The general linear model (and PROC GLM)

Solving systems of linear equations

$$3\beta_0 + 1\beta_1 = 12$$
$$4\beta_0 + 2\beta_1 = 8$$

can be written as

$$\begin{bmatrix} 3 & 1 \\ 4 & 2 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 12 \\ 8 \end{bmatrix} \quad \text{or} \quad \mathbf{A}\boldsymbol{\beta} = \text{c} \; .$$

Now look at the matrix $\begin{bmatrix} -\frac{2}{10} & \frac{4}{10} \\ \frac{4}{10} & -\frac{3}{10} \end{bmatrix}$. One can see that

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

$$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \mathbf{I} \, , \text{ the identity matrix.}$$

The matrix $\mathbf{A}^{-1} = \begin{bmatrix} -\frac{2}{10} & \frac{4}{10} \\ \frac{4}{10} & -\frac{3}{10} \end{bmatrix}$ is called the inverse matrix for the matrix $\mathbf{A}$. If such a matrix can be found, then the solution to the linear equations becomes

$$\mathbf{A}^{-1}\mathbf{A}\boldsymbol{\beta} = \mathbf{I}\boldsymbol{\beta} = \boldsymbol{\beta} = \mathbf{A}^{-1}\mathbf{c}$$

This solution is a unique solution.

Now suppose we have instead

$$\begin{aligned} 3\beta_0 + 1\beta_1 &= 12 \\ 6\beta_0 + 2\beta_1 &= 8 \end{aligned}$$

that is, $\mathbf{A} = \begin{bmatrix} 3 & 1 \\ 6 & 2 \end{bmatrix}$. Note that the equations have no solution. The **determinant** of **A** is

$$\det(\mathbf{A}) = 3(2) - 6(1) = 0 \,.$$

A is said to be **singular**. If we just consider the first equation, there are an infinite number of solutions (an arbitrary solution to the equation can be found just by setting $\beta_1 =$ constant, say, $\beta_1 = 3$, and then $\beta_0 = 3$ results.

General linear model

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$$

where **Y** is the $N \times 1$ vector of random variables (the responses), **X** is the $N \times (k+1)$ **design matrix** (with $N > (k+1)$), $\boldsymbol{\beta}$ is the $(k+1) \times 1$ column vector of parameters, and $\boldsymbol{\epsilon}$ is an $N \times 1$ column vector of independent normal$(0, \sigma^2)$ random variables. Let **y** denote the vector of data (recorded values of the variables in **Y**). Parameter estimates satisfy a system of linear equations (the **normal equations**):

$$(\mathbf{X}'\mathbf{X})\widehat{\boldsymbol{\beta}} = \mathbf{X}'\mathbf{y}$$

If $(\mathbf{X}'\mathbf{X})$ is nonsingular, its inverse $(\mathbf{X}'\mathbf{X})^{-1}$ exists, and the equations have a solution.

Matrix theory result: if the columns of **X** are linearly independent, that is, if one column cannot be written as a linear combination of the other columns, then $(\mathbf{X}'\mathbf{X})$ is nonsingular.

AOV: means coding

Suppose there are 3 treatment levels, with means μ_1, μ_2, and μ_3, and 4 observations in each cell. The model is

$$Y_{ij} = \mu_i + \epsilon_{ij}$$

The “means coding” for the design matrix is

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

Note that an indicator column is left off for the third treatment; otherwise the three last rows add up to the first row.
The parameters are

$$\beta_0 = \mu_3$$
$$\beta_0 + \beta_1 = \mu_1$$
$$\beta_0 + \beta_2 = \mu_2$$

AOV: effects coding

Suppose there are 3 treatment levels, with means μ_1, μ_2, and μ_3 now parameterized as $\mu_1 = \mu + \alpha_1$, $\mu_2 = \mu + \alpha_2$, $\mu_3 = \mu + \alpha_3$, where the α_is sum to zero, and 4 observations in each cell. The model is

$$Y_{ij} = \mu + \alpha_i + \epsilon_{ij}$$

The "effects coding" for the design matrix is

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

The parameters are

$$\beta_0 = \mu$$
$$\beta_1 = \alpha_1$$
$$\beta_2 = \alpha_2$$
$$-(\beta_1 + \beta_2) = \alpha_3$$

Two factors: suppose there is another factor with 2 levels, in a completely randomized factorial design, 2 observations per cell. Main effects model is

$$Y_{ijk} = \mu + \alpha_i + \gamma_j + \epsilon_{ijk}$$

The design matrix (effects coding) is

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

Parameters are

$$\beta_0 = \mu$$
$$\beta_1 = \alpha_1$$
$$\beta_2 = \alpha_2$$
$$-(\beta_1 + \beta_2) = \alpha_3$$
$$\beta_3 = \gamma_1$$
$$-\beta_3 = \gamma_2$$