P}\mathbf{Y}$, where $\mathbf{P} = \mathbf{X}(\mathbf{G}'\mathbf{G})^{-1}\mathbf{X}'$.

In Class Exercise: One-way ANOVA with 2 groups – revisited: Apply Method 3 using the constraint $\alpha_1 + \alpha_2 = 0$. To make life easier, assume $n_1 = n_2 = \frac{1}{2}n$. What is $\mathbf{H}$? What is $\hat{\boldsymbol{\beta}}$? What is $\hat{\mathbf{Y}}$?

You "might" find the following result useful:

$$\begin{pmatrix} n & \frac{n}{2} & \frac{n}{2} \\ \frac{n}{2} & \frac{n}{2}+1 & 1 \\ \frac{n}{2} & 1 & \frac{n}{2}+1 \end{pmatrix}^{-1} = \begin{pmatrix} \frac{n+4}{4n} & -\frac{1}{4} & -\frac{1}{4} \\ -\frac{1}{4} & \frac{n+4}{4n} & \frac{n-4}{4n} \\ -\frac{1}{4} & \frac{n-4}{4n} & \frac{n+4}{4n} \end{pmatrix}$$

See the end of this document for the solution.

9.6. METHOD $3\frac{1}{2}$: Impose identifiability constraints, implemented differently.

Consider the example of one-way ANOVA with, say, 5 groups. Suppose we use the constraint $\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 + \alpha_5 = 0$. We can use this constraint to formulate a full-rank model, which we illustrate by this example. (Thus the method might be considered "in between" methods 2 and 3.)

Consider one "replication" – five observations from each of the five groups. The design matrix $\mathbf{X}$ will have five columns (not six), for $\mu, \alpha_1, \alpha_2, \alpha_3, \alpha_4$, (but not α_5).

$$\mathbf{X} = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 1 & -1 & -1 & -1 & -1 \end{pmatrix}$$

The row of $\mathbf{X}$ for the observation from group 5 is gotten from using $\alpha_5 = -\alpha_1 - \alpha_2 - \alpha_3 - \alpha_4$.

Solution to Exercise for Method 3

$$\mathbf{H}\boldsymbol{\beta} \equiv (0,1,1)\begin{pmatrix}\mu\\ \alpha_1\\ \alpha_2\end{pmatrix} = 0$$

$$\mathbf{G}'\mathbf{G} = \begin{pmatrix} n & n/2 & n/2 \\ n/2 & n/2+1 & 1 \\ n/2 & 1 & n/2+1 \end{pmatrix}$$

$$(\mathbf{G}'\mathbf{G})^{-1} = \begin{pmatrix} \frac{n+4}{4n} & \frac{-1}{4} & \frac{-1}{4} \\ \frac{-1}{4} & \frac{n+4}{4n} & \frac{n-4}{4n} \\ \frac{-1}{4} & \frac{n-4}{4n} & \frac{n+4}{4n} \end{pmatrix}$$

$$\hat{\boldsymbol{\beta}} = (\mathbf{G}'\mathbf{G})^{-1}\mathbf{X}'\mathbf{Y} = \begin{pmatrix} \bar{Y}_{..} \\ \frac{1}{2}(\bar{Y}_{1.} - \bar{Y}_{2.}) \\ \frac{1}{2}(\bar{Y}_{2.} - \bar{Y}_{1.}) \end{pmatrix}$$

Clearly $\hat{\boldsymbol{\beta}}$ satisfies the constraint $\alpha_1 + \alpha_2 = 0$. Again, we have

$$\hat{\mathbf{Y}} = \mathbf{X}\hat{\boldsymbol{\beta}} = \begin{pmatrix} \bar{Y}_{1.} \\ \vdots \\ \bar{Y}_{1.} \\ \bar{Y}_{2.} \\ \vdots \\ \bar{Y}_{2.} \end{pmatrix}.$$

Solution to Exercise for Method 2

The third column of $\mathbf{X}$ is the first column minus the second column. Therefore, $\mathbf{F} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$

$$\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_1\mathbf{F}) = \underbrace{\mathbf{X}_1}_{\mathbf{K}} \underbrace{\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix}}_{\mathbf{L}}$$

$E(\mathbf{Y}) = \mathbf{X}\boldsymbol{\beta} = \mathbf{KL}\boldsymbol{\beta} = \mathbf{K}\boldsymbol{\alpha}$ where $\boldsymbol{\alpha} = \mathbf{L}\boldsymbol{\beta}$

$$\boldsymbol{\alpha} = \mathbf{L}\boldsymbol{\beta} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} \mu \\ \alpha_1 \\ \alpha_2 \end{pmatrix} = \begin{pmatrix} \mu + \alpha_2 \\ \alpha_1 - \alpha_2 \end{pmatrix}.$$

$$\begin{aligned} \hat{\boldsymbol{\alpha}} &= (\mathbf{K}'\mathbf{K})^{-1}\mathbf{K}'\mathbf{Y} \\ &= \begin{pmatrix} n & n_1 \\ n_1 & n_1 \end{pmatrix}^{-1} \begin{pmatrix} \sum_j Y_{1j} + \sum_j Y_{2j} \\ \sum_j Y_{1j} \end{pmatrix} \\ &= \begin{pmatrix} n_2^{-1} & -n_2^{-1} \\ -n_2^{-1} & n_1^{-1} + n_2^{-1} \end{pmatrix} \begin{pmatrix} \sum_j Y_{1j} + \sum_j Y_{2j} \\ \sum_j Y_{1j} \end{pmatrix} \\ &= \begin{pmatrix} \bar{Y}_2 \\ \bar{Y}_1 - \bar{Y}_2 \end{pmatrix} \end{aligned}$$

$$\hat{\mathbf{Y}} = \mathbf{X}_1\hat{\boldsymbol{\alpha}} = \begin{pmatrix} \bar{Y}_{1.} \\ \vdots \\ \bar{Y}_{1.} \\ \bar{Y}_{2.} \\ \vdots \\ \bar{Y}_{2.} \end{pmatrix}, \text{ as before}$$