STAT 511 Homework 3

Due Date: 11:00 A.M., Friday, February 3

1. A result from linear algebra states that $rank\,(\boldsymbol{AB}) \leq \min\{rank\,(\boldsymbol{A})\,, rank\,(\boldsymbol{B})\}$ for any matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ of appropriate order. (The number of columns of $\boldsymbol{A}$ must match the number of rows of $\boldsymbol{B}$ so that the product $\boldsymbol{AB}$ can be formed.) You may wish to use this result to help you prove the following.

   (a) For any matrix $\boldsymbol{X}$, $rank\,(\boldsymbol{X}) = rank\,(\boldsymbol{X}'\boldsymbol{X})$.

   (b) For any matrix $\boldsymbol{X}$, $rank\,(\boldsymbol{X}) = rank\,(\boldsymbol{P_X})$.

   (c) Suppose the Gauss-Markov linear model holds, and suppose $\boldsymbol{C\beta}$ is estimable. Furthermore, suppose the rows of $\boldsymbol{C}$ are linearly independent. Show that the rank of $\boldsymbol{C}\,(\boldsymbol{X}'\boldsymbol{X})^{-}\,\boldsymbol{C}'$ is q, where q is the number of rows of C.

   Note that $\boldsymbol{C}\,(\boldsymbol{X}'\boldsymbol{X})^{-}\,\boldsymbol{C}'$ is a $q \times q$ matrix. Thus, $rank\left(\boldsymbol{C}\,(\boldsymbol{X}'\boldsymbol{X})^{-}\,\boldsymbol{C}'\right) = q$ is equivalent to $\boldsymbol{C}\,(\boldsymbol{X}'\boldsymbol{X})^{-}\,\boldsymbol{C}'$ being non-singular. In class, we showed that $\text{Var}(\boldsymbol{C}\hat{\boldsymbol{\beta}}) = \sigma^2\boldsymbol{C}\,(\boldsymbol{X}'\boldsymbol{X})^{-}\,\boldsymbol{C}'$. Thus, completing this problem will show that $[\text{Var}(\boldsymbol{C}\hat{\boldsymbol{\beta}})]^{-1}$ exists when $\boldsymbol{C\beta}$ is estimable and $\boldsymbol{C}$ has full-row rank. This is a fact that we need when conducting inference for $\boldsymbol{C\beta}$. One way to prove the result is to proceed as follows.

$$\begin{aligned}
q &= rank\,(\boldsymbol{C}) \qquad (How\ do\ we\ know\ this?) \\
&= rank\,(\boldsymbol{AX})\ for\ some\ q \times n\ matrix\ \boldsymbol{A} \qquad (How\ do\ we\ know\ this?) \\
&= rank\,(\boldsymbol{AP_X X})\ (How\ do\ we\ know\ this?) \\
&\leq\ ??? \\
&=\ ??? = \cdots = rank\left(\boldsymbol{C}\left(\boldsymbol{X}'\boldsymbol{X}\right)^{-}\boldsymbol{C}'\right) \\
&\leq rank\,(\boldsymbol{C}) \qquad (How\ do\ we\ know\ this?) \\
&= q.
\end{aligned}$$

   Your task is to fill in the missing steps. If you do so, the argument above shows that $rank\left(\boldsymbol{C}\,(\boldsymbol{X}'\boldsymbol{X})^{-}\,\boldsymbol{C}'\right)$ is bounded both below and above by q and is thus equal to q. You might wish to use the result form part (a) somewhere along the way.

2. Consider an experiment designed to study the effect of two dietary factors, protein source and protein amount, on weight gain in pigs. A total of 12 pigs were randomly assigned to treatment with one of six combinations of protein source (1 or 2) and protein amount (1, 2, or 3 units). A completely randomized design was used with two individually penned pigs per treatment group. Let y_{ijk} denote the amount of weight gained during the study period by the k^{th} pig fed j units of protein from source i ($i = 1, 2; j = 1, 2, 3; k = 1, 2$). Consider the model

$$y_{ijk} = \mu + s_i + \beta x_j + \epsilon_{ijk}, \quad (i = 1, 2; j = 1, 2, 3; k = 1, 2)$$

   where μ, s_1, s_2,and β are unknown real-valued parameters, $x_j = j - 2$, the ϵ_{ijk}'s are $iid\ N(0, \sigma^2)$, and σ^2 is an unknown parameter in $I\!R^+$. Suppose

$$\boldsymbol{y} = [y_{111}, y_{112}, y_{121}, y_{122}, \ldots, y_{231}, y_{232}]',$$

$$\boldsymbol{\epsilon} = [\epsilon_{111}, \epsilon_{112}, \epsilon_{121}, \epsilon_{122}, \ldots, \epsilon_{231}, \epsilon_{232}]',$$

$$\text{and } \boldsymbol{\beta} = [\mu, s_1, s_2, \beta]'.$$

(a) Provide the appropriate design matrix $\boldsymbol{X}$ so that the model may be written as $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$.

(b) For each of the quantities below, state whether the quantity is estimable and prove that your answer is correct.

   i. $\mu + s_1$
   ii. $\mu + s_1 + 10\beta$
   iii. $s_1 - s_2$
   iv. μ

(c) Write down a full-column-rank matrix that has the same column space as $\boldsymbol{X}$ in part (a).

(d) Use your answer to part (c) to find a simplified expression for the BLUE of $\mathrm{E}(\mathrm{y}_{111})$ in terms of the y_{ijk} values.

(e) Provide the least squares estimate of each estimable quantity in part (b).