Applied Data Analysis

Exercise Sheet 3 - Solutions

Exercise 11

Let $i \in \{1, \ldots, p\}$.

$\Longrightarrow$: Let β_i be identifiable and suppose $B_i \in \operatorname{span}\big(\{B_j \,|\, j \in \{1, \ldots, p\} \backslash \{i\}\}\big)$.
Then, for $j \in \{1, \ldots, p\} \backslash \{i\}$, there exists $\alpha_j \in \mathbb{R}$ with

$$(1) \qquad B_i = \sum_{j \in \{1,\ldots,p\} \backslash \{i\}} \alpha_j B_j \iff \sum_{j \in \{1,\ldots,p\} \backslash \{i\}} \alpha_j B_j - B_i = \mathbf{0}$$

Setting $\alpha_i := -1$ and $\boldsymbol{\alpha} := (\alpha_1, \ldots, \alpha_p)'$, we obtain:

$$(2) \qquad B\boldsymbol{\alpha} = \sum_{j=1}^{p} \alpha_j B_j \overset{(1)}{=} \mathbf{0} \quad , \quad \text{i.e. } \operatorname{Ker}(B) \supsetneq \{\mathbf{0}\} \text{ since } \boldsymbol{\alpha} \neq \mathbf{0}$$

Now, let $\boldsymbol{\beta}^{(1)} \in \mathbb{R}^p$ and $\boldsymbol{\beta}^{(2)} := \boldsymbol{\beta}^{(1)} + \boldsymbol{\alpha}$. Then, on the one hand, considering the given linear model for $\boldsymbol{\beta}^{(1)}$ and $\boldsymbol{\beta}^{(2)}$, respectively, we get:

$$(3) \quad \mathrm{E}_{\boldsymbol{\beta}^{(2)}}(\boldsymbol{Y}) \underset{\text{I.4.5}}{\overset{\text{Ass.}}{=}} B\boldsymbol{\beta}^{(2)} \overset{\text{Def.}}{=} B\big(\boldsymbol{\beta}^{(1)} + \boldsymbol{\alpha}\big) = B\boldsymbol{\beta}^{(1)} + B\boldsymbol{\alpha} \overset{(2)}{=} B\boldsymbol{\beta}^{(1)} \underset{\text{I.4.5}}{\overset{\text{Ass.}}{=}} \mathrm{E}_{\boldsymbol{\beta}^{(1)}}(\boldsymbol{Y})$$

On the other hand, setting $g(\boldsymbol{x}) := \boldsymbol{e_{i,p}}'\boldsymbol{x}$, $\boldsymbol{x} \in \mathbb{R}^p$, we obtain:

$$(4) \qquad g\big(\boldsymbol{\beta}^{(2)}\big) \overset{\text{Def.}}{=} \boldsymbol{e_{i,p}}'\boldsymbol{\beta}^{(2)} = \beta_i^{(2)} \overset{\text{Def.}}{=} \beta_i^{(1)} - 1 \neq \beta_i^{(1)} = \boldsymbol{e_{i,p}}'\boldsymbol{\beta}^{(1)} \overset{\text{Def.}}{=} g\big(\boldsymbol{\beta}^{(1)}\big)$$

Thus, by Definition I.4.16, $\beta_i = g(\boldsymbol{\beta})$ is *not* identifiable leading to a contradiction, and therefore $B_i \notin \operatorname{span}\big(\{B_j \,|\, j \in \{1, \ldots, p\} \backslash \{i\}\}\big)$.

$\Longleftarrow$**:** Now, let $B_i \notin \operatorname{span}\big(\{B_j \,\big|\, j \in \{1,\dots,p\}\backslash\{i\}\}\big)$.

Then, a representation of B_i as linear combination of B_j, $j \in \{1,\dots,p\}\backslash\{i\}$ as in (1) does *not* exits. Thus for each $\boldsymbol{\alpha} = (\alpha_1,\dots,\alpha_p)' \in \mathbb{R}^p$, we obtain:

$$B\,\boldsymbol{\alpha} \;=\; \sum_{j=1}^{p} \alpha_j\, B_j \;=\; \mathbf{0} \quad\Longrightarrow\quad \alpha_i \;=\; 0 \tag{5}$$

Now, let $\boldsymbol{\beta}^{(1)}, \boldsymbol{\beta}^{(2)} \in \mathbb{R}^p$ with

$$B\,\boldsymbol{\beta}^{(1)} \underset{\text{I.4.5}}{\overset{\text{Ass.}}{=}} \mathrm{E}_{\boldsymbol{\beta}^{(1)}}(\boldsymbol{Y}) \;=\; \mathrm{E}_{\boldsymbol{\beta}^{(2)}}(\boldsymbol{Y}) \underset{\text{I.4.5}}{\overset{\text{Ass.}}{=}} B\,\boldsymbol{\beta}^{(2)}$$

leading to

$$B\left(\boldsymbol{\beta}^{(1)} - \boldsymbol{\beta}^{(2)}\right) \;=\; \mathbf{0} \tag{6}$$

Using the function g defined above, we obtain:

$$g\big(\boldsymbol{\beta}^{(1)}\big) \overset{\text{Def.}}{=} \beta_i^{(1)} \overset{(5),(6)}{=} \beta_i^{(2)} \overset{\text{Def.}}{=} g\big(\boldsymbol{\beta}^{(1)}\big) \tag{7}$$

Thus, by Definition I.4.16, $\beta_i = g(\boldsymbol{\beta})$ is identifiable.

Exercise 12

(a) First, by the assumption $\operatorname{rank}(B) = p \leq n$, we get:

(1) $$B'\,B \;\in\; \mathbb{R}^{p\times p} \quad \text{with} \quad \operatorname{rank}(B'\,B) \;\overset{\text{I.1.4,(6)}}{=}\; \operatorname{rank}(B) \;=\; p$$

Thus $B'\,B$ is regular (as already stated in Remark I.4.4).

Further, using $B = Q\,R$ with $Q \in \mathbb{R}^{n\times p}$ consisting of p orthonormal column vectors, we obtain:

(2) $$B'\,B \;=\; (Q\,R)'\,Q\,R \;=\; R'\,\underbrace{Q'\,Q}_{=\,I_p}\,R \;=\; R'\,R$$

Thus, $R'\,R = B'\,B$ is regular too, implying

(3) $$0 \;\neq\; \det(R'\,R) \;\overset{\text{I.1.4,(7)}}{=}\; \det(R')\,\det(R)$$

By (3), $R \in \mathbb{R}^{p\times p}$ is also regular. Then, we obtain:

(4) $$\hat{\boldsymbol{\beta}} \;\overset{\text{I.4.12}}{=}\; (B'\,B)^{-1}\,B'\,\boldsymbol{Y} \;\underset{B\,=\,Q\,R}{\overset{(2)}{=}}\; (R'\,R)^{-1}\,(Q\,R)'\,\boldsymbol{Y}$$
$$\overset{R\text{ reg.}}{=}\; R^{-1}\,\underbrace{(R')^{-1}\,R'}_{=\,I_p}\,Q'\,\boldsymbol{Y} \;\overset{\text{I.4.12}}{=}\; R^{-1}\,Q'\,\boldsymbol{Y}$$

(b) By the setting of part (b), we consider the 3×2-matrix

(5) $$A \;=\; [\,\boldsymbol{a_1}\,\boldsymbol{a_2}\,] \quad \text{with} \quad \boldsymbol{a_1} \;=\; \begin{pmatrix}1\\1\\0\end{pmatrix} \quad \text{and} \quad \boldsymbol{a_2} \;=\; \begin{pmatrix}1\\2\\1\end{pmatrix}$$

Then, following the Gram-Schmidt-procedure described on page 3 of Exercise Sheet 3, we obtain:

(6) $$\boldsymbol{e_1} \;\overset{\text{Def.}}{=}\; \frac{\boldsymbol{u_1}}{||\boldsymbol{u_1}||} \;\overset{\text{Def.}}{=}\; \frac{\boldsymbol{a_1}}{||\boldsymbol{a_1}||} \;\overset{(5)}{=}\; \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\\0\end{pmatrix}$$

(7) $$\boldsymbol{e_2} \;\overset{\text{Def.}}{=}\; \frac{\boldsymbol{u_2}}{||\boldsymbol{u_2}||} \;\overset{\text{Def.}}{=}\; \frac{\boldsymbol{a_2} - (\boldsymbol{a_2}'\,\boldsymbol{e_1})\,\boldsymbol{e_1}}{||\boldsymbol{a_2} - (\boldsymbol{a_2}'\,\boldsymbol{e_1})\,\boldsymbol{e_1}||}$$

Calculation of $\boldsymbol{u_2}$ and $||\boldsymbol{u_2}||$ yields:

(8) $$\boldsymbol{u_2} \;\overset{\text{Def.}}{=}\; \boldsymbol{a_2} - (\boldsymbol{a_2}'\,\boldsymbol{e_1})\,\boldsymbol{e_1} \;\overset{(5),(6)}{=}\; \begin{pmatrix}1\\2\\1\end{pmatrix} - \frac{1}{\sqrt{2}}\,(1\,2\,1)\begin{pmatrix}1\\1\\0\end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\\0\end{pmatrix}$$
$$=\; \begin{pmatrix}1\\2\\1\end{pmatrix} - \frac{3}{2}\begin{pmatrix}1\\1\\0\end{pmatrix} \;=\; \frac{1}{2}\begin{pmatrix}-1\\1\\2\end{pmatrix}$$

(9) $$||\boldsymbol{u_2}|| \;\overset{(8)}{=}\; \frac{1}{2}\sqrt{(-1)^2 + 1^2 + 2^2} \;=\; \frac{\sqrt{6}}{2}$$

Thus, we obtain:

(10) $$\boldsymbol{e_2} \;\overset{(7),(8),(9)}{=}\; \frac{1}{\sqrt{6}}\begin{pmatrix}-1\\1\\2\end{pmatrix}$$

Then, according to the Gram-Schmidt-procedure, the wanted matrix $Q \in \mathbb{R}^{3\times 2}$ is given by

(11) $$Q \;:=\; [\,\boldsymbol{e_1}\,\boldsymbol{e_2}\,] \;\overset{(6),(10)}{=}\; \frac{1}{\sqrt{6}} \begin{pmatrix} \sqrt{3} & -1 \\ \sqrt{3} & 1 \\ 0 & 2 \end{pmatrix}$$

Further, we get:

(12) $$\boldsymbol{a_1}'\,\boldsymbol{e_1} \;\overset{(5),(6)}{=}\; \frac{1}{\sqrt{2}}\,(1\,1\,0) \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \;=\; \frac{2}{\sqrt{2}}$$

(13) $$\boldsymbol{a_2}'\,\boldsymbol{e_1} \;\underset{\text{cf. (8)}}{\overset{(5),(6)}{=}}\; \frac{1}{\sqrt{2}}\,(1\,2\,1) \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \;=\; \frac{3}{\sqrt{2}}$$

(14) $$\boldsymbol{a_2}'\,\boldsymbol{e_2} \;\overset{(5),(10)}{=}\; \frac{1}{\sqrt{6}}\,(1\,2\,1) \begin{pmatrix} -1 \\ 1 \\ 2 \end{pmatrix} \;=\; \frac{3}{\sqrt{6}} \;=\; \frac{\sqrt{3}}{\sqrt{2}}$$

Then, according to the Gram-Schmidt-procedure, the triangular matrix $R \in \mathbb{R}^{2\times 2}$ is given by

(15) $$R \;:=\; \begin{pmatrix} \boldsymbol{a_1}'\,\boldsymbol{e_1} & \boldsymbol{a_2}'\,\boldsymbol{e_1} \\ 0 & \boldsymbol{a_2}'\,\boldsymbol{e_2} \end{pmatrix} \;\overset{(12),(13),(14)}{=}\; \frac{1}{\sqrt{2}} \begin{pmatrix} 2 & 3 \\ 0 & \sqrt{3} \end{pmatrix}$$

Thus, finally, we obtain the following QR-decomposition of the matrix A:

(16) $$A \;=\; Q\,R \quad \text{with} \quad Q \;=\; \frac{1}{\sqrt{6}} \begin{pmatrix} \sqrt{3} & -1 \\ \sqrt{3} & 1 \\ 0 & 2 \end{pmatrix} \quad \text{and} \quad R \;=\; \frac{1}{\sqrt{2}} \begin{pmatrix} 2 & 3 \\ 0 & \sqrt{3} \end{pmatrix}$$

(c) First, applying the rule for inverting a regular (2×2)-Matrix, we get:

(17) $$R^{-1} \;\overset{(15)}{=}\; \left(\frac{1}{\sqrt{2}} \begin{pmatrix} 2 & 3 \\ 0 & \sqrt{3} \end{pmatrix} \right)^{-1} \;=\; \frac{\sqrt{2}}{2\sqrt{3}} \begin{pmatrix} \sqrt{3} & -3 \\ 0 & 2 \end{pmatrix} \;=\; \frac{1}{\sqrt{6}} \begin{pmatrix} \sqrt{3} & -3 \\ 0 & 2 \end{pmatrix}$$

Then, applying the results of (a) and (b), we obtain the following least-square-estimate $\hat{\boldsymbol{\beta}}$ of $\boldsymbol{\beta}$ for the given observation $\boldsymbol{y} = (y_1, y_2, y_3)' = (1, 1, 1)'$:

(18) $$\begin{aligned} \hat{\boldsymbol{\beta}} \;&\overset{(a)}{=}\; R^{-1}\,Q'\,\boldsymbol{y} \\ &\overset{(17),(17)}{=}\; \frac{1}{\sqrt{6}} \begin{pmatrix} \sqrt{3} & -3 \\ 0 & 2 \end{pmatrix} \frac{1}{\sqrt{6}} \begin{pmatrix} \sqrt{3} & \sqrt{3} & 0 \\ -1 & 1 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \\ &=\; \frac{1}{6} \begin{pmatrix} 6 & 0 & -6 \\ -2 & 2 & 4 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \;=\; \frac{1}{6} \begin{pmatrix} 0 \\ 4 \end{pmatrix} \;=\; \begin{pmatrix} 0 \\ \frac{2}{3} \end{pmatrix} \end{aligned}$$

Exercise 13

The parts (b) – (f) of this Exercise could be solved by proceeding analogously to the corresponding derivations for the linear model of I.4.2 given in the Lecture.

Instead of this way, we can use the transformation considered in the solution of part (a) to trace back the linear model (1) given on page 2 of Exercise Sheet 3 to a linear model according to I.4.2 and apply the corresponding theorems of the Lecture already derived for this model.

(a) First, $\Sigma \in \mathbb{R}^{n\times n}$ is positive definite as a regular variance-covariance matrix. Then, by Exercise 2, there exists a positive definite – and thus, especially symmetric and regular – matrix $\Sigma^{1/2} \in \mathbb{R}^{n\times n}$ with $\Sigma = \Sigma^{1/2}\,\Sigma^{1/2}$.

We set $A := \Sigma^{-1/2} \overset{\text{Ex. 2}}{=} (\Sigma^{1/2})^{-1}$ and

$$\boldsymbol{Z} \;:=\; A\,\boldsymbol{Y} \;\overset{\text{Def.}}{=}\; \Sigma^{-1/2}\,\boldsymbol{Y} \tag{1}$$

Then, applying the given linear model $\boldsymbol{Y} = X\,\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ of this Exercise, we obtain:

$$\boldsymbol{Z} \;=\; \Sigma^{-1/2}\big(X\,\boldsymbol{\beta} + \boldsymbol{\varepsilon}\big) \;=\; B\,\boldsymbol{\beta} + \boldsymbol{\eta} \quad \text{with} \tag{2}$$

$$B \;:=\; \Sigma^{-1/2}\,X \;\in\; \mathbb{R}^{n\times p} \quad \text{and} \quad \eta \;:=\; \Sigma^{-1/2}\,\boldsymbol{\varepsilon} \tag{3}$$

According to the rules for expectations and variance-covariance matrices, we get:

$$\mathrm{E}(\boldsymbol{\eta}) \;\overset{\text{Def.}}{=}\; \mathrm{E}\big(\Sigma^{-1/2}\,\boldsymbol{\varepsilon}\big) \;\overset{\text{I.1.12}}{=}\; \Sigma^{-1/2}\,\underbrace{\mathrm{E}\big(\boldsymbol{\varepsilon}\big)}_{=\mathbf{0}} \;\overset{\text{Ass.}}{=}\; \mathbf{0} \tag{4}$$

$$\begin{aligned}\mathrm{Cov}(\boldsymbol{\eta}) \;&\overset{\text{Def.}}{=}\; \mathrm{Cov}\big(\Sigma^{-1/2}\,\boldsymbol{\varepsilon}\big) \;\underset{\Sigma^{-1/2}\text{ symm.}}{\overset{\text{I.1.16,(2)}}{=}}\; \Sigma^{-1/2}\,\mathrm{Cov}\big(\boldsymbol{\varepsilon}\big)\,\Sigma^{-1/2} \\ &\overset{\text{Ass.}}{=}\; \Sigma^{-1/2}\,\sigma^2\,\Sigma\,\Sigma^{-1/2} \;\overset{\text{Ex. 2}}{=}\; \sigma^2\,I_n\end{aligned} \tag{5}$$

Thus, (2) together with (3) – (5) yields a (ordinary) linear model for $\boldsymbol{Z}$ with design matrix B as given by I.4.2.

The two minimization problems given on one hand for the linear model (2) (with objective funktion ψ given in I.4.8) and on the other hand given on page 2 of Exercise Sheet 3 have the same solutions, as will be shown next.

For that purpose, in the sequel, let $\boldsymbol{y} \in \mathbb{R}^n$ denote a realization of $\boldsymbol{Y}$ and $\boldsymbol{z} := \Sigma^{-1/2}\,\boldsymbol{y}$ denote the corresponding realization of $\boldsymbol{Z} = \Sigma^{-1/2}\,\boldsymbol{Y}$. Then, for $\boldsymbol{\beta} \in \mathbb{R}^p$, we obtain:

$$\begin{aligned}\psi(\boldsymbol{\beta}) \;&\underset{\text{(I.4.8)}}{\overset{\text{Def.}}{=}}\; ||\,\boldsymbol{z} - B\,\boldsymbol{\beta}\,||^2 \;\overset{\text{Def. }\boldsymbol{z},(3)}{=}\; ||\,\Sigma^{-1/2}\,\boldsymbol{y} - \Sigma^{-1/2}\,X\,\boldsymbol{\beta}\,||^2 \\ &=\; \big(\Sigma^{-1/2}\,(\boldsymbol{y} - X\,\boldsymbol{\beta})\big)'\,\big(\Sigma^{-1/2}\,(\boldsymbol{y} - X\,\boldsymbol{\beta})\big) \;\underset{\text{symm.}}{\overset{\Sigma^{-1/2}}{=}}\; (\boldsymbol{y} - X\,\boldsymbol{\beta})'\,\Sigma^{-1/2}\,\Sigma^{-1/2}\,(\boldsymbol{y} - X\,\boldsymbol{\beta}) \\ &\overset{\text{Ex. 2}}{=}\; (\boldsymbol{y} - X\,\boldsymbol{\beta})'\,\Sigma^{-1}\,(\boldsymbol{y} - X\,\boldsymbol{\beta}) \;\overset{\text{Def.}}{=}\; ||\,\boldsymbol{y} - X\,\boldsymbol{\beta}\,||_\Sigma^2 \;\overset{\text{Def.}}{=}\; \widetilde{\psi}(\boldsymbol{\beta})\end{aligned} \tag{6}$$

By (6), the two minimization problems

$$\text{(OLS)} \quad \psi(\boldsymbol{\beta}) \;\longrightarrow\; \min_{\boldsymbol{\beta}\in\mathbb{R}^p} \quad \text{and} \quad \text{(GLS)} \quad \widetilde{\psi}(\boldsymbol{\beta}) \;\longrightarrow\; \min_{\boldsymbol{\beta}\in\mathbb{R}^p} \tag{7}$$

are equivalent.

Thus, in the sequel, each statement concerning a solution of the minimization problem (GLS) considered in this Exercise can be traced back to the corresponding statement of the minimization problem (OLS) given by the Lecture.

(b) Each solution $\widehat{\boldsymbol{\beta}}$ of (GLS) and thus also of (OLS) has to satisfy the following normal equations (according to Theorem I.4.9):

$$B'\,B\,\widehat{\boldsymbol{\beta}} \;=\; B'\,\boldsymbol{z} \;\underset{\text{Def. }\boldsymbol{z}}{\overset{(3)}{\Longleftrightarrow}}\; (\Sigma^{-1/2}\,X)'\,\Sigma^{-1/2}\,X\,\widehat{\boldsymbol{\beta}} \;=\; (\Sigma^{-1/2}\,X)'\,\Sigma^{-1/2}\,\boldsymbol{y}$$

$$\underset{\text{symm.}}{\overset{\Sigma^{-1/2}}{\Longleftrightarrow}}\; X'\,\Sigma^{-1/2}\,\Sigma^{-1/2}\,X\,\widehat{\boldsymbol{\beta}} \;=\; X'\,\Sigma^{-1/2}\,\Sigma^{-1/2}\,\boldsymbol{y} \;\overset{\text{Ex. 2}}{\Longleftrightarrow}$$

$$X'\,\Sigma^{-1}\,X\,\widehat{\boldsymbol{\beta}} \;=\; X'\,\Sigma^{-1}\,\boldsymbol{y} \tag{8}$$

(c) First, the assumption $\operatorname{rank}(X) = p \le n$ implies:

$$\operatorname{rank}\big(X'\,\Sigma^{-1}\,X\big) \overset{\text{Ex. 2}}{=} \operatorname{rank}\big(X'\,\Sigma^{-1/2}\,\Sigma^{-1/2}\,X\big) \underset{\text{symm.}}{\overset{\Sigma^{-1/2}}{=}} \operatorname{rank}\big((\Sigma^{-1/2}\,X)'\,\Sigma^{-1/2}\,X\big) \tag{9}$$

$$\overset{\text{I.1.4,(4),(6)}}{=} \operatorname{rank}\big((\Sigma^{-1/2}\,X)'\big) \underset{\text{symm.}}{\overset{\Sigma^{-1/2}}{=}} \operatorname{rank}\big(X'\,\Sigma^{-1/2}\big) \underset{\Sigma^{-1/2}\text{ reg.}}{\overset{\text{I.1.4,(4),(5)}}{=}} \operatorname{rank}(X) \overset{\text{Ass.}}{=} p$$

By (9), $X'\,\Sigma^{-1}\,X \in \mathbb{R}^{p\times p}$ is regular. Thus, by (8), we obtain immediately that the solution $\widehat{\boldsymbol{\beta}}$ of (GLS) is uniquely determined and given by

$$\widehat{\boldsymbol{\beta}} \;=\; (X'\,\Sigma^{-1}\,X)^{-1}\,X'\,\Sigma^{-1}\,\boldsymbol{Y} \tag{10}$$

Alternatively, this representation could be derived by applying Corollary I.4.12 to the minimization problem (OLS) and using the definitions of B and $\boldsymbol{Z}$.

(d) First, for $\boldsymbol{\beta} \in \mathbb{R}^p$, we get by (10) (analogously to the corresponding derivations for the linear model considered in the Lecture):

$$\widehat{\boldsymbol{\beta}} \overset{(10)}{=} (X'\,\Sigma^{-1}\,X)^{-1}\,X'\,\Sigma^{-1}\,\boldsymbol{Y} \overset{\text{Model (1)}}{=} (X'\,\Sigma^{-1}\,X)^{-1}\,X'\,\Sigma^{-1}\,(X\,\boldsymbol{\beta} + \boldsymbol{\varepsilon}) \tag{11}$$

$$= \underbrace{(X'\,\Sigma^{-1}\,X)^{-1}\,X'\,\Sigma^{-1}\,X}_{=\,I_p}\,\boldsymbol{\beta} + \underbrace{(X'\,\Sigma^{-1}\,X)^{-1}\,X'\,\Sigma^{-1}}_{=:\,M}\,\boldsymbol{\varepsilon} \;=\; \boldsymbol{\beta} + M\,\boldsymbol{\varepsilon}$$

Then, applying the rules for expectations and variance-covariance matrices, we obtain for $\boldsymbol{\beta} \in \mathbb{R}^p$:

$$\mathrm{E}(\widehat{\boldsymbol{\beta}}) \overset{(11)}{=} \mathrm{E}(\boldsymbol{\beta} + M\,\boldsymbol{\varepsilon}) \overset{\text{I.1.12}}{=} \boldsymbol{\beta} + M\,\underbrace{\mathrm{E}(\boldsymbol{\varepsilon})}_{=\,\mathbf{0}} \overset{\text{Ass.}}{=} \boldsymbol{\beta} \tag{12}$$

$$\mathrm{Cov}(\widehat{\boldsymbol{\beta}}) \overset{(11)}{=} \mathrm{Cov}(\boldsymbol{\beta} + M\,\boldsymbol{\varepsilon}) \overset{\text{I.1.16}}{=} M\,\underbrace{\mathrm{Cov}(\boldsymbol{\varepsilon})}_{=\,\sigma^2\,\Sigma}\,M' \tag{13}$$

$$\underset{\text{Ass.}}{\overset{\text{Def. }M}{=}} \sigma^2\,(X'\,\Sigma^{-1}\,X)^{-1}\,X'\,\underbrace{\Sigma^{-1}\,\Sigma}_{=\,I_n}\,\big((X'\,\Sigma^{-1}\,X)^{-1}\,X'\,\Sigma^{-1}\big)'$$

$$\underset{\text{symm.}}{\overset{\Sigma^{-1}}{=}} \sigma^2\,\underbrace{(X'\,\Sigma^{-1}\,X)^{-1}\,X'\,\Sigma^{-1}\,X}_{=\,I_p}\,\big(\underbrace{(X'\,\Sigma^{-1}\,X)^{-1}}_{\text{symm.}}\big)' \;=\; \sigma^2\,(X'\,\Sigma^{-1}\,X)^{-1}$$

By (12), under the given assumptions of this Exercise, (also) the GLS estimator $\widehat{\boldsymbol{\beta}}$ given by (10), is an unbiased estimator of $\boldsymbol{\beta} \in \mathbb{R}^p$.

According to the setting of this Exercise, for the following two parts, let us additionally assume

$$\boldsymbol{\varepsilon} \sim N_n(0, \sigma^2\,\Sigma) \tag{14}$$

(e) Applying I.2.11, by (14) and Model (1) of this Exercise Sheet with the corresponding assumptions, it holds:

$$\boldsymbol{Y} \;\sim\; N_n(X\,\boldsymbol{\beta}\,,\, \sigma^2\,\Sigma) \tag{15}$$

Thus, since Σ is regular, for a given realization $\boldsymbol{y} \in \mathbb{R}^n$ of $\boldsymbol{Y}$, by (15) and I.2.5 the likelihood function (i.e. density) $L^{\boldsymbol{Y}}(\bullet\,|\,\boldsymbol{y}) : \mathbb{R}^p \longrightarrow [0,\infty)$ of $\boldsymbol{Y}$ is given by

$$L^{\boldsymbol{Y}}(\boldsymbol{\beta}\,|\,\boldsymbol{y}) \;=\; \frac{1}{\sqrt{(2\pi)^n\,\det(\sigma^2\,\Sigma)}}\;\exp\left(-\frac{1}{2}\,(\boldsymbol{y}-X\,\boldsymbol{\beta})'\,(\sigma^2\,\Sigma)^{-1}\,(\boldsymbol{y}-X\,\boldsymbol{\beta})\right)\;,\;\boldsymbol{\beta}\in\mathbb{R}^p \tag{16}$$

Then, the corresponding log-likelihood function $l^{\boldsymbol{Y}}(\bullet\,|\,\boldsymbol{y}) : \mathbb{R}^p \longrightarrow \mathbb{R}$ of $\boldsymbol{Y}$ is given by

$$\begin{aligned}
(17)\quad l^{\boldsymbol{Y}}(\boldsymbol{\beta}\,|\,\boldsymbol{y}) \;&:=\; \ln\big(L^{\boldsymbol{Y}}(\boldsymbol{\beta}\,|\,\boldsymbol{y})\big)\\
&\overset{(16)}{=}\; -\ln\big((2\pi)^{n/2}\big) - \ln\Big(\big(\sigma^{2n}\,\det(\Sigma)\big)^{1/2}\Big) - \frac{1}{2}\,(\boldsymbol{y}-X\,\boldsymbol{\beta})'\,\frac{1}{\sigma^2}\,\Sigma^{-1}\,(\boldsymbol{y}-X\,\boldsymbol{\beta})\\
&\overset{\text{Ex. 2}}{=}\; \underbrace{-\frac{n}{2}\,\ln(2\pi) - n\,\ln(\sigma) - \frac{1}{2}\,\ln\big(\det(\Sigma)\big)}_{=:\,c(\sigma)} - \frac{1}{2\,\sigma^2}\,(\boldsymbol{y}-X\,\boldsymbol{\beta})'\,\Sigma^{-1/2}\,\Sigma^{-1/2}\,(\boldsymbol{y}-X\,\boldsymbol{\beta})\\
&\underset{\text{symm.}}{\overset{\Sigma^{-1/2}}{=}}\; c(\sigma) - \frac{1}{2\,\sigma^2}\,(\Sigma^{-1/2}\,\boldsymbol{y}-\Sigma^{-1/2}\,X\,\boldsymbol{\beta})'\,(\Sigma^{-1/2}\,\boldsymbol{y}-\Sigma^{-1/2}\,X\,\boldsymbol{\beta})\\
&\overset{(3)}{=}\; c(\sigma) - \frac{1}{2\,\sigma^2}\,(\boldsymbol{z}-B\,\boldsymbol{\beta})'\,(\boldsymbol{z}-B\,\boldsymbol{\beta})\;\;,\;\;\boldsymbol{\beta}\in\mathbb{R}^p\;\;,
\end{aligned}$$

with $\boldsymbol{z} := \Sigma^{-1/2}\,\boldsymbol{y}$ denoting the corresponding realization of $\boldsymbol{Z} = \Sigma^{-1/2}\,\boldsymbol{Y}$.

On the other hand, applying I.2.11, by (14) and (2) – (5), it holds:

$$\boldsymbol{Z} \;\sim\; N_n(B\,\boldsymbol{\beta}\,,\, \sigma^2\,I_n) \tag{18}$$

Then, by the derivation of the log-likelihood function for the linear model considered in the Lecture or by the derivation given above, replacing X by B and Σ by I_n, we obtain the following representation for the log-likelihood function $l^{\boldsymbol{Z}}(\bullet\,|\,\boldsymbol{z}) : \mathbb{R}^p \longrightarrow \mathbb{R}$ of $\boldsymbol{Z}$ for a given realization $\boldsymbol{z} \in \mathbb{R}^n$ of $\boldsymbol{Z}$:

$$\begin{aligned}
(19)\quad & l^{\boldsymbol{Z}}(\boldsymbol{\beta}\,|\,\boldsymbol{z})\\
&=\; -\frac{n}{2}\,\ln(2\pi) - n\,\ln(\sigma) - \frac{1}{2}\,\underbrace{\ln\big(\det(I_n)\big)}_{=\ln(1)=0} - \frac{1}{2\,\sigma^2}\,(\boldsymbol{z}-B\,\boldsymbol{\beta})'\,(\boldsymbol{z}-B\,\boldsymbol{\beta})\\
&\overset{\text{Def. }c(\sigma)}{=}\; c(\sigma) + \frac{1}{2}\,\ln\big(\det(\Sigma)\big) - \frac{1}{2\,\sigma^2}\,(\boldsymbol{z}-B\,\boldsymbol{\beta})'\,(\boldsymbol{z}-B\,\boldsymbol{\beta})\;\;,\;\;\boldsymbol{\beta}\in\mathbb{R}^p
\end{aligned}$$

By (17) and (19), for a given realization $\boldsymbol{y} \in \mathbb{R}^n$ of $\boldsymbol{Y}$ and $\boldsymbol{z} := \Sigma^{-1/2}\,\boldsymbol{y}$, the two log-likelihood functions $l^{\boldsymbol{Y}}(\bullet\,|\,\boldsymbol{y})$ and $l^{\boldsymbol{Z}}(\bullet\,|\,\boldsymbol{z})$ only distinguish by the contant term $\frac{1}{2}\,\ln\big(\det(\Sigma)\big)$ and thus, they achieve their maxima at the same parameter vectors $\boldsymbol{\beta} \in \mathbb{R}^p$.

Then, by Theorem I.4.31 and by part (a), (also) for the linear model considered in this Exercise, the maximum likelihood estimator coincides with the GLS estimator given by (10).

(f) Consider a matrix $X_0 \in \mathbb{R}^{n\times p_0}$ with $\mathrm{Im}(X_0) \subseteq \mathrm{Im}(X)$ and $\mathrm{rank}(X_0) = p_0$ (which implies $p_0 = \dim(\mathrm{Im}(X_0)) \le \dim(\mathrm{Im}(X)) = p \le n$). Further, let

$$\boldsymbol{Y} = X_0 \boldsymbol{\gamma} + \boldsymbol{\varepsilon} \tag{20}$$

with parameter vector $\gamma \in \mathbb{R}^{p_0}$ and (as before) error term $\boldsymbol{\varepsilon} \sim N_n(0, \sigma^2 \Sigma)$ denote the *reduced* (normal) GLS model associated with X_0.

Then (see (2),(3)),

$$\boldsymbol{Z} = B\boldsymbol{\beta} + \boldsymbol{\eta} = \Sigma^{-1/2} X \boldsymbol{\beta} + \boldsymbol{\eta} \tag{21}$$

represents the corresponding *full* normal OLS model associated with the design matrix $B = \Sigma^{-1/2} X$ (considered in the previous parts) and

$$\boldsymbol{Z} = B_0 \boldsymbol{\gamma} + \boldsymbol{\eta} = \Sigma^{-1/2} X_0 \boldsymbol{\gamma} + \boldsymbol{\eta} \tag{22}$$

represents the corresponding *reduced* normal OLS model associated with the design matrix $B_0 := \Sigma^{-1/2} X_0$.

By the assumed normality of $\boldsymbol{\varepsilon}$, I.2.11 and (4),(5), it holds:

$$\boldsymbol{\eta} \overset{(3)}{=} \Sigma^{-1/2} \boldsymbol{\varepsilon} \underset{(4),(5)}{\overset{\text{I.2.11}}{\sim}} N_n(\mathbf{0}\,,\, \sigma^2 I_n) \tag{23}$$

Then, all assumptions of Theorem I.4.39 (and thus, of Theorem I.4.40) are fulfilled by finally defining the projections accordingly as follows:

$$Q := B\,(B'\,B)^{-1}\,B' \underset{\Sigma^{-1/2}\text{ symm.}}{\overset{(3),(8)}{=}} \Sigma^{-1/2}\,X\,(X'\,\Sigma^{-1}\,X)^{-1}\,X'\,\Sigma^{-1/2} \tag{24}$$

$$Q_0 := B_0\,(B_0'\,B_0)^{-1}\,B_0' \underset{\Sigma^{-1/2}\text{ symm.}}{\overset{\text{Def. }B_0}{=}} \Sigma^{-1/2}\,X_0\,(X_0'\,\Sigma^{-1}\,X_0)^{-1}\,X_0'\,\Sigma^{-1/2} \tag{25}$$

Then, applying Theorem I.4.40, the null hypothesis

$$H_0 \,:\, \mathrm{E}(\boldsymbol{Y}) = X_0\,\boldsymbol{\gamma} \quad \text{for some} \quad \boldsymbol{\gamma} \in \mathbb{R}^{p_0} \tag{26}$$

is rejected at significance level $\alpha \in (0,1)$ if and only if

$$\begin{aligned} \frac{\boldsymbol{Y}'\,\Sigma^{-1/2}\,(Q-Q_0)\,\Sigma^{-1/2}\,\boldsymbol{Y}/(p-p_0)}{\boldsymbol{Y}'\,\Sigma^{-1/2}\,(I_n-Q)\,\Sigma^{-1/2}\,\boldsymbol{Y}/(n-p)} &\overset{(1)}{=} \frac{\boldsymbol{Z}'\,(Q-Q_0)\,\boldsymbol{Z}/(p-p_0)}{\boldsymbol{Z}'\,(I_n-Q)\,\boldsymbol{Z}/(n-p)} \\ &> F_{1-\alpha}(p-p_0, n-p) \end{aligned} \tag{27}$$

with $F_{1-\alpha}(p-p_0, n-p)$ denoting the $(1-\alpha)$-quantile of the F-distribution $F(p-p_0, n-p)$.

Notice that, due to the assumptions $p_0 = \mathrm{rank}(X_0) \le \mathrm{rank}(X) = p \le n$ in the given situation of this Exercise, by (24) and (25), it holds:

$$\begin{aligned} \mathrm{rank}(Q) &= \mathrm{rank}(B) = \mathrm{rank}(X) = p \\ \mathrm{rank}(Q_0) &= \mathrm{rank}(B_0) = \mathrm{rank}(X_0) = p_0 \end{aligned}$$