Züchtungslehre - Varianzkomponentenschätzung - Teil 2

Peter von Rohr

2016-12-02

Varianzanalyse

- Erwartungswerte von Summenquadraten sind Funktionen von Varianzkomponenten
- Negative Schätzwerte aufgrund von Datenkonstellationen
- keine Konsistenz, d.h. bei steigenden Datenmengen keine Steigerung der Qualität
- keine Berücksichtigung der Verteilung der Daten

Likelihood

- Anpassung einer Verteilung an die Daten
- Verteilungsparameter sollen so geschätzt werden, dass Verteilung "optimal" zu Daten passt
- Als Kriterium dient die **Likelihood** L
- Definition

$$L(\theta) = f(y|\theta)$$

Maximierung der Likelihood

- Nach Beobachtung der Daten ist y fix gegeben
- Deshalb wird L als Funktion der Parameter θ aufgegasst
- **Ziel**: finde Parameter θ, so dass L möglichst gut zu beobachteten Daten passt
- Umsetzung: Maximierung von L im Bezug auf θ, setzte θ beim Maximum von L als Schätzwert

$$\hat{\theta}_{ML} = argmax_{\theta} \; L(\theta)$$

Beispiel

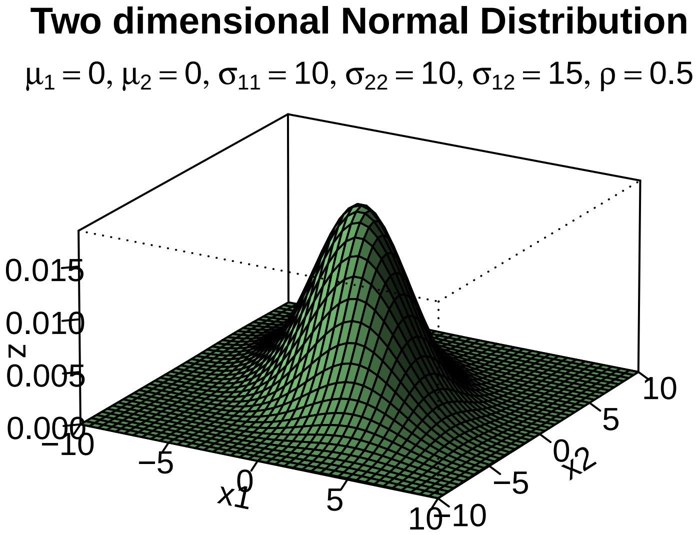
- Regressionmodell

$$y = Xb + e$$

- Ziel: Schätzung der Restvarianz σ^2 mit ML
- Annahme: Beobachtungen sind normalverteilt

$$f_Y(y|b,\sigma^2) = (2\pi\sigma^2)^{-n/2} exp\left(-\frac{1}{2\sigma^2}(y - Xb)^T \ (y - Xb)\right)$$

Multivariate Normalverteilung



Parameter

- Für unser Beispiel sind b und σ^2 unbekannte Parameter
- Schätzung für b: Maximierung von L im Bezug auf b
- Schätzung für σ^2: Maximierung von L im Bezug auf σ^2
- Maximierung von L:
  - Partielle Ableitung
  - Ableitung 0 setzen
  - nach gesuchtem Parameter auflösen
- Vorbereitung: Transformation von L zu l, wobei

$$l(\theta) = \log(L(\theta))$$

ML Schätzung für b

- Transformation

$$l(b,\sigma^2) = \log(L(b,\sigma^2))$$
$$= -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}(y - Xb)^T\ (y - Xb)$$

- Partielle Ableitung:

$$\begin{aligned} \frac{\partial l(b,\sigma^2)}{\partial b} &= -\frac{1}{2\sigma^2}(-(y^TX)^T - X^Ty + 2X^TXb) \\ &= -\frac{1}{2\sigma^2}(-2X^Ty + 2X^TXb) \qquad (1) \end{aligned}$$

Bestimmung von $\hat{b}_{ML}$

- Nullstelle der Ableitung

$$-\frac{1}{2\sigma^2}(-2X^T y + 2X^T X b) = 0$$

- Normalgleichung

$$X^T y = X^T X \hat{b}$$

- Resultat ist gleich wie bei LS

$$\hat{b} = (X^T X)^{-1} X^T y$$

ML Schätzung für σ^2

- Analoges Vorgehen wie bei b
- Partielle Ableitung von l nach σ^2

$$\frac{\partial l(b,\sigma^2)}{\partial\sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}(y - Xb)^T\,(y - Xb) \tag{2}$$

- Vorausgesetzt, $\sigma^2 \neq 0$

$$\frac{1}{\hat{\sigma}^2}(y - Xb)^T\,(y - Xb) - n = 0$$

- Nullstelle der Ableitung

$$\hat{\sigma}^2 = \frac{1}{n}(y - Xb)^T\,(y - Xb) \tag{3}$$

Bestimmung von $\hat{\sigma}^2$

- Da b unbekannt, wird $\hat{b}_{ML}$ eingesetzt
- In Summennotation:

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} (y_i - x_i^T \hat{b})^2 \tag{4}$$

- Schätzung aufgrund der Residuen

$$\hat{\sigma}^2_{Res} = \frac{1}{n-p} \sum_{i=1}^{n} r_i^2 \tag{5}$$

- Somit $E\left[\hat{\sigma}^2_{ML}\right] \neq \sigma^2$

Lineares gemischtes Modell

- Gleiches Vorgehen
- Modell

$$y = Xb + Zu + e \tag{6}$$

$$var(e) = R = I * \sigma_e^2$$

$$var(u) = G.$$

$$E[e] = 0 \text{ und } E[u] = 0$$

$$E[y] = Xb \text{ und } var(y) = V$$

$$y \sim \mathcal{N}(Xb, V)$$

ML-Schätzung

- Log-Likelihood

$$l(b, V) = \log(L(b, V)) = -\frac{n}{2}\log(2\pi) - \frac{1}{2}\log(det(V)) - \frac{1}{2}(y - Xb)^T V$$

- Partielle Ableitungen bilden

$$\frac{\partial l(b, V)}{\partial b}$$

$$\frac{\partial l(b, V)}{\partial \sigma^2}$$

- Nullstellen finden
- Schätzer bestimmen

Restricted (Residual) Maximum Likelihood (REML)

- ML-Schätzer für Varianzkomponenten nicht erwartungstreu
- Problem: gleichzeitiges Schätzen von fixen Effekten, was bei Schätzung von σ^2 nicht berücksichtigt
- Lösung:
  - vorgängige transformation von y zu $\tilde{y} = Ky$, so dass gilt $E\left[\tilde{y}\right] = 0$
  - dann gleiches Vorgehen mit $l(\theta) = f(\tilde{y}|\theta)$
- Resultat entspricht **REML**-Schätzung, wobei $\hat{\sigma}^2_{REML}$ erwartungstreu