Aufgabe 12 (Fortsetzung)

b) Gemäß Bem. 2.7 existiert der ML-Schätzer im LMM genau dann, wenn

$\nexists \vec{\beta} \in \mathbb{R}^p, \vec{b}_1, \dots, \vec{b}_m \in \mathbb{R}^q: \quad \vec{y}_i = X_i \vec{\beta} + Z_i \vec{b}_i \quad \forall i = 1, \dots, m$

Hier: " " $\quad \vec{y}_i = Z_i A_i \vec{\beta} + Z_i \vec{b}_i = Z_i (A_i \beta + \vec{b}_i) \; \forall i = 1, \dots, m$ ⊗

Da hierbei $\vec{b}_i \in \mathbb{R}^q$ bel. $\forall i = 1, \dots, m$, ist auch $A_i \beta + b_i \in \mathbb{R}^q$ bel., so dass die Existenzbedingung ⊗ gleichwertig damit ist, dass sich mindestens ein $\vec{y}_i$ nicht als $Z_i \vec{v}$ für ein $\vec{v} \in \mathbb{R}^q$ schreiben läßt – und dies ist der erste Teil der Behauptung.

Da es ein i_0 mit $n_{i_0} > q$ gibt nach Vor. und $\sigma^2 > 0$ gelten soll, liegt $\vec{y}_{i_0}$ mit Ws'keit 1 nicht im q-dimensionalen Unterraum des $\mathbb{R}^{n_{i_0}}$, der von den Spalten von Z_{i_0} aufgespannt wird, so dass die Existenzbedingungen mit Ws'keit 1 erfüllt ist, vgl. Bem 2.8a.

c) In der Form $\vec{y}_i = Z_i A_i \vec{\beta} + Z_i \vec{b}_i + \varepsilon_i$, $\vec{b}_i \sim N(0, \sigma^2 \mathbb{B})$, $\vec{\varepsilon}_i \sim N(0, \sigma^2 \Sigma)$, läßt sich das Modell schreiben mittels

$$Z_i = \begin{pmatrix} 1 & t_1 & t_1^2 \\ \vdots & \vdots & \vdots \\ 1 & t_n & t_n^2 \end{pmatrix}, \quad \vec{b}_i = \begin{pmatrix} a_i \\ c_i \\ d_i \end{pmatrix} \sim N(\vec{0}, \sigma^2 \mathbb{B}) \text{ mit } \mathbb{B} = \begin{pmatrix} \frac{\sigma_a^2}{\sigma^2} & 0 & 0 \\ 0 & \frac{\sigma_c^2}{\sigma^2} & 0 \\ 0 & 0 & \frac{\sigma_d^2}{\sigma^2} \end{pmatrix}$$

$$A_i = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}, \text{ so dass } X_i = Z_i A_i = \begin{pmatrix} 1 & t_1 \\ \vdots & \vdots \\ 1 & t_n \end{pmatrix}, \quad \vec{\beta} = \begin{pmatrix} \alpha \\ \gamma \end{pmatrix}, \; \Sigma_i = \mathbb{I}_n, \; i = 1, \dots, m$$

In diesem Modell wird angenommen, dass die mittlere Größe eines Kindes in der Population linear über die Zeit anwächst (über den Beobachtungszeitraum). α ist hierbei die mittlere Größe bei Geburt, falls auch vor t_1 Linearität angenommen werden kann, und γ das mittlere Wachstum während einer Zeiteinheit.

Für jedes einzelne Kind kann die individuelle „Wachstumskurve" $X_i \beta + Z_i b_i$ von der mittleren Kurve $X_i \beta$ in der Population (hier X_i, Z_i unabh. von i) abweichen. Ein Kind kann schneller wachsen als der Durchschnitt ($c_i > 0$)

bei der Geburt vergleichsweise groß oder klein sein ($a_i > 0$ bzw. $a_i < 0$), und es können sich systematische Abweichungen von einem linearen Verlauf in „parabel-förmiger Gestalt" ergeben ($d_i \neq 0$). $\varepsilon_{i1}, \dots, \varepsilon_{in}$ sind dagegen singuläre Effekte (Messfehler etc.), die jeweils nur die Messung zu einem Zeitpunkt beeinflussen.

d) – Das Modell unterscheidet nicht zwischen Jungen und Mädchen. Unterschiede in der Größe bei Geburt könnten durch eine Dummy-Variable $g_i = \begin{cases} 1 & \text{Kind ist Junge} \\ 0 & \text{Mädchen} \end{cases}$ und einen weiteren festen Parameter δ als zusätzlicher Summand $g_i \delta$ im Modell eingebracht werden. Dann würde man annehmen, dass Jungs bei Geburt im Mittel δ Größeneinheiten größer als Mädchen sind, und dieser Unterschied über die Zeit konstant bestehen bleibt. Durch Hinzufügen dieses Terms würde A_i zu $\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ falls Kind i Junge, und $A_i = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$, falls Kind i Mädchen, und $\vec{\beta} = \begin{pmatrix} \alpha \\ \gamma \\ \delta \end{pmatrix}$.
Sollten Jungs im Mittel mit einer anderen Steigung wachsen als Mädchen, könnte dies durch einen Zusatzterm $g_i t_j \beta$ modelliert werden, mit $A_i = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$ falls Kind i Junge, $A_i = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ bei Mädchen, und $\vec{\beta} = \begin{pmatrix} \alpha \\ \gamma \\ \beta \end{pmatrix}$. Natürlich lassen sich auch beide Erweiterungen gleichzeitig berücksichtigen, mittels $A_i = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ für Junge und $\vec{\beta} = \begin{pmatrix} \alpha \\ \gamma \\ \delta \\ \beta \end{pmatrix}$

– Ein weiterer fehlender Umstand ist mögliche Korrelation zwischen benachbarten Zeitpunkten: Die für jedes Kind individuell unterstellte Parabel ist recht unflexibel. (Auto-)Korrelation könnte mittels eines so genannten AR(1)-Modells für $\varepsilon_{i1}, \dots, \varepsilon_{in}$ modelliert werden, um flexiblere Verläufe modellieren zu können.

Dies bedeutet: $\varepsilon_{ij} = \phi \varepsilon_{ij-1} + u_{ij}$, $j = 2, \dots, n$, $i \in \{1, \dots, m\}$

mit u_{ij} u.i.v. $\forall i, j$ mit $u_{ij} \sim N(0, \sigma_u^2)$.

Um Stationarität der Verteilung von ε_{ij} über die Zeit zu erreichen (→ Zeitreihen), kann man $\varepsilon_{i1} \sim N\left(0, \frac{\sigma_u^2}{1-\phi^2}\right)$ initialisieren $\forall i = 1, \dots, m$, so dass $\varepsilon_{ij} \sim N\left(0, \frac{\sigma_u^2}{1-\phi^2}\right) \forall i, j$

wegen $Var(\varepsilon_{ij}) = Var(\phi \varepsilon_{ij-1} + u_{ij}) = \phi^2 Var(\varepsilon_{ij-1}) + \sigma_u^2$

$$\text{(induktiv)} = \phi^2 \frac{\sigma_u^2}{1-\phi^2} + \sigma_u^2 = \frac{\sigma_u^2}{1-\phi^2} \quad \forall i, j$$

Hier: $\Sigma_i = \begin{pmatrix} 1 & \phi & \cdots & \phi \\ \phi & \ddots & & \vdots \\ \vdots & & \ddots & \phi \\ \phi & \cdots & \phi & 1 \end{pmatrix}$

e) Damit Z_i vollen Rang $q = 3$ hat, müssen $n \geq 3$ verschiedene Messzeitpunkte vorliegen, wobei sogar $n \geq 4$ gelten muss, damit mindestens ein $n_i > q$. Dies ist bereits ausreichend, da $\sum_{i=1}^{m} A_i' A_i = \sum_{i=1}^{m} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} = m \mathbb{I}_2$ stets nicht-singulär.

In der 1. Erweiterung aus c wäre

$$A_i' A_i = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \text{ für Mädchen}$$

$$A_i' A_i = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \text{ für Junge,}$$

so dass zusätzlich mindestens ein Junge und mind. ein Mädchen in der Stichprobe sein müßte, damit $\sum_{i=1}^{m} A_i' A_i = \begin{pmatrix} m & 0 & \tilde{m} \\ 0 & m & 0 \\ \tilde{m} & 0 & \tilde{m} \end{pmatrix}$ (mit $\tilde{m}$ = Anzahl der Jungs in der Stichprobe) nicht-singulär ist.

f) Siehe Zusatzblatt

Aufgabe 13:

In Analogie zu Aufgabe 11 können wir $\hat{\beta}_m^{KQ} - \vec{\beta}$ schreiben als

$$\hat{\beta}_m^{KQ} - \vec{\beta} = \left(\frac{1}{m}\sum_{i=1}^m \vec{X}_i \vec{X}_i'\right)^{-1} \frac{1}{m}\sum_{i=1}^m \vec{X}_i' \varepsilon_i \quad (*)$$

diesmal aber mit $(\vec{X}_1, \vec{\varepsilon}_1), (\vec{X}_2, \vec{\varepsilon}_2), \ldots$ u.i.v.

a) Somit auch $\vec{X}_1\vec{X}_1', \vec{X}_2\vec{X}_2', \ldots$ u.i.v., so dass mit dem starken Gesetz der großen Zahlen das arithmetische Mittel von $\frac{1}{m}\sum_{i=1}^m \vec{X}_1\vec{X}_1' \xrightarrow{f.s.} E(\vec{X}_1\vec{X}_1') = \mathbb{W}$. Da $\mathbb{W}$ invertierbar, muss für "fast alle" $\omega \in \Omega$ auch $\frac{1}{m}\sum_{i=1}^m \vec{X}_1(\omega)\vec{X}_1'(\omega)$ invertierbar sein, da nahe bei $\mathbb{W}$, und damit $g\left(\frac{1}{m}\sum_{i=1}^m \vec{X}_i\vec{X}_i'\right) = \left(\frac{1}{m}\sum_{i=1}^m \vec{X}_i\vec{X}_i'\right)^{-1} \xrightarrow{f.s.} \mathbb{W}^{-1}$, da g als Inversion stetig ist.

Gleichzeitig aber $\frac{1}{m}\sum_{i=1}^m \vec{X}_i'\varepsilon_i \xrightarrow{f.s.} E(\vec{X}_1'\varepsilon_1) = E(\vec{X}_1)'\underbrace{E(\varepsilon_1)}_{\vec{0}} = \vec{0}$,

da $\vec{X}_1'\varepsilon_1, \vec{X}_2'\varepsilon_2, \ldots$ u.i.v. und $\vec{X}_1, \vec{\varepsilon}_1$ unabhängig.

Insgesamt: $(*) = \left(\frac{1}{m}\sum_{i=1}^m \vec{X}_1\vec{X}_1'\right)^{-1} \frac{1}{m}\sum_{i=1}^m \vec{X}_i'\varepsilon_i \xrightarrow{f.s.} \mathbb{W}^{-1}\cdot\vec{0} = \vec{0}$.

b) Wegen $E(\vec{X}_i'\varepsilon_i) = 0$, $\mathrm{Var}(\vec{X}_i'\varepsilon_i) = E(\vec{X}_i'\varepsilon_i^2\vec{X}_i) = E(\varepsilon_2^2\vec{X}_i'\vec{X}_i) = \sigma^2\mathbb{W}$ gleich aufgrund der Unabhängigkeit von ε_i^2 und $\vec{X}_i'\vec{X}_i$, und $\vec{X}_1'\varepsilon_1, \vec{X}_2'\varepsilon_2, \ldots$ u.i.v. gilt mit ZGWS: $\frac{1}{\sqrt{m}}\sum_{i=1}^m \vec{X}_i'\varepsilon_i \xrightarrow{V} N(\vec{0}, \sigma^2\mathbb{W})$.

Mit dem Satz von Slutsky:

$$\sqrt{m}\left(\hat{\beta}_m^{KQ} - \vec{\beta}\right) = \left(\frac{1}{m}\sum_{i=1}^m \vec{X}_i\vec{X}_i'\right)^{-1} \frac{1}{\sqrt{m}}\sum_{i=1}^m \vec{X}_i'\varepsilon_i \xrightarrow{V} N(\vec{0}, \sigma^2\underbrace{\mathbb{W}^{-1}\mathbb{W}\mathbb{W}^{-1}}_{\mathbb{W}^{-1}})$$