DIRECTIONS: There are 3 questions. Each question is worth 10 points. Show all of your work. Explain all of your reasoning. This part is open-book, open-notes. If you use software (e.g., Maple, R, SAS, etc), attach all output. **You may not talk with anyone besides the instructor!**

1. Consider two linear models for the same data:

$$\text{Model 1:} \quad \mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{e}$$
$$\text{Model 2:} \quad \mathbf{y} = \mathbf{W}\mathbf{c} + \mathbf{e}.$$

Here, $\mathbf{X}$ is $N \times p$, $\mathbf{W}$ is $N \times q$, $\mathbf{b}$ is $p \times 1$, and $\mathbf{c}$ is $q \times 1$. Suppose that $\mathcal{C}(\mathbf{W}) \subset \mathcal{C}(\mathbf{X})$. For Model 1, let $\widehat{\mathbf{y}}_X$, $\widehat{\mathbf{e}}_X$, and $\mathbf{P_X}$ denote the vector of (least-squares) fitted values, the vector of (least-squares) residuals, and the perpendicular projection matrix onto $\mathcal{C}(\mathbf{X})$. The quantities $\widehat{\mathbf{y}}_W$, $\widehat{\mathbf{e}}_W$, and $\mathbf{P_W}$ are defined analogously. For each part, you may use the previous parts in your answer; e.g., to prove (c), you may use the results from (a) and (b), etc.
(a) Show that $\mathbf{W} = \mathbf{XC}$, for some $p \times q$ matrix $\mathbf{C}$.
(b) Show that $\mathbf{P_X P_W} = \mathbf{P_W}$.
(c) Show that $(\widehat{\mathbf{y}}_X - \widehat{\mathbf{y}}_W)'\widehat{\mathbf{y}}_W = 0$.
(d) Show that $\mathbf{y}'\mathbf{y} = \widehat{\mathbf{y}}'_W\widehat{\mathbf{y}}_W + (\widehat{\mathbf{y}}_X - \widehat{\mathbf{y}}_W)'(\widehat{\mathbf{y}}_X - \widehat{\mathbf{y}}_W) + \widehat{\mathbf{e}}'_X\widehat{\mathbf{e}}_X$.

2. The effectiveness of three skin creams was studied in an experiment on N subjects. On the forearm of each subject, three locations were specified. The three creams were randomly allocated to locations on each subject; that is, each subject received a complete set of three treatments. It is assumed that the three observations on the same individual are correlated and that observations on different subjects are uncorrelated. A statistical model for this experiment is

$$y_{ij} = \mu_i + e_{ij},$$

where

- y_{ij} denotes the ith measurement on subject j, and
- μ_i denotes the mean response for the ith skin cream.

Assume that e_{ij}, for $i = 1, 2, 3$ and $j = 1, 2, ..., N$, are random variables with $E(e_{ij}) = 0$, $\text{var}(e_{ij}) = \sigma^2$, and $\text{corr}(e_{i'j}, e_{ij}) = \rho$, for $i \neq i'$. Note that $\text{corr}(e_{ij}, e_{ij'}) = 0$ when $j \neq j'$ (regardless of i) because e_{ij} and $e_{ij'}$ correspond to different subjects.
(a) Assuming that μ_i is fixed (not random), express this model as $\mathbf{y} = \mathbf{Xb} + \mathbf{e}$. Define all vectors and matrices. Your design matrix $\mathbf{X}$ should be full rank.
(b) Compute $\text{cov}(\mathbf{y})$.
(c) Find $\widehat{\mathbf{b}}$, the least-squares estimator of $\mathbf{b}$. Your answer should be a vector (I don't want to see matrices in your final answer).
(d) Compute $\text{cov}(\widehat{\mathbf{b}})$.

3. Let $\mathbf{P_X}$ denote the perpendicular projection matrix onto $\mathcal{C}(\mathbf{X})$.
(a) Give a detailed argument showing that $\mathbf{I}-\mathbf{P_X}$ is the perpendicular projection matrix onto $\mathcal{N}(\mathbf{X}')$.
(b) Let

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

Compute $\mathbf{P_X}$ and $\mathbf{I}-\mathbf{P_X}$.
(c) Express

$$\mathbf{y} = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{pmatrix}$$

as the sum of two vectors, one in $\mathcal{C}(\mathbf{X})$ and one in $\mathcal{N}(\mathbf{X}')$.
(d) For the $\mathbf{X}$ matrix in part (b), describe in words what $\mathcal{C}(\mathbf{X})$ and $\mathcal{N}(\mathbf{X}')$ are.