1. For problem 1, please see the solution of exam 1 in 2012 page 35-51.
   Available at: https://dnett.github.io/S510/exam1sol2012.pdf.

2. Prove that $\mathcal{C}(\boldsymbol{X}) = \mathcal{C}(\boldsymbol{X}\boldsymbol{B}^{-1})$:

$$
\begin{aligned}
\boldsymbol{a} \in \mathcal{C}(\boldsymbol{X}) &\iff \underset{n\times 1}{\boldsymbol{a}} = \boldsymbol{X}\boldsymbol{b} && \text{for some } \underset{p\times 1}{\boldsymbol{b}} \\
&\iff \boldsymbol{a} = \boldsymbol{X} \underset{p\times p}{\boldsymbol{I}} \boldsymbol{b} && \text{for some } \boldsymbol{b}\text{x} \\
&\iff \boldsymbol{a} = \boldsymbol{X}\boldsymbol{B}^{-1} \underbrace{\boldsymbol{B}\boldsymbol{b}}_{p\times 1} && \text{treat as } \boldsymbol{X}\boldsymbol{B}^{-1} \text{ product a } p\times 1 \text{ vector} \\
&\implies \boldsymbol{a} \in \mathcal{C}(\boldsymbol{P_X})
\end{aligned}
$$

So $\mathcal{C}(\boldsymbol{X}) \subseteq \mathcal{C}(\boldsymbol{X}\boldsymbol{B}^{-1})$.
Then similarly,

$$
\begin{aligned}
\boldsymbol{g} \in \mathcal{C}(\boldsymbol{X}\boldsymbol{B}^{-1}) &\iff \boldsymbol{g} = \boldsymbol{X}\boldsymbol{B}^{-1}\boldsymbol{h} && \text{for some } p\times 1 \text{ vector } \boldsymbol{h} \\
&\iff \boldsymbol{g} = \boldsymbol{X} \underbrace{\boldsymbol{B}^{-1}\boldsymbol{h}}_{p\times 1} && \text{treat as } \boldsymbol{X} \text{ product a } p\times 1 \text{ vector} \\
&\implies \boldsymbol{g} \in \mathcal{C}(\boldsymbol{X})
\end{aligned}
$$

So $\mathcal{C}(\boldsymbol{X}\boldsymbol{B}^{-1}) \subseteq \mathcal{C}(\boldsymbol{X})$.
According to the results above, $\mathcal{C}(\boldsymbol{X}) = \mathcal{C}(\boldsymbol{X}\boldsymbol{B}^{-1})$.

3. Prove that $\boldsymbol{P_X} = \boldsymbol{P_W}$, i.e. $\boldsymbol{P_X} - \boldsymbol{P_W} = \boldsymbol{0}$:

$$
\text{Key: } \mathcal{C}(\boldsymbol{X}) = \mathcal{C}(\boldsymbol{W}) \implies \begin{cases} \boldsymbol{X} = \boldsymbol{W}\boldsymbol{A} \text{ for some } \boldsymbol{A} \\ \boldsymbol{W} = \boldsymbol{X}\boldsymbol{B} \text{ for some } \boldsymbol{B} \end{cases}
$$

From homework 1 problem 7 (a), we also know that

$$
\boldsymbol{P_X} - \boldsymbol{P_W} = \boldsymbol{0} \iff (\boldsymbol{P_X} - \boldsymbol{P_W})'(\boldsymbol{P_X} - \boldsymbol{P_W}) = \boldsymbol{0}
$$

So it is equivalent to prove $(\boldsymbol{P_X} - \boldsymbol{P_W})'(\boldsymbol{P_X} - \boldsymbol{P_W}) = \boldsymbol{0}$.

$$
\begin{aligned}
(\boldsymbol{P_X} - \boldsymbol{P_W})'(\boldsymbol{P_X} - \boldsymbol{P_W}) &= (\boldsymbol{P}'_{\boldsymbol{X}} - \boldsymbol{P}'_{\boldsymbol{W}})(\boldsymbol{P_X} - \boldsymbol{P_W}) \text{ by the property of transpose operation} \\
&= (\boldsymbol{P_X} - \boldsymbol{P_W})(\boldsymbol{P_X} - \boldsymbol{P_W}) \\
&= \boldsymbol{P_X}\boldsymbol{P_X} - \boldsymbol{P_X}\boldsymbol{P_W} - \boldsymbol{P_W}\boldsymbol{P_X} + \boldsymbol{P_W}\boldsymbol{P_W} \\
&= \boldsymbol{P_X} - \boldsymbol{P_X}\boldsymbol{W}(\boldsymbol{W}'\boldsymbol{W})^{-}\boldsymbol{W}' - \boldsymbol{P_W}\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}' + \boldsymbol{P_W} \\
&= \boldsymbol{P_X} - \underbrace{\boldsymbol{P_X}\boldsymbol{X}}\boldsymbol{B}(\boldsymbol{W}'\boldsymbol{W})^{-}\boldsymbol{W}' - \underbrace{\boldsymbol{P_W}\boldsymbol{W}}\boldsymbol{A}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}' + \boldsymbol{P_W} \\
&= \boldsymbol{P_X} - \boldsymbol{X}\boldsymbol{B}(\boldsymbol{W}'\boldsymbol{W})^{-}\boldsymbol{W}' - \boldsymbol{W}\boldsymbol{A}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}' + \boldsymbol{P_W} \\
&= \boldsymbol{P_X} - \boldsymbol{P_W} - \boldsymbol{P_X} + \boldsymbol{P_W} \\
&= \boldsymbol{0}
\end{aligned}
$$

Therefore the equivalent statement $\boldsymbol{P_X} = \boldsymbol{P_W}$ holds.

4. (a) The corresponding design matrix is

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

(b) Recall that a scalar $\boldsymbol{c}'\boldsymbol{\beta}$ is estimable if there exists a vector $\boldsymbol{a}'$ such that

$$\boldsymbol{c}'\boldsymbol{\beta} = \boldsymbol{a}' \, \mathrm{E}(\boldsymbol{y}),$$

since $\mathrm{E}(\boldsymbol{y}) = \boldsymbol{X}\boldsymbol{\beta}$. Hence, $\tau_1 - \tau_2$ is estimable if it can be written as a linear combination of the expected value of $\boldsymbol{y}$. Put

$$\boldsymbol{a}' = (1, 1, -1, -1, 1, 1, -1, -1)/4.$$

Then

$$\begin{aligned} \boldsymbol{a}' \, \mathrm{E}(\boldsymbol{y}) &= \frac{1}{4} \sum_{i=1}^{2} \sum_{k=1}^{2} \big[\mathrm{E}(y_{i1k}) - \mathrm{E}(y_{i2k}) \big] \\ &= \frac{1}{4} \sum_{i=1}^{2} \sum_{k=1}^{2} \big[\mathrm{E}(\mu + \lambda_i + \tau_1 + \varepsilon_{i1k}) - \mathrm{E}(\mu + \lambda_i + \tau_2 + \varepsilon_{i2k}) \big] \\ &= \frac{1}{4} \sum_{i=1}^{2} \sum_{k=1}^{2} \big[(\mu + \lambda_i + \tau_1 + 0) - (\mu + \lambda_i + \tau_2 + 0) \big] \\ &= \frac{1}{4} (2 \cdot 2) \big[\tau_1 - \tau_2 \big] \\ &= \tau_1 - \tau_2. \end{aligned}$$

Alternatively, and more simply, $\mathrm{E}(y_{111}) - \mathrm{E}(y_{121}) = \tau_1 - \tau_2$, so $\tau_1 - \tau_2$ is estimable.

(c) Notice that $\mathrm{rank}(\boldsymbol{X}) = 3$, so that $\boldsymbol{X}^*$ needs to be 8×3 to have full column rank and the same column space as $\boldsymbol{X}$. One possible choice is

$$\boldsymbol{X}^* = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & 1 & -1 \\ 1 & -1 & 1 \\ 1 & -1 & 1 \\ 1 & -1 & -1 \\ 1 & -1 & -1 \end{bmatrix}.$$

(d) Let $\boldsymbol{\beta}^* = (\beta_1^*, \beta_2^*, \beta_3^*)'$. Using $\boldsymbol{X}^*$ as given in part (c), equating expected values gives

$$\begin{aligned} \beta_1^* + \beta_2^* + \beta_3^* &= \mu + \lambda_1 + \tau_1, \\ \beta_1^* + \beta_2^* - \beta_3^* &= \mu + \lambda_1 + \tau_2, \\ \beta_1^* - \beta_2^* + \beta_3^* &= \mu + \lambda_2 + \tau_1, \\ \beta_1^* - \beta_2^* - \beta_3^* &= \mu + \lambda_2 + \tau_2. \end{aligned}$$

Then

$$
\begin{aligned}
2\beta_1^* &= 2\mu + \lambda_1 + \lambda_2 + \tau_1 + \tau_2, \\
2\beta_2^* &= \lambda_1 - \lambda_2, \\
2\beta_3^* &= \tau_1 - \tau_2,
\end{aligned}
$$

which implies

$$
\begin{aligned}
\beta_1^* &= \mu + (\lambda_1 + \lambda_2)/2 + (\tau_1 + \tau_2)/2, \\
\beta_2^* &= (\lambda_1 - \lambda_2)/2, \\
\beta_3^* &= (\tau_1 - \tau_2)/2.
\end{aligned}
$$

(e) Since $\tau_1 - \tau_2 = 2\beta_3^* = \begin{pmatrix}0, & 0, & 2\end{pmatrix} \boldsymbol{\beta}^*$,

$$
\begin{aligned}
\left(\widehat{\tau_1 - \tau_2}\right)_{\text{OLS}} &= \begin{pmatrix}0, & 0, & 2\end{pmatrix} \hat{\boldsymbol{\beta}}^*_{\text{OLS}} \\
&= \begin{pmatrix}0, & 0, & 2\end{pmatrix} (\boldsymbol{X}^{*\prime}\boldsymbol{X}^*)^{-1}\boldsymbol{X}^{*\prime}\boldsymbol{y} \\
&= \begin{pmatrix}0, & 0, & 2\end{pmatrix} \begin{pmatrix} 8 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 8 \end{pmatrix}^{-1} \boldsymbol{X}^{*\prime}\boldsymbol{y} \\
&= \begin{pmatrix}0, & 0, & 2\end{pmatrix} \begin{pmatrix} 1/8 & 0 & 0 \\ 0 & 1/8 & 0 \\ 0 & 0 & 1/8 \end{pmatrix} \boldsymbol{X}^{*\prime}\boldsymbol{y} \\
&= \begin{pmatrix}0, & 0, & 1/4\end{pmatrix} \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \end{pmatrix} \boldsymbol{y} \\
&= \frac{1}{4} \begin{pmatrix} 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \end{pmatrix} \begin{pmatrix} y_{111} \\ y_{112} \\ y_{121} \\ y_{122} \\ y_{211} \\ y_{212} \\ y_{221} \\ y_{222} \end{pmatrix} \\
&= \frac{1}{4} \sum_{i=1}^{2} \sum_{k=1}^{2} [y_{i1k} - y_{i2k}] \\
&= \bar{y}_{\cdot 1 \cdot} - \bar{y}_{\cdot 2 \cdot}.
\end{aligned}
$$