Prof. Dr. E. Cramer, Prof. Dr. M. Kateri
T. v. Bentum, M.Sc., L. Kaufmann, M.Sc.

RWTH Aachen, SS 2020

publication: May 14th solutions: May 21st

Applied Data Analysis

Exercise Sheet 3 - Solutions

Exercise 11

If the set $\{B_j, j \in \{1,\dots,p\}\backslash\{i\}\}$ is not linearly independent from B_i, then there exists $\boldsymbol{\alpha} = (\alpha_1,\dots,\alpha_p)$ such that $B_i = \sum_{k\neq i}\alpha_k B_k$. In particular, the kernel of B is non-trivial and there exists an $\boldsymbol{\alpha}$ with $\alpha_i \neq 0$ and $B\boldsymbol{\alpha} = \mathbf{0}$. Letting $\boldsymbol{\beta}_* = \boldsymbol{\beta} + \boldsymbol{\alpha}$ we hae $\beta_{*i} \neq \beta_i$ and

$$B\boldsymbol{\beta}_* = B\boldsymbol{\beta} + \mathbf{0} = B\boldsymbol{\beta}.$$

Thus, β_i is not identifiable.

If, on the other hand, the set is linearly independent from B_i, then there does not exist some $\boldsymbol{\alpha} = (\alpha_1,\dots,\alpha_p)$ such that $B_i = \sum_{k\neq i}\alpha_k B_k$. Thus, for any $\boldsymbol{\alpha}$ with $B\boldsymbol{\alpha} = \mathbf{0}$ we must have $\alpha_i = 0$. Therefore, if for some $\boldsymbol{\beta}_*$ we have

$$B\boldsymbol{\beta} = B\boldsymbol{\beta}_* \quad\Longleftrightarrow\quad B(\boldsymbol{\beta} - \boldsymbol{\beta}_*) = \mathbf{0},$$

then $\beta_i - \beta_{*i} = 0$ and β_i is identifiable.

Exercise 12

(a) Since by assumption $\operatorname{rank}(B) = p$ we get by Corollary I.4.12

$$\begin{aligned}\hat{\boldsymbol{\beta}} &= (B'B)^{-1}B'\boldsymbol{Y}\\ &= (R'Q'QR)^{-1}R'Q'\boldsymbol{Y}\\ &= (R'R)^{-1}R'Q'\boldsymbol{Y}\\ &= R^{-1}(R')^{-1}R'Q'\boldsymbol{Y}\\ &= R^{-1}Q'\boldsymbol{Y}.\end{aligned}$$

(b) Following the Gram-Schmidt procedure we get for $a_1 = \begin{pmatrix}1 & 1 & 0\end{pmatrix}'$

$$u_1 = \begin{pmatrix}1 & 1 & 0\end{pmatrix}', \qquad ||u_1|| = \sqrt{2}, \qquad e_1 = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1 & 0\end{pmatrix}'.$$

For $a_2 = \begin{pmatrix}1 & 2 & 1\end{pmatrix}'$ we get

$$u_2 = a_2 - \langle a_2, e_1\rangle e_1 = \begin{pmatrix}1\\2\\1\end{pmatrix} - \frac{3}{\sqrt{2}}\cdot\frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\\0\end{pmatrix} = \frac{1}{2}\begin{pmatrix}-1\\1\\2\end{pmatrix}$$

$$||u_2|| = \sqrt{\frac{1}{4} + \frac{1}{4} + 1} = \sqrt{\frac{3}{2}}$$

$$e_2 = \sqrt{\frac{1}{6}}\begin{pmatrix}-1\\1\\2\end{pmatrix}$$

and $Q = [e_1 \mid e_2]$.

Since

$$\langle a_1, e_1 \rangle = \sqrt{2}, \qquad \langle a_2, e_1 \rangle = \frac{3}{\sqrt{2}}, \qquad \langle a_2, e_2 \rangle = 3\sqrt{\frac{1}{6}}$$

we get

$$R = \frac{1}{\sqrt{2}} \begin{pmatrix} 2 & 3 \\ 0 & \sqrt{3} \end{pmatrix}$$

(c) Since by the elementary inversion formula for 2×2 matrices

$$R^{-1} = \frac{1}{\sqrt{3}} \cdot \frac{1}{\sqrt{2}} \begin{pmatrix} \sqrt{3} & -3 \\ 0 & 2 \end{pmatrix}$$

we get by (a) and (b)

$$\begin{aligned} \hat{\boldsymbol{\beta}} =& \frac{1}{\sqrt{6}} \begin{pmatrix} \sqrt{3} & -3 \\ 0 & 2 \end{pmatrix} \cdot \frac{1}{\sqrt{6}} \begin{pmatrix} \sqrt{3} & \sqrt{3} & 0 \\ -1 & 1 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \\ =& \frac{1}{6} \begin{pmatrix} 6 & 0 & -6 \\ -2 & 2 & 4 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \\ =& \begin{pmatrix} 0 \\ \frac{2}{3} \end{pmatrix}. \end{aligned}$$

Exercise 13

The exercises can be solved in the given order. In this case, the solution in (a) is – with the exception of a suitable transformation – essentially analogous to the proof for Theorem I.4.9 given in the lecture. Since the transformation holds similarly for (d), we will solve (d) first.

(d) First, let according to assumption $\Sigma \in \mathbb{R}^{n \times n}_{>0}$. Then, by Exercise 2 there exists an $A \in \mathbb{R}^{n \times n}_{>0}$ such that $\Sigma = A'A$, thus showing the hint.

Since A is regular, we can define

$$\boldsymbol{Z} = A^{-1}\boldsymbol{Y} = \Sigma^{-1/2}\boldsymbol{Y}, \quad \text{say},$$

and a similar transformation yields

$$B = \Sigma^{-1/2}X, \qquad \boldsymbol{\eta} = \Sigma^{-1/2}\boldsymbol{\varepsilon}.$$

Note that according to the rules for expectations and variances we have

$$\mathrm{E}(\boldsymbol{\eta}) = \Sigma^{-1/2}\,\mathrm{E}(\boldsymbol{\varepsilon}) = 0$$

and

$$\mathrm{Cov}(\boldsymbol{\eta}) = \mathrm{Cov}(\Sigma^{-1/2}\boldsymbol{\varepsilon}) = (\Sigma^{-1/2})'\mathrm{Cov}(\boldsymbol{\varepsilon})\Sigma^{-1/2} = \sigma^2\Sigma^{-1/2}\Sigma^{1/2}\Sigma^{1/2}\Sigma^{-1/2} = \sigma^2 I_p,$$

thus transforming the GLS model into an OLS model.

Note that the same arguments hold if we drop the assumption of regularity and replace the inverse by the corresponding Moore-Penrose inverse. This allows us in the following to essentially reduce the model of GLS to an equivalent OLS model.

(a)+(b) The object function $\psi(\boldsymbol{\beta})$ for the transformed OLS model is given by

$$\begin{aligned}\psi(\boldsymbol{\beta}) =&(\boldsymbol{Z} - B\boldsymbol{\beta})'(\boldsymbol{Z} - B\boldsymbol{\beta})\\ =&(\Sigma^{-1/2}\boldsymbol{Y} - \Sigma^{-1/2}X\boldsymbol{\beta})'(\Sigma^{-1/2}\boldsymbol{Y} - \Sigma^{-1/2}X\boldsymbol{\beta})\\ =&(\boldsymbol{Y} - X\boldsymbol{\beta})\Sigma^{-1/2}\Sigma^{-1/2}(\boldsymbol{Y} - X\boldsymbol{\beta})\\ =&||\boldsymbol{Y} - X\boldsymbol{\beta}||^2_\Sigma\end{aligned}$$

and therefore coincides with the object function of the corresponding GLS model.

Then, according to Theorem I.4.9 the normal equation is given by

$$\begin{aligned}& B'\boldsymbol{z} = B'B\hat{\boldsymbol{\beta}}\\ \Leftrightarrow\quad & X'(\Sigma^{-1/2})'\Sigma^{-1/2}X\hat{\boldsymbol{\beta}} = X'(\Sigma^{-1/2})'\Sigma^{-1/2}\boldsymbol{y}\\ \Leftrightarrow\quad & X'\Sigma^{-1}X\hat{\boldsymbol{\beta}} = X'\Sigma^{-1}\boldsymbol{y}.\end{aligned}$$

Similarly, under the assumption of regularity, we instantly get the GLS estimator given in (b).

(c) By linearity of expectations

$$\mathrm{E}(\hat{\boldsymbol{\beta}}) = (X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}\mathrm{E}(\boldsymbol{Y}) = (X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}X\boldsymbol{\beta} = \boldsymbol{\beta}.$$

By Theorem I.4.25

$$\mathrm{Cov}(\hat{\boldsymbol{\beta}}) = \sigma^2(B'B)^{-1} = \sigma^2(X'\Sigma^{-1}X)^{-1}.$$

(e) By assumption $\boldsymbol{Y}$ has density from the distribution $N_n(X\boldsymbol{\beta}, \sigma^2\Sigma)$, that is

$$\begin{aligned}l(\boldsymbol{\beta}|\boldsymbol{y}) \propto& -\ln(\sigma^2) - \frac{1}{2}\ln(|\Sigma|) - \frac{1}{2\sigma^2}(\boldsymbol{y} - X\boldsymbol{\beta})'\Sigma^{-1}(\boldsymbol{y} - X\boldsymbol{\beta})\\ \propto& -\frac{1}{2\sigma^2}(\Sigma^{-1/2}\boldsymbol{y} - \Sigma^{-1/2}X\boldsymbol{\beta})'I_n(\Sigma^{-1/2}\boldsymbol{y} - \Sigma^{-1/2}X\boldsymbol{\beta})\\ =& -\frac{1}{2\sigma^2}(\boldsymbol{z} - B\boldsymbol{\beta})'(\boldsymbol{z} - B\boldsymbol{\beta}).\end{aligned}$$

Since this is proportional to the log-likelihood for the OLS model, the likelihood equations are equivalent to the least squares equations and the proposition then follows by Theorem I.4.31.

(f) Consider a matrix X_0 with $\mathcal{I}m(X_0) \subseteq \mathcal{I}m(X)$ and denote by

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

the reduced GLS model associated with X_0. Then, according to (d), we can find a normal OLS and a reduced normal OLS with associated matrices $B = \Sigma^{-1}X$ and $B_0 = \Sigma^{-1}X_0$, respectively, such that

$$\boldsymbol{Z} = B\boldsymbol{\beta} + \boldsymbol{\eta}$$

and

$$\boldsymbol{Z} = B_0\boldsymbol{\gamma} + \boldsymbol{\eta}$$

are the transformed full and reduced normal OLS model, respectively.

If

$$Q = B(B'B)^{-1}B' = \Sigma^{-1}X(X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}$$

and

$$Q_0 = B_0(B_0'B_0)^{-1}B_0' = \Sigma^{-1}X_0(X_0'\Sigma^{-1}X_0)^{-1}X_0'\Sigma^{-1}$$

are the associated orthogonal projectors, respectively, then the appropriate decision rule for

$$H_0\colon \mathrm{E}\boldsymbol{Y} = X_0\boldsymbol{\gamma} \text{ for some } \boldsymbol{\gamma}$$

is given by Theorem I.4.40.