1. Let $H_0: C\beta = 0$ be testable. Show that $C(X^TX)C^T$ is unique, i.e. does not depend on the choice of conditional inverse for $X^TX$. (Hint: Let $(X^TX)_i^T$ be another conditional inverse for $X^TX$ and show that $C(X^TX)C^T = C(X^TX)_i^T C^T$. Remember that since $H_0$ is testable, $C(X^TX)^c X^T X = C$ for any conditional inverse of $X^TX$).

Solution:

\[
C(X^TX)^c C^T = C(X^TX)_i^T C(X^TX)^cX^T X \equiv C(X^TX)_i^T C^T.
\]

2. An industrial psychologist is investigating absenteeism among production-line workers, based on different types of work hours: (1) 4-day week with a 10-hour day, (2) 5-day week with a flexible 8-hour day, and (3) 5-day week with a structured 8-hour day. A study was conducted and the following data obtained of the average number of days missed:

<table>
<thead>
<tr>
<th>Work plan</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>9</td>
<td>6.2</td>
<td>10.1</td>
</tr>
<tr>
<td>Number</td>
<td>100</td>
<td>85</td>
<td>90</td>
</tr>
</tbody>
</table>

We also have $s^2 = 110.15$. Test the hypothesis that the work plan has no effect on the absenteeism. The 95% critical value for an $F$ distribution with 2 and 272 degrees of freedom is 3.03.

Solution:

\[
C = \begin{bmatrix}
0 & 1 & -1 & 0 \\
0 & 0 & 1 & -1 \\
\end{bmatrix}
\]

\[
Cb = \begin{bmatrix}
2.8 \\
-3.9 \\
\end{bmatrix}
\]

\[
C(X^TX)^c C^T = \begin{bmatrix}
0.0218 & -0.0118 \\
-0.0118 & 0.0229 \\
\end{bmatrix}
\]

\[
\frac{(Cb)^T [C(X^TX)^c C^T]^{-1} Cb}{s^2} = 3.2.
\]

Therefore we reject the null hypothesis.
3. For the one-way classification model, given the expressions for $X^T y$ and $(X^T X)^c$ in the lecture slides, show that

$$SS_{Reg} = \sum_{i=1}^{k} (\bar{y}_i)^2 n_i.$$  

**Solution:**

$$SS_{Reg} = y^T X (X^T X)^c X^T y$$

$$= (X^T y)^T b$$

$$= \left[ \sum_{ij} y_{ij} \sum_{j} y_{1j} \sum_{j} y_{2j} \ldots \sum_{j} y_{kj} \right] \left[ \begin{array}{c} 0 \\ \bar{y}_1 \\ \bar{y}_2 \\ \vdots \\ \bar{y}_k \end{array} \right]$$

$$= \sum_i \left( \sum_j y_{ij} \bar{y}_i \right)$$

$$= \sum_i (\bar{y}_i)^2 n_i.$$  

4. We are interested in looking at the effect of breed and diet on the milk yield of cows. A study is conducted and the following data obtained:

<table>
<thead>
<tr>
<th>Breed</th>
<th>Diet 1</th>
<th>Diet 2</th>
<th>Diet 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>18.8</td>
<td>16.7</td>
<td>19.8</td>
</tr>
<tr>
<td></td>
<td>21.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>22.3</td>
<td>15.9</td>
<td>21.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>19.2</td>
<td></td>
</tr>
</tbody>
</table>

(a) Express this as a two-factor model with no interaction in matrix form.  

**Solution:** $y = X \beta + \varepsilon$, where

$$y = \begin{bmatrix} 18.8 \\ 21.2 \\ 16.7 \\ 19.8 \\ 23.9 \\ 22.3 \\ 15.9 \\ 19.2 \\ 21.8 \end{bmatrix}, \quad X = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}, \quad \beta = \begin{bmatrix} \mu \\ \tau_1 \\ \tau_2 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix}$$

and $\varepsilon$ is as expected.
(b) Express this as a two-factor model with interaction in matrix form.

**Solution:** \( y = X\beta + \varepsilon \), where

\[
\begin{bmatrix}
18.8 \\
21.2 \\
16.7 \\
19.8 \\
23.9 \\
22.3 \\
15.9 \\
19.2 \\
21.8
\end{bmatrix} =
\begin{bmatrix}
1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\
1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
\begin{bmatrix}
\mu \\
\tau_1 \\
\tau_2 \\
\beta_1 \\
\beta_2 \\
\beta_3 \\
\xi_{11} \\
\xi_{12} \\
\xi_{13} \\
\xi_{21} \\
\xi_{22} \\
\xi_{23}
\end{bmatrix}
\]

and \( \varepsilon \) is as expected.

(c) Express the hypothesis that there is no interaction in terms of your parameters. Eliminate any redundancies.

**Solution:** We know that we require \((a - 1)(b - 1) = 2\) hypotheses, so we take the obviously non-redundant hypotheses

\[
\begin{align*}
(\xi_{11} - \xi_{12}) - (\xi_{21} - \xi_{22}) &= 0 \\
(\xi_{11} - \xi_{13}) - (\xi_{31} - \xi_{33}) &= 0.
\end{align*}
\]