STAT 510 Homework 5
Due Date: 11:00 A.M., Wednesday, February 26

1. Suppose $\boldsymbol{X}$ is an $n \times p$ matrix and $\boldsymbol{B}$ is a $p \times p$ non-singular matrix. Prove that

$$\mathcal{C}(\boldsymbol{X}) = \mathcal{C}\left(\boldsymbol{X}\boldsymbol{B}^{-1}\right).$$

2. Consider the Gauss-Markov model with normal errors $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$, where $\boldsymbol{\epsilon} \sim N(\boldsymbol{0}, \sigma^2\boldsymbol{I})$. For any nonsingular $p \times p$ matrix $\boldsymbol{B}$, the model can be reparameterized by

$$\boldsymbol{X}\boldsymbol{\beta} = \boldsymbol{X}\boldsymbol{B}^{-1}\boldsymbol{B}\boldsymbol{\beta} = \boldsymbol{W}\boldsymbol{\alpha}, \text{ where } \boldsymbol{W} = \boldsymbol{X}\boldsymbol{B}^{-1} \text{ and } \boldsymbol{\alpha} = \boldsymbol{B}\boldsymbol{\beta}.$$

From problem 1, we know the column spaces of $\boldsymbol{X}$ and $\boldsymbol{W}$ are identical so that $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$ and $\boldsymbol{y} = \boldsymbol{W}\boldsymbol{\alpha} + \boldsymbol{\epsilon}$ are the same models. Suppose $\boldsymbol{C}$ is a $q \times p$ matrix of rank $q < p$. Then there exists a $(p-q) \times p$ matrix $\boldsymbol{A}$ such that

$$\boldsymbol{B} = \begin{bmatrix} \boldsymbol{A} \\ \boldsymbol{C} \end{bmatrix} \text{ has rank } p \text{ and is therefore nonsingular.}$$

Then we can write

$$\boldsymbol{B}\boldsymbol{\beta} = \begin{bmatrix} \boldsymbol{A} \\ \boldsymbol{C} \end{bmatrix} \boldsymbol{\beta} = \begin{bmatrix} \boldsymbol{A}\boldsymbol{\beta} \\ \boldsymbol{C}\boldsymbol{\beta} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\alpha}_1 \\ \boldsymbol{\alpha}_2 \end{bmatrix} = \boldsymbol{\alpha},$$

where $\boldsymbol{\alpha}_1 = \boldsymbol{A}\boldsymbol{\beta}$ and $\boldsymbol{\alpha}_2 = \boldsymbol{C}\boldsymbol{\beta}$. If we let $\boldsymbol{W}_1$ be the matrix consisting of the first $p-q$ columns of $\boldsymbol{W} = \boldsymbol{X}\boldsymbol{B}^{-1}$ and $\boldsymbol{W}_2$ be the matrix consisting of the last q columns of $\boldsymbol{W} = \boldsymbol{X}\boldsymbol{B}^{-1}$, then

$$\boldsymbol{W}\boldsymbol{\alpha} = [\boldsymbol{W}_1, \boldsymbol{W}_2] \begin{bmatrix} \boldsymbol{\alpha}_1 \\ \boldsymbol{\alpha}_2 \end{bmatrix} = \boldsymbol{W}_1\boldsymbol{\alpha}_1 + \boldsymbol{W}_2\boldsymbol{\alpha}_2.$$

Now consider testing

$$H_0 : \boldsymbol{C}\boldsymbol{\beta} = \boldsymbol{0} \text{ vs. } H_A : \boldsymbol{C}\boldsymbol{\beta} \neq \boldsymbol{0}.$$

(a) Rewrite these hypotheses in terms of the $\boldsymbol{\alpha}$ parameter vector.

(b) If you wanted to fit a reduced model corresponding to the null hypothesis, what model matrix would you use?

(c) Consider the unbalanced experiment described in slide set 8. Assume the full model given on slide 6. Provide a matrix $\boldsymbol{C}$ for testing the main effect of temperature.

(d) Provide a matrix $\boldsymbol{A}$ so that

$$\boldsymbol{B} = \begin{bmatrix} \boldsymbol{A} \\ \boldsymbol{C} \end{bmatrix}$$

is a 4×4 matrix of rank 4.

(e) Provide a model matrix for a reduced model that corresponds to the null hypothesis of no temperature main effect.

(f) Find the error sum of squares for the reduced and full models.

(g) Find the degrees of freedom associated with the sums of squares in part (f).

(h) Compute the F-statistic for testing the null hypothesis of no temperature main effect using the sums of squares and degrees of freedom computed in parts (f) and (g).

3. An experiment was conducted to study the effect of two diets (1 and 2) and two drugs (1 and 2) on blood pressure in rats. A total of 40 rats were randomly assigned to the 4 combinations of diet and drug, with 10 rats per combination. Let y_{ijk} be the decrease in blood pressure from the beginning of the study to the end of the study for diet i, drug j, and rat k ($i = 1, 2$; $j = 1, 2$; $k = 1, \dots, 10$). Suppose

$$y_{ijk} = \mu_{ij} + \epsilon_{ijk}, \tag{1}$$

where the μ_{ij} terms are unknown parameters and the ϵ_{ijk} terms are independent and identically distributed as $N(0, \sigma^2)$ for some unknown variance parameter $\sigma^2 > 0$.

A researcher suspects the mean reduction in blood pressure will be the same for all combinations of diet and drug except for the combination of diet 1 with drug 1. This leads to consideration of the null hypothesis

$$H_0 : \mu_{12} = \mu_{21} = \mu_{22}.$$

Assuming model (1) holds, determine the distribution of the F statistic you would use to test this null hypothesis. State the degrees of freedom of the statistic and provide a fully simplified expression for the noncentrality parameter of the statistic in terms of model (1) parameters.