Aufgabenblatt 1

Aufgabe 1

a) $\vec{y} = \vec{1}\mu + \vec{1}b + \vec{\varepsilon}$, $\quad b \sim N(0, \tau^2)$, $\vec{\varepsilon} \sim N(\vec{0}, \sigma^2 \mathbb{I})$

Gemäß Bem. 1.9 gilt $\vec{y} \sim N(\vec{1}\mu, \underbrace{\vec{1}\tau^2\vec{1}' + \sigma^2\mathbb{I}}_{})$

"uniform correlation" ("compound symmetry")

$$\begin{pmatrix} \tau^2 & \cdots & \tau^2 \\ \vdots & \ddots & \vdots \\ \tau^2 & \cdots & \tau^2 \end{pmatrix} + \begin{pmatrix} \sigma^2 & & 0 \\ & \ddots & \\ 0 & & \sigma^2 \end{pmatrix} = \begin{pmatrix} \tau^2+\sigma^2 & \tau^2 & \cdots & \tau^2 \\ \tau^2 & \ddots & & \vdots \\ \vdots & & \ddots & \tau^2 \\ \tau^2 & \cdots & \tau^2 & \tau^2+\sigma^2 \end{pmatrix}$$

Eigenwerte von $Cov(\vec{y})$ (und Eigenvektoren): $\vec{1}$ ist offensichtlich ein EV zum EW $\sigma^2 + n\tau^2$, da alle Zeilensummen gleich $\sigma^2 + n\tau^2$ sind.

Raten: $(0 \ldots 0\, 1\, {-1}\, 0 \ldots 0)$ ist EV zum EW σ^2, hiervon gibt es $(n-1)$ offensichtlich untereinander linear unabhängige, aber nicht paarweise orthogonale Vektoren, die gemeinsam den Raum $\{\vec{x} \in \mathbb{R}^n : \vec{1}'\vec{x} = 0\}$ der zu $\vec{1}$ orthogonalen Vektoren aufspannen. Eine Orthonormalbasis $\vec{u}_2, \ldots, \vec{u}_n$ bekommen wir mit dem Gram-Schmidt-Verfahren.

Quadratwurzel von $Cov(\vec{y}) = \mathbb{U}\mathbb{D}\mathbb{U}'$ mit $\mathbb{U} = (\vec{u}_1\ \vec{u}_2 \ldots \vec{u}_n)$, $u_1 = \frac{1}{\sqrt{n}}\vec{1}$

$\mathbb{D} = diag(\sigma^2 + n\tau^2, \sigma^2, \ldots, \sigma^2)$

ist $\mathbb{U}\mathbb{D}^{\frac{1}{2}}$ wegen $\underbrace{(\mathbb{U}\mathbb{D}^{\frac{1}{2}})}_{V}\underbrace{(\mathbb{D}^{\frac{1}{2}}\mathbb{U}')}_{V'} = Cov(\vec{y})$

b) $$Cov(\vec{y}) = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} \tau^2 & 0 \\ 0 & \tau^2 \end{pmatrix}\begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} + \sigma^2\mathbb{I} = \tau^2\begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix} + \sigma^2\mathbb{I} = \begin{pmatrix} \sigma^2+\tau^2 & \tau^2 & & 0 \\ \tau^2 & \sigma^2+\tau^2 & & \\ & & \sigma^2+\tau^2 & \tau^2 \\ 0 & & \tau^2 & \sigma^2+\tau^2 \end{pmatrix}$$

Block-Korrelationsstruktur

Eigenwerte von $Cov(\vec{y})$: Wir können den linken oberen und den rechten unteren Block getrennt betrachten, da unkorreliert:

$(1\,1\,0\,0)$ und $(0\,0\,1\,1)$ sind orthogonale EV zum EW $\sigma^2 + 2\tau^2$

$(1\,{-1}\,0\,0)$ und $(0\,0\,1\,{-1})$ " " " " " σ^2

Quadratwurzel von $Cov(\vec{y}) = \mathbb{U}\mathbb{D}\mathbb{U}'$ mit $\mathbb{U} = (\vec{u}_1, \vec{u}_2, \vec{u}_3, \vec{u}_4)$, $\vec{u}_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}$ etc.

$\mathbb{D} = diag(\sigma^2 + 2\tau^2, \sigma^2, \sigma^2 + 2\tau^2, \sigma^2)$

$$\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 \end{pmatrix}\begin{pmatrix} \sqrt{\sigma^2+2\tau^2} & & & 0 \\ & \sigma & & \\ & & \sqrt{\sigma^2+2\tau^2} & \\ 0 & & & \sigma \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} \sqrt{\sigma^2+2\tau^2} & \sigma & 0 & 0 \\ \sqrt{\sigma^2+2\tau^2} & -\sigma & 0 & 0 \\ & & \sqrt{\sigma^2+2\tau^2} & \sigma \\ 0 & & \sqrt{\sigma^2+2\tau^2} & -\sigma \end{pmatrix}$$

c) $$\begin{pmatrix} y_{11} \\ y_{12} \\ y_{21} \\ y_{22} \end{pmatrix} = \begin{pmatrix} \mu \\ \mu \\ \mu \\ \mu \end{pmatrix} + \underbrace{\begin{pmatrix} 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}}_{\mathbb{Z}} \begin{pmatrix} a_1 \\ a_2 \\ b_1 \\ b_2 \end{pmatrix} + \begin{pmatrix} \varepsilon_{11} \\ \varepsilon_{12} \\ \varepsilon_{21} \\ \varepsilon_{22} \end{pmatrix}$$

$$\mathrm{Cov}(\vec{y}) = \mathbb{Z}\, \mathrm{diag}(\sigma_a^2, \sigma_a^2, \sigma_b^2, \sigma_b^2)\, \mathbb{Z}' + \sigma_\varepsilon^2 \mathbb{I}$$

$$= \begin{pmatrix} \sigma_a^2 & 0 & \sigma_b^2 & 0 \\ \sigma_a^2 & 0 & 0 & \sigma_b^2 \\ 0 & \sigma_a^2 & \sigma_b^2 & 0 \\ 0 & \sigma_a^2 & 0 & \sigma_b^2 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix} + \sigma_\varepsilon^2 \mathbb{I} = \begin{pmatrix} \sigma_a^2+\sigma_b^2+\sigma_\varepsilon^2 & \sigma_a^2 & \sigma_b^2 & 0 \\ \sigma_a^2 & & 0 & \sigma_b^2 \\ \sigma_b^2 & 0 & & \sigma_a^2 \\ 0 & \sigma_b^2 & \sigma_a^2 & \sigma_a^2+\sigma_b^2+\sigma_\varepsilon^2 \end{pmatrix}$$

(oder mittels Betrachtung der paarweisen Kovarianzen)

Eigenwerte & Eigenvektoren von $\mathrm{Cov}(\vec{y})$:

Wiederum ist $\vec{1}$ ein EV, diesmal zum EW $2\sigma_a^2 + 2\sigma_b^2 + \sigma_\varepsilon^2$

$(1\ {-1}\ {-1}\ 1)$ ist EV zum EW σ_ε^2

$(1\ 1\ {-1}\ {-1})$ ist EV zum EW $2\sigma_a^2 + \sigma_\varepsilon^2$

$(1\ {-1}\ 1\ {-1})$ ist EV zum EW $2\sigma_b^2 + \sigma_\varepsilon^2$

Quadratwurzel von $\mathrm{Cov}(\vec{y}) = \mathbb{U}\mathbb{D}\mathbb{U}'$ mit $\mathbb{U} = (\vec{u}_1, \vec{u}_2, \vec{u}_3, \vec{u}_4)$, $\vec{u}_1 = \frac{1}{2}\vec{1}$ etc

$$\mathbb{D} = \mathrm{diag}(2\sigma_a^2 + 2\sigma_b^2 + \sigma_\varepsilon^2, \sigma_\varepsilon^2, 2\sigma_a^2 + \sigma_\varepsilon^2, 2\sigma_b^2 + \sigma_\varepsilon^2)$$

$$\mathbb{U}\mathbb{D}^{\frac{1}{2}} = \frac{1}{2} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & 1 & -1 & -1 \end{pmatrix} \begin{pmatrix} 2\sigma_a^2+2\sigma_b^2+\sigma_\varepsilon^2 & & & 0 \\ & \sigma_\varepsilon^2 & & \\ & & 2\sigma_a^2+\sigma_\varepsilon^2 & \\ 0 & & & 2\sigma_b^2+\sigma_\varepsilon^2 \end{pmatrix}$$

$$= \frac{1}{2} \begin{pmatrix} 2\sigma_a^2+2\sigma_b^2+\sigma_\varepsilon^2 & \sigma_\varepsilon^2 & 2\sigma_a^2+\sigma_\varepsilon^2 & 2\sigma_b^2+\sigma_\varepsilon^2 \\ " & -\sigma_\varepsilon^2 & " & -2\sigma_b^2-\sigma_\varepsilon^2 \\ " & -\sigma_\varepsilon^2 & -(") & 2\sigma_b^2+\sigma_\varepsilon^2 \\ " & \sigma_\varepsilon^2 & -(") & -(") \end{pmatrix}$$

Aufgabe 2

Die Likelihood lautet

$$L(\vec{\beta}, \sigma^2; \vec{y}) = \frac{1}{\sqrt{(2\pi\sigma^2)^n \det \mathbb{V}}} \exp\left(-\frac{1}{2\sigma^2}(\vec{y}-\mathbb{X}\vec{\beta})'\mathbb{V}^{-1}(\vec{y}-\mathbb{X}\vec{\beta})\right)$$

$\vec{\beta}$ geht nur über die mit $\mathbb{V}^{-1}$ gewichtete Summe $(\vec{y}-\mathbb{X}\vec{\beta})'\mathbb{V}^{-1}(\vec{y}-\mathbb{X}\vec{\beta})$ ein, die zu minimieren ist (gewichtetes KQ-Kriterium).

Jetzt geht's entweder mit üblichen Ableitungsregeln für quadratische Formen weiter, oder wir transformieren die Daten mittels einer Matrix $\mathbb{W}$, die $\mathbb{W}'\mathbb{W} = \mathbb{V}^{-1}$ erfüllt: $(\vec{y}-\mathbb{X}\vec{\beta})'\mathbb{V}^{-1}(\vec{y}-\mathbb{X}\vec{\beta}) = (\vec{y}-\mathbb{X}\vec{\beta})'\mathbb{W}'\mathbb{W}(\vec{y}-\mathbb{X}\vec{\beta})$
(vgl. Aufgabe 1)

$$= (\mathbb{W}\vec{y}-\mathbb{W}\mathbb{X}\vec{\beta})'(\mathbb{W}\vec{y}-\mathbb{W}\mathbb{X}\vec{\beta}),$$

welches dem KQ-Kriterium im Modell $\mathbb{W}\vec{y} = \mathbb{W}\mathbb{X}\vec{\beta} + \mathbb{W}\vec{e}$ für die mit $\mathbb{W}$ transformierten Daten $\mathbb{W}\vec{y}$ und Designmatrix $\mathbb{W}\mathbb{X}$ ist, mit $\mathbb{W}\vec{e} \sim N(\mathbb{W}\vec{0}, \mathbb{W}\sigma^2\mathbb{V}\mathbb{W}') = N(\vec{0}, \sigma^2\mathbb{I})$ wegen $\mathbb{W}'\mathbb{W}\mathbb{V}\mathbb{W}' = \mathbb{W}'$, d.h. $\mathbb{W}\mathbb{V}\mathbb{W}' = \mathbb{I}$ aufgrund der Invertierbarkeit der Matrizen.

Folglich wird das KQ-Kriterium im transformierten Modell (= gewichtetes KQ-Kriterium im ursprünglichen Modell) minimiert durch

$$\hat{\vec{\beta}} = [(\mathbb{W}\mathbb{X})'(\mathbb{W}\mathbb{X})]^{-1}(\mathbb{W}\mathbb{X})'(\mathbb{W}\vec{y}) = [\mathbb{X}'\underbrace{\mathbb{W}'\mathbb{W}}_{\mathbb{V}^{-1}}\mathbb{X}]^{-1}\mathbb{X}'\underbrace{\mathbb{W}'\mathbb{W}}_{\mathbb{V}^{-1}}\vec{y}$$

$$E(\hat{\vec{\beta}}) = E\left([\mathbb{X}'\mathbb{V}^{-1}\mathbb{X}]^{-1}\mathbb{X}'\mathbb{V}^{-1}\vec{y}\right) = [\mathbb{X}'\mathbb{V}^{-1}\mathbb{X}]^{-1}\mathbb{X}'\mathbb{V}^{-1}\underbrace{E(\vec{y})}_{\mathbb{X}\vec{\beta}} = \vec{\beta}$$

$$Cov(\hat{\vec{\beta}}) = \underbrace{[\mathbb{X}'\mathbb{V}^{-1}\mathbb{X}]^{-1}\mathbb{X}'\mathbb{V}^{-1}\underbrace{Cov(\vec{y})}_{\sigma^2\mathbb{V}}\mathbb{V}^{-1}}_{\sigma^2\mathbb{V}^{-1}}\mathbb{X}[\mathbb{X}'\mathbb{V}^{-1}\mathbb{X}]^{-1} = \sigma^2(\mathbb{X}'\mathbb{V}^{-1}\mathbb{X})^{-1}$$

$\hat{\vec{\beta}}$ ist als lineare Transformation einer (multivariat) normalverteilten ZV ($\vec{y}$) natürlich ebenfall (multivariat) normalverteilt.

Obige Likelihood ist gleich der Likelihood im transformierten Modell, so dass $\hat{\sigma}^2 = \frac{1}{n}\hat{\vec{e}}_W'\hat{\vec{e}}_W = \frac{1}{n}(\mathbb{W}\vec{y}-\mathbb{W}\mathbb{X}\hat{\vec{\beta}})'(\mathbb{W}\vec{y}-\mathbb{W}\mathbb{X}\hat{\vec{\beta}}) =$

$$= \frac{1}{n}\underbrace{(\vec{y}-\mathbb{X}\hat{\vec{\beta}})'}_{\hat{\vec{e}}}\underbrace{\mathbb{W}'\mathbb{W}}_{\mathbb{V}^{-1}}\underbrace{(\vec{y}-\mathbb{X}\hat{\vec{\beta}})}_{\hat{\vec{e}}}$$

Aufgabe 3

b) Bekannt: KQ-Schätzer hat im multiplen linearen Regressionsmodell mit unkorrelierten Fehlern eine im Sinne der Loewner-Halbordnung minimale Kovarianzmatrix unter allen linearen unverzerrten Schätzern von $\vec{\beta}$. Außerdem ist der gewichtete KQ-Schätzer im Modell mit korrelierten Fehlern gleich dem KQ-Schätzer im transformierten Modell mit unkorrelierten Fehlern, vgl. Aufgabe 2.

Annahme: Es gibt eine $(p \times n)$-Matrix $\mathbb{A}$, so dass $\mathbb{A}\vec{Y}$ unverzerrt für $\vec{\beta}$, aber $Cov(\mathbb{A}\vec{Y}) = \mathbb{A}\mathbb{V}\mathbb{A}'$ nicht im Sinne der Loewner-Halbordnung von $(\mathbb{X}'\mathbb{V}^{-1}\mathbb{X})^{-1}$, der Kovarianzmatrix des gewichteten KQ-Schätzers, dominiert wird.

Dies führt aber unmittelbar zum Widerspruch, da dann $\mathbb{A}\vec{Y} = (\mathbb{A}\mathbb{W}^{-1})(\mathbb{W}\vec{Y})$ auch erwartungstreu im transformierten Modell mit unkorrelierten Fehlern ist, ohne vom KQ-Schätzer in diesem Modell im Sinne minimaler Kovarianzmatrix dominiert zu sein

a) Gewichteter KQ-Schätzer: $(\mathbb{X}'\mathbb{V}^{-1}\mathbb{X})^{-1} = \frac{3}{2}$ $(\sigma^2 = 1)$

da $\mathbb{V}^{-1} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}^{-1} = \begin{pmatrix} \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} \end{pmatrix}$; $\mathbb{X}'\mathbb{V}^{-1}\mathbb{X} = (1\ 1)\begin{pmatrix} \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{2}{3}$

Gewöhnlicher KQ-Schätzer: $(\mathbb{X}'\mathbb{X})^{-1}\mathbb{X}'\mathbb{V}\mathbb{X}(\mathbb{X}'\mathbb{X})^{-1} = \frac{1}{2} \cdot 6 \cdot \frac{1}{2} = \frac{6}{4} = \underline{\frac{3}{2}}$

da $\mathbb{X}'\mathbb{V}\mathbb{X} = (1\ 1)\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = 6$; $\mathbb{X}'\mathbb{X} = 2$

In der Tat sind in diesem Modell der gewichtete KQ-Schätzer und der gewöhnliche KQ-Schätzer für μ sogar gleich:

$(\mathbb{X}'\mathbb{V}^{-1}\mathbb{X})^{-1}\mathbb{X}'\mathbb{V}^{-1} = \frac{3}{2}(1\ 1)\begin{pmatrix} \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} \end{pmatrix} = \frac{3}{2}\left(\frac{1}{3}\ \ \frac{1}{3}\right) = \left(\frac{1}{2}\ \ \frac{1}{2}\right)$,

also ist der gewichtete KQ-Schätzer für μ gleich dem gewöhnlichen arithmetischen Mittel und damit gleich dem gewöhnlichen KQ-Schätzer (aus Symmetriegründen ist es hier plausibel, keine Beobachtung höher zu gewichten als die anderen).

Dennoch lohnt die Anpassung eines gemischten linearen Modells, da wir sonst den Standardfehler der Schätzung falsch einschätzen: Konfidenzintervalle, Tests ...