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

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

2. Consider the simple linear regression model

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i \quad i = 1, \ldots, n$$

where $\epsilon_1, \ldots, \epsilon_n$ are iid and have a normal distribution $N\left(0, \sigma^2\right)$ and $\beta_0, \beta_1$, and $\sigma^2 > 0$ are unknown parameters. The design matrix for this Normal theory Gauss-Markov linear model can be written as $X = [1, x]$, where $1$ is an $n \times 1$ vector of ones and $x = (x_1, \ldots, x_n)'$.

   (a) We have learned that the least squares estimator of $\beta$ in a Normal theory Gauss-Markov linear model with a full-rank design matrix is given by $\hat{\beta} = (X'X)^{-1} X'y$. Simplify this expression for the special case of simple linear regression to obtain expression for the least squares estimators of $\beta_0$ and $\beta_1$. Express your final answers using summation notation.

   (b) There are some computational advantages to working with a design matrix whose columns are orthogonal. For the simple linear regression problem consider the design matrix

   $$W = [1, x - \bar{x}1]$$

   This design matrix is obtained by "centering" the explanatory variable around its mean $\bar{x} = \sum_{i=1}^{n} x_i/n$. Find a matrix $B$ so that $XB^{-1} = W$. It will follow that

   $$X\beta = XB^{-1}B\beta = W\alpha$$

   where $B\beta = \alpha$.

   (c) Derive expressions for the least squares estimators of $\alpha_0$ and $\alpha_1$ (where $\alpha = (\alpha_0, \alpha_1)'$ from part (b)) using $\hat{\alpha} = (W'W)^{-1} W'y$.

   (d) Multiply $\hat{\alpha}$ from part (c) by $B^{-1}$ from part (b) to obtain expressions for $\hat{\beta}_0$ and $\hat{\beta}_1$.

   (e) Show that your answer to part (a) matches your answer to part (d).

3. An experiment was conducted to study the durability of coated fabric subjected to abrasive tests. Three factors were considered. One factor was filler type with two levels (F1 and F2). Another was surface treatment with two levels (S1 and S2). The third factor was proportion of filler with three levels (25%, 50%, and 75%). Using a completely randomized design with two fabric samples per treatment, the amount of fabric lost in milligrams for each fabric sample was recorded following testing. Data are available in a tab delimited text file at http://www.public.iastate.edu/~dnett/S511/FabricLoss.txt.

   (a) Consider a cell means model for this data. Estimate the mean and standard error for the treatment corresponding to S1, F2, and 25% filler.

   (b) Again considering a cell means model, find what SAS would call lsmeans for the factor filler type.

   (c) Provide a standard error for each of the estimates computed in part (b).

(d) Can the model be simplified by treating proportion of filler as a quantitative variable with a linear effect on the response? Conduct a test for lack of fit to answer this question. You should fit the same model as previously considered except that the factor proportion of filler is replaced with a quantitative variable that is linearly associated with the response within each combination of filler type and surface treatment. Provide an appropriate $F$ statistic, its degrees of freedom, a $p$-value, and a conclusion that addresses this question.

(e) Working with proportion of filler as a quantitative variable rather than a qualitative factor, find a simplified model that adequately fits the data by removing non-significant interactions from the model. Based on the fit of your simplified model, provide estimates of the intercept and slope for the relationship between fabric loss and filler proportion for each combination of filler type and surface treatment.

(f) Suppose the company conducting this experiment wants to know how to best protect their fabric from weight loss due to abrasion. What advice would you give based on the model arrived at in part (e)?