# Gemischte lineare Modelle
## Aufgabenblatt 5

06. Juli 2009

11. Zeigen Sie Korollar 3.9:
Im gewöhnlichen linearen Regressionsmodell $Y_i = \vec{x}_i'\vec{\beta} + \epsilon_i$, $i = 1, 2, \ldots$, mit $\epsilon_1, \epsilon_2, \ldots$ u.i.v. und $\epsilon_1 \sim (0, \sigma^2)$ ist der KQ-Schätzer

$$\hat{\vec{\beta}}_m^{KQ} = \left( \sum_{i=1}^{m} \vec{x}_i \vec{x}_i' \right)^{-1} \sum_{i=1}^{m} \vec{x}_i y_i$$

asymptotisch normalverteilt,

$$\sqrt{m} \left( \hat{\vec{\beta}}_m^{KQ} - \vec{\beta} \right) \xrightarrow{V} N \left( \vec{0}, \sigma^2 \boldsymbol{W}^{-1} \right),$$

wenn es ein $a > 0$ gibt mit $||x_i|| < a \forall i \in \mathbb{N}$ und $\lim_{m \to \infty} \frac{1}{m} \sum_{i=1}^{m} \vec{x}_i \vec{x}_i' =: \boldsymbol{W} > 0$ gilt.

12. **Lineares Wachstumskurvenmodell:** Ein LMM $\vec{Y}_i = \boldsymbol{X}_i \vec{\beta} + \boldsymbol{Z}_i \vec{b}_i + \vec{\epsilon}_i$ mit $\boldsymbol{X}_i = \boldsymbol{Z}_i \boldsymbol{A}_i$ für eine $(q \times p)$-Matrix $\boldsymbol{A}_i$, $\vec{b}_i \sim N(\vec{0}, \sigma^2 \boldsymbol{B})$, $\vec{\epsilon}_i \sim N(\vec{0}, \sigma^2 \boldsymbol{\Sigma}_i)$, $i = 1, \ldots, m$, bezeichnet man als lineares Wachstumskurvenmodell. Wir fordern, dass $\sum_{i=1}^{m} \boldsymbol{A}_i' \boldsymbol{A}_i$ nicht-singulär ist und dass alle Matrizen $\boldsymbol{Z}_i$ vollen Rang $q \leq n_i$ haben, mit $q < n_i$ für mindestens ein $i$.

   a) Zeigen Sie, dass das Modell unter diesen Bedingungen identifizierbar ist.
   TIPP: Nutzen Sie, dass $\boldsymbol{X}$ Rang $p$ hat $\iff$ $\boldsymbol{X}'\boldsymbol{X}$ hat Rang $p$.

   b)* Zeigen Sie, dass der ML-Schätzer in diesem Modell genau dann existiert, wenn mindestens ein $\vec{Y}_i$ nicht im Spaltenraum der zugehörigen Matrix $\boldsymbol{Z}_i$ liegt. Begründen Sie weiter, dass dies im Falle $\sigma^2 > 0$ und $\boldsymbol{\Sigma}_i$ nicht-singulär für alle $i = 1, \ldots, m$ mit Wahrscheinlichkeit 1 passiert.

   c)* Um das Wachstum von $m$ Kindern zu beschreiben, wollen wir folgendes Modell unterstellen:

$$Y_{ij} = \alpha + \gamma t_j + a_i + c_i t_j + d_i t_j^2 + e_{ij}, \quad j = 1, \ldots, n, i = 1, \ldots, m,$$

   wobei $Y_{ij}$ die $j$-te Beobachtung für Individuum $i$ ist,
   $t_j$ der $j$-te Beobachtungszeitpunkt, $j = 1, \ldots, n$,
   $\alpha$ und $\gamma$ unbekannte relle Konstanten,
   $a_i, c_i, d_i, e_{i1}, \ldots, e_{in}$ Zufallsvariablen, die unabhängig untereinander und auch unabhängig für verschiedene $i$ sind,
   mit $a_1, \ldots, a_m$ identisch $N(0, \sigma_a^2)$,
   $c_1, \ldots, c_m$ identisch $N(0, \sigma_c^2)$,
   $d_1, \ldots, d_m$ identisch $N(0, \sigma_d^2)$, und
   $e_{11}, \ldots, e_{1n}, \ldots, e_{21}, \ldots, e_{mn}$ identisch $N(0, \sigma^2)$ verteilt.
   Schreiben Sie das Modell in der obigen Form, d.h., geben Sie die Matrizen $\boldsymbol{Z}_i$, $\boldsymbol{A}_i$, $\boldsymbol{B}$ und $\boldsymbol{\Sigma}_i$ sowie die Vektoren $\vec{\beta}$ und $\vec{b}_i$ an für $i = 1, \ldots, m$.
   Interpretieren Sie sodann die Größen $\alpha$ und $\gamma$ sowie die Zufallsvariablen $a_i, c_i, d_i, e_{i1}, \ldots, e_{in}$, $i = 1, \ldots, m$.

d)* Nennen Sie mindestens einen Defekt des Modells aus c), d.h. einen realistischen Sachverhalt, der nicht vom Modell beschrieben wird. Kann man das Modell erweitern, so dass dieser Defekt behoben wird, und wenn ja, wie?

e)* Welche Bedingungen müssen die Zeitpunkte $t_1, \ldots, t_n$ erfüllen, damit das Modell aus c) die Bedingungen erfüllt, dass $\sum_{i=1}^m \boldsymbol{A}_i' \boldsymbol{A}_i$ nicht-singulär ist und alle Matrizen $\boldsymbol{Z}_i$ vollen Rang $q \leq n_i$ haben mit $q < n_i$ für mindestens ein $i$?

f)* Generieren Sie $m = 10$ Wachstumskurven $y_{i,1}, \ldots, y_{i,n}$ der Länge $n = 11$ vom Modell aus c) mit $t_j = j - 1$, $j = 1, \ldots, 11$, $\alpha = 50, \gamma = 8$, $\sigma_a^2 = 1$, $\sigma_c^2 = 0.2^2$, $\sigma_d^2 = 0.1^2$ und $\sigma^2 = 1$. Variieren Sie sodann einige der Parameterwerte, generieren Sie Wachstumskurven mit den geänderten Werten, zeichnen Sie die neuen Kurven und kommentieren Sie.

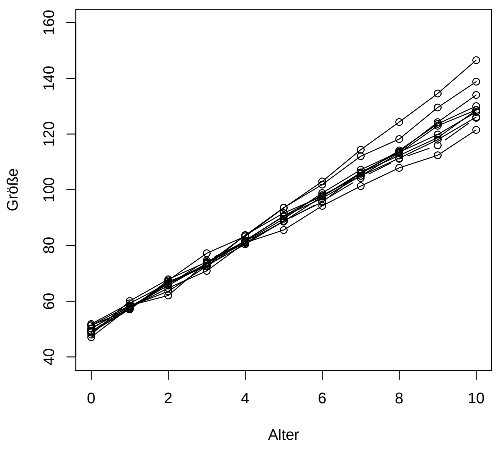
13.* Zeigen Sie Behauptung 3.10:
Seien $\vec{X}_1, \vec{X}_2, \ldots$ u.i.v. mit $\vec{X}_1 \sim F_{\vec{\xi}}$ und $E(\vec{X}_1 \vec{X}_1') =: \boldsymbol{W}$ invertierbar. Weiter seien $\epsilon_1, \epsilon_2, \ldots$ u.i.v. und unabhängig von $\vec{X}_1, \vec{X}_2, \ldots$ mit $\epsilon_1 \sim (0, \sigma^2)$. Es gelte $\vec{Y}_i = \vec{X}_i' \vec{\beta} + \epsilon_i$, $i = 1, 2, \ldots$, wobei kein Zusammenhang zwischen $\vec{\beta}$ und $\vec{\xi}$ bestehe. Dann gilt für den KQ-Schätzer $\hat{\vec{\beta}}_m^{KQ}$, dass

a) $\hat{\vec{\beta}}_m^{KQ} \xrightarrow{f.s.} \vec{\beta}$

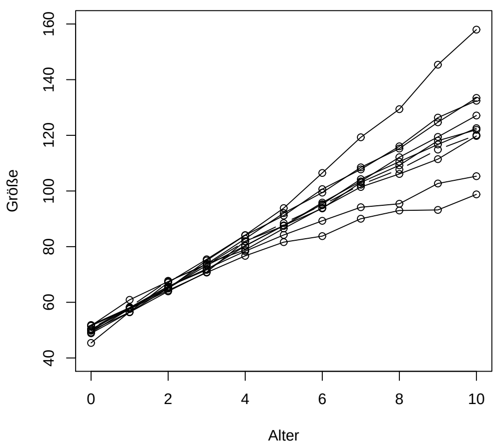
b) $\sqrt{m}(\hat{\vec{\beta}}_m^{KQ} - \vec{\beta}) \xrightarrow{V} N(\vec{0}, \sigma^2 \boldsymbol{W}^{-1})$

Die mit * gekennzeichneten Aufgabenteile können bis zum 16.07.2009, 16:00, in den Briefkasten Nr. 142 eingeworfen werden. Die Rückgabe erfolgt in der nächsten Übung am 20.07.2009.

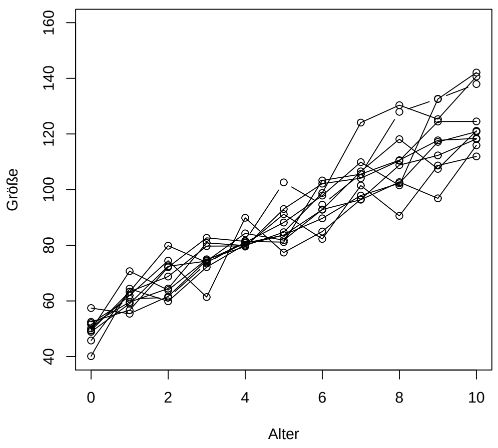
**Am kommenden Montag, den 13.07., findet als Ersatz für die ausgefallene Vorlesung vom 16.06. eine Rechner-Übung im Rechnerraum M / U 18 im Untergeschoss des Mathe-Gebäudes statt.**




Wachstumskurven mit den angegebenen Parameterwerten.




Kurven für $\sigma_c^2 = 0.1^2$ und $\sigma_d^2 = 0.2^2$: Die Kurven sind mehr gekrümmt als zuvor mit größeren Unterschieden zwischen den Kurven (da $d_i$ über das Quadrat von $t_j$ eingeht).




Kurven für $\sigma^2 = 5^2$: Die Kurven sind unruhiger als zuvor und nicht mehr monoton.

Abbildung 1: Wachstumskurven für verschiedene Parameterwerte.