Alternative Parameterizations

- Recall that the Gauss-Markov Linear Model simply says that $\mathrm{E}(\boldsymbol{y}) \in \mathcal{C}(\boldsymbol{X})$ and $\mathrm{Var}(\boldsymbol{y}) = \sigma^2 \boldsymbol{I}$ for some $\sigma^2 > 0$.
- Thus, as long as $\mathcal{C}(\boldsymbol{X}) = \mathcal{C}(\boldsymbol{W})$, the following models are identical.

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon} \qquad \boldsymbol{\epsilon} \sim (\boldsymbol{0}, \sigma^2 \boldsymbol{I})$$

$$\boldsymbol{y} = \boldsymbol{W}\boldsymbol{\alpha} + \boldsymbol{\epsilon} \qquad \boldsymbol{\epsilon} \sim (\boldsymbol{0}, \sigma^2 \boldsymbol{I})$$

- For example y_{ij} $i = 1, 2, 3$ $j = 1, 2$

Treatment Effects

$\mathrm{E}(y_{ij}) = \mu + \tau_i$

$$\begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu \\ \tau_1 \\ \tau_2 \\ \tau_3 \end{bmatrix}$$

$\boldsymbol{X}$ $\boldsymbol{\beta}$

Cell Means

$\mathrm{E}(y_{ij}) = \mu_i$

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix}$$

$\boldsymbol{W}$ $\boldsymbol{\alpha}$

- Both models state that

$$
\mathrm{E}(\boldsymbol{y}) \in \left\{ \begin{bmatrix} a_1 \\ a_1 \\ a_2 \\ a_2 \\ a_3 \\ a_3 \end{bmatrix} \quad : \quad a_1, a_2, a_3 \in \mathbb{R} \right\}
$$

$$
= \quad \mathcal{C}(\boldsymbol{X}) \quad = \quad \mathcal{C}(\boldsymbol{W}).
$$

- Thus, the two models are identical.

- In models like the treatment effects model where the design matrix is not of full column rank, "identifiability constraints" are often imposed.

- The most common constraints are

"set first to zero"	$\tau_1 = 0$	R default
"set last to zero"	$\tau_3 = 0$	SAS effective default
"sum to zero"	$\tau_1 + \tau_2 + \tau_3 = 0$	

- Such constraints are not necessary.
- Placing such constriants on the parameters can be viewed as equivalent to choosing a particular full column rank design matrix.
- Set first to zero:

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu \\ \tau_2 \\ \tau_3 \end{bmatrix} = \begin{bmatrix} \mu \\ \mu \\ \mu + \tau_2 \\ \mu + \tau_2 \\ \mu + \tau_3 \\ \mu + \tau_3 \end{bmatrix}$$

- Set last to zero:

$$\begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \mu \\ \tau_1 \\ \tau_2 \end{bmatrix} = \begin{bmatrix} \mu + \tau_1 \\ \mu + \tau_1 \\ \mu + \tau_2 \\ \mu + \tau_2 \\ \mu \\ \mu \end{bmatrix}$$

- Sum to zero: $\tau_1 + \tau_2 + \tau_3 = 0 \iff \tau_3 = -\tau_1 - \tau_2$

$$\begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & -1 & -1 \\ 1 & -1 & -1 \end{bmatrix} \begin{bmatrix} \mu \\ \tau_1 \\ \tau_2 \end{bmatrix} = \begin{bmatrix} \mu + \tau_1 \\ \mu + \tau_1 \\ \mu + \tau_2 \\ \mu + \tau_2 \\ \mu - \tau_1 - \tau_2 \\ \mu - \tau_1 - \tau_2 \end{bmatrix}$$

- Instead of viewing "identifiability constraints" as constraints on parameters, we can view them as constraints on our solutions to the normal equations.

- For example,

$$\boldsymbol{X} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix} \quad \boldsymbol{\beta} = \begin{bmatrix} \mu \\ \tau_1 \\ \tau_2 \\ \tau_3 \end{bmatrix} \quad \boldsymbol{y} = \begin{bmatrix} y_{11} \\ y_{12} \\ y_{21} \\ y_{22} \\ y_{31} \\ y_{33} \end{bmatrix}$$

Normal Equations $\boldsymbol{X'Xb} = \boldsymbol{X'y}$

$$\begin{bmatrix} n. & n_1 & n_2 & n_3 \\ n_1 & n_1 & 0 & 0 \\ n_2 & 0 & n_2 & 0 \\ n_3 & 0 & 0 & n_3 \end{bmatrix} \begin{bmatrix} ? \\ ? \\ ? \\ ? \end{bmatrix} = \begin{bmatrix} y_{..} \\ y_{1.} \\ y_{2.} \\ y_{3.} \end{bmatrix}$$

Set first to zero

$$\begin{bmatrix} \bar{y}_{1.} \\ 0 \\ \bar{y}_{2.} - \bar{y}_{1.} \\ \bar{y}_{3.} - \bar{y}_{1.} \end{bmatrix}$$

Set last to zero

$$\begin{bmatrix} \bar{y}_{3.} \\ \bar{y}_{1.} - \bar{y}_{3.} \\ \bar{y}_{2.} - \bar{y}_{3.} \\ 0 \end{bmatrix}$$

Sum to zero

$$\begin{bmatrix} (\bar{y}_{1.} + \bar{y}_{2.} + \bar{y}_{3.})/3 \\ \bar{y}_{1.} - (\bar{y}_{1.} + \bar{y}_{2.} + \bar{y}_{3.})/3 \\ \bar{y}_{2.} - (\bar{y}_{1.} + \bar{y}_{2.} + \bar{y}_{3.})/3 \\ \bar{y}_{3.} - (\bar{y}_{1.} + \bar{y}_{2.} + \bar{y}_{3.})/3 \end{bmatrix}$$

- As noted before, such constraints are not necessary; we do not need to consider constraints when we work with generalized inverses.

- In 611, we study the issue of constraints much more carefully. We show that constraints of the form $\boldsymbol{Mb} = \mathbf{0}$ produce a unique solution to the normal equations without restricting that solution to a proper subset of $\mathcal{C}(\boldsymbol{X})$ if and only if
$$\mathcal{C}(\boldsymbol{X}') \cap \mathcal{C}(\boldsymbol{M}') = \{\mathbf{0}\} \text{ and } \operatorname{rank}(\boldsymbol{M}) = p - \operatorname{rank}(\boldsymbol{X}).$$

- Such constraints have no impact on what linear functions of $\boldsymbol{\beta}$ are estimable or on inferences about estimable functions of $\boldsymbol{\beta}$. Thus, the same analysis results are obtained with or without constraints.

- In 511, we often start with a non-full-rank design matrix for ease and symmetry when specifying a model.

- For example, it is nice to be able to write

  $y_{ijk} = \mu + \alpha_i + \beta_j + \epsilon_{ijk}$ for $i = 1, 2, 3; j = 1, 2, 3, 4; k = 1, \ldots, 10$

  if we want to specify a two-factor additive model.

- The design matrix that matches this model specification does not have full-column rank.

- R (and other statistical software) will automatically pick a full-column-rank design matrix, which is equivalent to placing certain constraints on the solutions to the normal equations.

- It does not matter which full-column-rank design matrix (or corresponding constraint set) is chosen as long as the column space of the selected design matrix is the same as the column space of the design matrix for the model as originally specified.

- However, it is important to understand the design matrix used so that the parameter estimates corresponding to coefficients in the linear combination of the columns of the design matrix can be properly interpreted.

- For example, suppose $E(y_{ij}) = \mu + \tau_i$ for $i = 1, 2, 3$ and $j = 1, \ldots, n_i$
- What does the parameter τ_2 represent?

Constraints	Interpretation of τ_2
none	non-estimable
set first to zero	trt 2 mean − trt 1 mean
set last to zero	trt 2 mean − trt 3 mean
sum to zero	trt 2 mean − average of trt 1, 2, 3 means

- Recall that all linear functions of $\mathrm{E}(\boldsymbol{y})$ are the only estimable quantities; i.e., the estimable quantities are given by $\{\boldsymbol{A}\mathrm{E}(\boldsymbol{y}) : \boldsymbol{A}$ an n-column matrix of constants$\}$.
- Thus, as long as models restrict $\mathrm{E}(\boldsymbol{y})$ to the same column space, the estimable quantities are identical.

- Then why, in the treatment effects formulation of the model, is it that τ_2 is estimable under "set first to zero" constraint but not estimable without constraints?
- Under "set first to zero" constraint, τ_2 is the estimable quantity "treatment 2 mean $-$ treatment 1 mean."
- With no constraints, that same quantity is also estimable, but it is $\tau_2 - \tau_1$.