

BLUP Zuchtwertschätzung

Peter von Rohr

2016-11-11

BLUP

- ▶ steht für **Best Linear Unbiased Prediction**
- ▶ **Linear**: fixe Effekte plus zufällige Effekte als Linearkombination der Beobachtungswerte geschätzt.
- ▶ **Unbiased**: Schätzungen für fixe Effekte plus zufällige Effekte unverzerrt (unbiased) oder erwartungstreu
- ▶ **Best**: unter allen linearen unverzerrten Schätzern weist der BLUP-Schätzer die kleinste Fehlervarianz auf

Modell

$$y = Xb + Zu + e \quad (1)$$

- mit
- y : Vektor der Beobachtungswerte
 - b : Vektor der fixen Effekte
 - u : Vektor der zufälligen Effekte
 - e : Vektor der zufälligen Resteffekte
 - X : Inzidenzmatrix für b
 - Z : Inzidenzmatrix für u

Erwartungswerte und Varianzen

Die Erwartungswerte und Varianzen der Zufallsvariablen im Modell (1) sind definiert als

$$E \begin{bmatrix} y \\ u \\ e \end{bmatrix} = \begin{bmatrix} Xb \\ 0 \\ 0 \end{bmatrix} \quad (2)$$

$$\text{var} \begin{bmatrix} y \\ u \\ e \end{bmatrix} = \begin{bmatrix} ZGZ^T + R & ZG & R \\ & GZ^T & 0 \\ & & R & R \end{bmatrix} \quad (3)$$

Schätzgleichungen

Aus den BLUP-Eigenschaften lassen sich die folgenden Schätzgleichungen ableiten.

$$\hat{b} = (X^T V^{-1} X)^{-1} X^T V^{-1} y \quad (4)$$

$$\hat{u} = GZ^T V^{-1} (y - X\hat{b}) \quad (5)$$

mit \hat{b} Lösungsvektor der fixen Effekte
 \hat{u} Lösungsvektor der zufälligen Effekte
 $()^{-}$ verallgemeinerte Inverse

Mischmodellgleichungen (Mixed Model Equation)

$$\begin{bmatrix} X^T R^{-1} X & X^T R^{-1} Z \\ Z^T R^{-1} X & Z^T R^{-1} Z + G^{-1} \end{bmatrix} \begin{bmatrix} \hat{b} \\ \hat{u} \end{bmatrix} = \begin{bmatrix} X^T R^{-1} y \\ Z^T R^{-1} y \end{bmatrix} \quad (6)$$

- ▶ Gleiche Lösungen, wie explizite Schätzgleichungen
- ▶ R (meist) diagonal, somit einfacher zu invertieren
- ▶ Varianzkomponenten müssen bekannt sein

Das Tiermodell

$$y = Xb + Za + e \quad (7)$$

- mit
- y phänotypischen Beobachtungen
 - b fixe Effekte
 - a zufällige Effekte - Zuchtwerte
 - e zufällige Resteffekte
 - X Inzidenzmatrix für b
 - Z Inzidenzmatrix für a

Varianzen

$$\text{Var}(a) = G = A \sigma_a^2 \text{ und } G^{-1} = A^{-1} \sigma_a^{-2} \quad (8)$$

$$\text{Var}(e) = R = I \sigma_e^2 \text{ und } R^{-1} = I \sigma_e^{-2} \quad (9)$$

wobei A additiv genetische Verwandtschaftsmatrix
 I Einheitsmatrix
 σ_a^2 additiv genetische Varianz
 σ_e^2 Restvarianz

Mischmodellgleichungen

- ▶ für ein Merkmal
- ▶ Erweiterung mit Restvarianz σ_e^2 führt zu vereinfachten Form

$$\begin{bmatrix} X^T X & X^T Z \\ Z^T X & Z^T Z + A^{-1} \alpha \end{bmatrix} \begin{bmatrix} \hat{b} \\ \hat{a} \end{bmatrix} = \begin{bmatrix} X^T y \\ Z^T y \end{bmatrix} \quad (10)$$

wobei α Verhältnis der Varianzen $\sigma_e^2/\sigma_a^2 = (1 - h^2)/h^2$
 h^2 Heritabilität

- ▶ kompakte Schreibweise

$$M * \hat{s} = r$$

wobei M Koeffizientenmatrix heisst
 \hat{s} Lösungsvektor
 r Vektor der rechten Handseite

Ein Beispiel für das Tiermodell

- ▶ Merkmal Gewichtszuwachs (WWG in kg) vor dem Absetzen bei Kälbern
- ▶ Ziel: Zuchtwerte für das Merkmal (WWG) schätzen
- ▶ Varianzkomponenten $\sigma_e^2 = 40$ und $\sigma_a^2 = 20$.
- ▶ Somit ist $\alpha = 40/20 = 2$

Kalb	Geschlecht	Vater	Mutter	WWG
4	M	1	NA	4.5
5	F	3	2	2.9
6	F	1	2	3.9
7	M	4	5	3.5
8	M	3	6	5.0

Modell für eine Beobachtung

$$y_{ij} = b_i + a_j + e_{ij}$$

- wobei
- y_{ij} Merkmal WWG für Kalb j mit Geschlecht i
 - b_i fixer Effekt für Geschlecht i
 - a_j Zuchtwert für Kalb j
 - e_{ij} Resteffekt für Kalb j mit Geschlecht i

Gleichungssystem

$$4.5 = b_M + a_4 + e_{M4}$$

$$2.9 = b_F + a_5 + e_{F5}$$

$$3.9 = b_F + a_6 + e_{F6}$$

$$3.5 = b_M + a_7 + e_{M7}$$

$$5.0 = b_M + a_8 + e_{M8}$$

- ▶ Matrix-Vektor-Schreibweise

$$y = Xb + Za + e \tag{11}$$

Vektoren im Modell

$$y = \begin{bmatrix} 4.50 \\ 2.90 \\ 3.90 \\ 3.50 \\ 5.00 \end{bmatrix}, \quad b = \begin{bmatrix} b_M \\ b_F \end{bmatrix}, \quad a = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \\ a_6 \\ a_7 \\ a_8 \end{bmatrix}, \quad e = \begin{bmatrix} e_{M4} \\ e_{F5} \\ e_{F6} \\ e_{M7} \\ e_{M8} \end{bmatrix}$$

Inzidenzmatrizen

$$X = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 1 & 0 \\ 1 & 0 \end{bmatrix}$$

$$Z = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Aufstellen der Mischmodellgleichungen

$$X^T X = \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}$$

$$X^T Z = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \end{bmatrix}$$

$$Z^T Z = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 \end{bmatrix}$$

Pedigree

```
##   sire  dam
## 1 <NA> <NA>
## 2 <NA> <NA>
## 3 <NA> <NA>
## 4   1  <NA>
## 5   3   2
## 6   1   2
## 7   4   5
## 8   3   6
```


Inverse Verwandtschaftsmatrix

$$A^{-1} = \begin{bmatrix} 1.833 & 0.500 & 0.000 & -0.667 & 0.000 & -1.000 & 0.000 & 0.000 \\ 0.500 & 2.000 & 0.500 & 0.000 & -1.000 & -1.000 & 0.000 & 0.000 \\ 0.000 & 0.500 & 2.000 & 0.000 & -1.000 & 0.500 & 0.000 & -1.000 \\ -0.667 & 0.000 & 0.000 & 1.833 & 0.500 & 0.000 & -1.000 & 0.000 \\ 0.000 & -1.000 & -1.000 & 0.500 & 2.500 & 0.000 & -1.000 & 0.000 \\ -1.000 & -1.000 & 0.500 & 0.000 & 0.000 & 2.500 & 0.000 & -1.000 \\ 0.000 & 0.000 & 0.000 & -1.000 & -1.000 & 0.000 & 2.000 & 0.000 \\ 0.000 & 0.000 & -1.000 & 0.000 & 0.000 & -1.000 & 0.000 & 2.000 \end{bmatrix}$$

Rechte Handseite

$$X^T y = \begin{bmatrix} 13.0 \\ 6.8 \end{bmatrix}$$

$$Z^T y = \begin{bmatrix} 0.0 \\ 0.0 \\ 0.0 \\ 4.5 \\ 2.9 \\ 3.9 \\ 3.5 \\ 5.0 \end{bmatrix}$$

Lösung

$$\hat{s} = M^{-1} * r$$

Die Zahlenwerte für den Lösungsvektor bekommen wir

$$\hat{s} = \begin{bmatrix} 4.348 \\ 3.392 \\ 0.158 \\ -0.018 \\ -0.054 \\ 0.002 \\ -0.263 \\ 0.280 \\ -0.332 \\ 0.287 \end{bmatrix}$$