Gemischte lineare Modelle

Aufgabenblatt 3

8. Juni 2009

8. **Balanciertes Modell mit zufälligen Koeffizienten** *(balanced random-coefficients)*
Ein LMM mit $n_1 = \ldots = n_m =: n$ und $\boldsymbol{Z}_i = \boldsymbol{Z}_1 \; \forall i = 1, \ldots, m$ nennt man *balanciert*. Falls $\boldsymbol{Z}_i = \boldsymbol{X}_i \; \forall i = 1, \ldots, m$, so spricht man von einem *Modell mit zufälligen Koeffizienten*.
In dieser Aufgabe betrachten wir ein balanciertes Modell mit zufälligen Koeffizienten, welches man also schreiben kann als

$$\vec{Y}_i = \boldsymbol{X}_1(\vec{\beta} + \vec{b}_i) + \vec{\epsilon}_i, \quad i = 1, \ldots, m,$$

mit $\vec{\beta}+\vec{b}_1, \ldots, \vec{\beta}+\vec{b}_m, \vec{\epsilon}_1, \ldots, \vec{\epsilon}_m$ unabhängig und $\vec{\beta}+\vec{b}_i \sim N(\vec{\beta}, \sigma^2\boldsymbol{B})$, $\vec{\epsilon}_i \sim N(\vec{0}, \sigma^2\boldsymbol{\Sigma})$. Hierbei wollen wir (wie oft üblich) $\boldsymbol{\Sigma} = \boldsymbol{I}$ unterstellen.

a) Geben Sie ein Szenario an, für das ein balanciertes Modell mit zufälligen Koeffizienten plausibel erscheint.

b) Sei $\overline{\vec{y}} = \sum_{i=1}^{m} \vec{y}_i/m$ das arithmetische Mittel der Beobachtungsvektoren für die verschiedenen Individuen. Zeigen Sie, dass der gewöhnliche KQ-Schätzer für $\vec{\beta}$ gegeben ist durch $\hat{\vec{\beta}}_{KQ} = (\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1\overline{\vec{y}}$.

c) Zeigen Sie, dass der 'Residuenvektor' $\hat{\vec{e}} = \overline{\vec{y}} - \boldsymbol{X}_1\hat{\vec{\beta}}_{KQ}$ orthogonal zu den Regressoren ist, d.h. $\boldsymbol{X}_1'\hat{\vec{e}} = \vec{0}$.

d) Zeigen Sie, dass der optimale gewichtete KQ-Schätzer $\hat{\vec{\beta}}_{gKQ}$ aus Beh. 1.15 in einem solchen Modell gleich dem gewöhnlichen KQ-Schätzer ist.

   TIPP: Zeigen Sie mittels quadratischer Ergänzung, dass $\hat{\vec{\beta}}_{gKQ}$ den Ausdruck $(\overline{\vec{y}} - \boldsymbol{X}_1\vec{\beta})'\boldsymbol{V}^{-1}(\overline{\vec{y}} - \boldsymbol{X}_1\vec{\beta})$ minimieren muss, und zerlegen Sie diesen Ausdruck unter Verwendung von $\hat{\vec{e}}$ weiter.

e) Folgern Sie hieraus $\hat{\vec{\beta}}_{ML} = \hat{\vec{\beta}}_{ReML} = \hat{\vec{\beta}}_{gKQ} = \hat{\vec{\beta}}_{KQ}$, d.h., dass der ML-, der ReML-, der (für $\boldsymbol{V}$ unbekannt nicht berechenbare) optimale gewichtete KQ und der gewöhnliche KQ-Schätzer von $\vec{\beta}$ alle gleich sind.

f) Zeigen Sie, dass der ML-Schätzer in diesem Modell genau dann existiert, wenn es ein $i \in \{1, \ldots, m\}$ gibt, so dass $\vec{y}_i$ nicht im Spaltenraum von $\boldsymbol{X}_1$ liegt.

9.* **Lineares Modell mit zufälligem Achsenabschnitt** *(random intercept model)*
Falls $q = 1$, $\boldsymbol{Z}_i = \vec{1}_{n_i}$, sowie die letzte Spalte von $\boldsymbol{X}_i$ eine Spalte von Einsen ist $i = 1, \ldots, m$, so spricht man von einem *Modell mit zufälligem Achsenabschnitt*. Ein solches Modell kann man also schreiben als

$$\vec{Y}_i = \boldsymbol{X}_i \vec{\beta} + \vec{1}_{n_i} b_i + \vec{\epsilon}_i, \quad i = 1, \ldots, m,$$

mit $b_1, \ldots, b_m, \vec{\epsilon}_1, \ldots, \vec{\epsilon}_m$ unabhängig und $b_i \sim N(0, \tau^2)$, $\vec{\epsilon}_i \sim N(\vec{0}, \sigma^2 \boldsymbol{\Sigma})$, wobei wir im Folgenden (wie oft üblich) $\boldsymbol{\Sigma} = \boldsymbol{I}$ unterstellen.

a) Geben Sie ein Szenario an, für das ein Modell mit zufälligen Koeffizienten plausibel erscheint.

b) Zeigen Sie, dass der optimale gewichtete KQ-Schätzer $\hat{\vec{\beta}}_{gKQ}$ aus 1.15 geschrieben werden kann als

$$\hat{\vec{\beta}}_{gKQ} = \left[\sum_{i=1}^{m} \left(\boldsymbol{X}_i' \boldsymbol{X}_i - \frac{n_i^2 \kappa}{1 + n_i \kappa} \overline{\boldsymbol{x}}_i \overline{\boldsymbol{x}}_i' \right) \right]^{-1} \sum_{i=1}^{m} \left(\boldsymbol{X}_i' \vec{y}_i - \frac{n_i^2 \kappa}{1 + n_i \kappa} \overline{\boldsymbol{x}}_i \overline{y}_i \right)$$

mit $\kappa = \tau^2/\sigma^2$, $\overline{\boldsymbol{x}}_i = \boldsymbol{X}_i' \vec{1}_{n_i}/n_i = \sum_{j=1}^{n_i} \boldsymbol{x}_{ij}/n_i$ und $\overline{y}_i = \vec{1}_{n_i}' \vec{y}_i/n_i = \sum_{j=1}^{n_i} y_{ij}/n_i$.

c) Zeigen Sie, dass der fixed effects estimator $\hat{\vec{\beta}}_\infty$ gleich dem gewöhnlichen KQ-Schätzer aus den zentrierten Daten $(\vec{y}_i - \overline{y}_i \vec{1}_{n_i}, \boldsymbol{X}_i - \vec{1}_{n_i} \overline{\boldsymbol{x}}_i')$, $i = 1, \ldots, m$, ist.

d) Begründen Sie, dass $\boldsymbol{X}_i (\boldsymbol{X}_i' \boldsymbol{X}_i)^{-1} \boldsymbol{X}_i' \vec{1}_{n_i} = \vec{1}_{n_i}$ gilt, $i = 1, \ldots, m$.
Hieraus folgt weiter, dass $\overline{\boldsymbol{x}}_i{}' (\boldsymbol{X}_i' \boldsymbol{X}_i)^{-1} \overline{\boldsymbol{x}}_i = n_i^{-1}$.

e) Zeigen Sie unter Verwendung von d), dass der optimale gewichtete KQ-Schätzer für $\vec{\beta}$ aus b) im *balancierten linearen Modell mit zufälligem Achsenabschnitt* gleich dem gewöhnlichen KQ-Schätzer für $\vec{\beta}$ ist. Hierbei nennen wir ein lineares Modell mit zufälligem Achsenabschnitt *balanciert*, wenn $n_1 = \ldots = n_m =: n$ und $\boldsymbol{X}_1 = \ldots = \boldsymbol{X}_m$ gilt.
TIPP: Nutzen Sie zur Matrizeninvertierung die Formel

$$(\boldsymbol{C} - d\vec{c}\vec{c}')^{-1} = \boldsymbol{C}^{-1} + d \frac{\boldsymbol{C}^{-1} \vec{c}\vec{c}' \boldsymbol{C}^{-1}}{1 - d\vec{c}' \boldsymbol{C}^{-1} \vec{c}} .$$

Für die Notation können Sie sich auch an A8b) orientieren.

f) Begründen Sie, dass der ML-Schätzer für $\vec{\theta}$ im balancierten Modell mit zufälligen Achsenabschnitten genau dann nicht existiert, wenn $\vec{y}_1, \ldots, \vec{y}_m$ alle im Spaltenraum von $\boldsymbol{X}_1$ liegen und nach Zentrieren gleich sind.

Die mit * gekennzeichnete Aufgabe kann bis zum Donnerstag, 18.06.2009, 15:00, in den Briefkasten Nr. 142 eingeworfen werden. Die Rückgabe erfolgt in der nächsten Übung am 22.06.2009.
HINWEIS: Am 16.06.2009 fällt die Vorlesung aus!