Multiple Traits

Peter von Rohr

13. November 2020

So Far . . .

- Prediction of Breeding Values for **one trait**

$\rightarrow$ **univariate** analyses

- In Livestock Breeding, populations are improved with respect to several traits

$\rightarrow$ **multi-trait** or multiple trait

- Different selection strategies and different approaches of how data is analysed are possible

Multiple Trait Selection

In principle, there are three possible strategies

- ▶ Selection index theory provides a tool for optimal integration of different sources of information   **(3)**
- ▶ But still other strategies are applied
  - ▶ Tandem selection   **(1)**
  - ▶ Selection based on independent thresholds   **(2)**

Tandem Selection

- Improve one trait at the time until they all reach a certain threshold
- Problem: For traits which are not improved
  - only correlated selection responses
  - can be negative ==> while improving one trait, other traits would get worse
- Populations with long generation intervals, response per year is very small

Example for problematic situation: trait1: milk yield and trait2: fertility in cattle
1. increasing milk yield would decrease fertility
2. improve fertility would lower milk yield

==> infinite loop without significant progress

Independent Selection Thresholds

Traits to improve
- mastitis resistence (above 95%)
- fertility (NRR above 80%)
- milk yield (above +400kg PBV)

- ▶ Applied before selection index
- ▶ Define selection thresholds in each of the traits
- ▶ Select animals as parents which are above thresholds for all traits
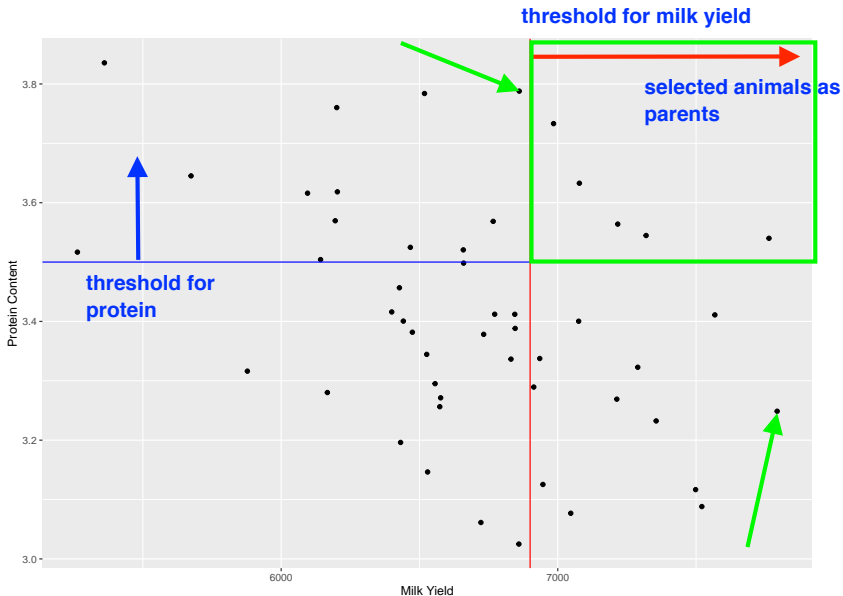
Example

Figure 1: Milk Yield and Protein Content For Dairy Cows

Pros and Cons

- ▶ Selection response in all traits
- ▶ Thresholds often set to only positive predicted breeding values in all traits

$\rightarrow$ exclusion of very many animals and reduction in genetic variability

- ▶ Genetic relationships between traits ignored

$\rightarrow$ genetic gain will not be as expected

3. Differences in the economic relevance ignored.

$\rightarrow$ threshold in all traits above positive predicted breeding values emphasizes traits with high heritability

Aggregate Genotype (Gesamtzuchtwert)

preferred selection strategy when populations are to be improved with respect to several traits at the same time.

▶ Define the set of important traits for which population should be improved
▶ Determine economic values w for these traits
▶ Aggregate genotype H follows as

Selection of animals as parents is best done when animals can be ranked according to a certain criterion. In our case this criterion is a single number. In univariate case with only one trait, this criterion is the predicted breeding value of each animal in the single trait.

$$H = w^T u$$

u is the vector of true breeding values

Selection Index

True Aggregate Genotype H cannot be observed, because the true breeding values u, cannot be observed. But we can predict H based on the PBV. This prediction of H is done using a method that is called selection index theory. and the predicted H is called index I

▶ Use index I to estimate H where I is a linear combination of information sources

$$I = b^T \hat{u}$$

▶ Index weights b are determined using selection index theory as

$$b = P^{-1} G w$$

▶ Information sources are predicted breeding values
▶ If traits in u and $\hat{u}$ are the same and $\hat{u}$ were estimated using BLUP, then $b = w$

Implementations

- First possible implementation
  - Do univariate predictions of breeding values using BLUP animal model
  - Combine $\hat{u}$ with appropriate b-values
- Imrprovement
  - get $\hat{u}$ from multivariate analysis

Multivariate Analysis

▶ Given two traits with univariate models

$$y_1 = X_1\beta_1 + Z_1u_1 + e_1$$
$$y_2 = X_2\beta_2 + Z_2u_2 + e_2$$

▶ Combine both univariate models by stacking one on top of the other, resulting in

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} Z_1 & 0 \\ 0 & Z_2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}$$

Multivariate Model

$$\left[\begin{array}{c} y_1 \\ y_2 \end{array} \right] = \left[\begin{array}{cc} X_1 & 0 \\ 0 & X_2 \end{array} \right] \left[\begin{array}{c} \beta_1 \\ \beta_2 \end{array} \right] + \left[\begin{array}{cc} Z_1 & 0 \\ 0 & Z_2 \end{array} \right] \left[\begin{array}{c} u_1 \\ u_2 \end{array} \right] + \left[\begin{array}{c} e_1 \\ e_2 \end{array} \right]$$

can be written as

$$y = X\beta + Zu + e$$

with $y = \left[\begin{array}{c} y_1 \\ y_2 \end{array} \right]$, $\beta = \left[\begin{array}{c} \beta_1 \\ \beta_2 \end{array} \right]$, $u = \left[\begin{array}{c} u_1 \\ u_2 \end{array} \right]$, $e = \left[\begin{array}{c} e_1 \\ e_2 \end{array} \right]$

$X = \left[\begin{array}{cc} X_1 & 0 \\ 0 & X_2 \end{array} \right]$, $Z = \left[\begin{array}{cc} Z_1 & 0 \\ 0 & Z_2 \end{array} \right]$

Multivariate Variance-Covariance Matrices

$$G_0 = \begin{bmatrix} \sigma^2_{g_1} & \sigma_{g_1,g_2} \\ \sigma_{g_1,g_2} & \sigma^2_{g_2} \end{bmatrix} = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix}$$

$$var(u) = var\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} g_{11}A & g_{12}A \\ g_{21}A & g_{22}A \end{bmatrix} = G_0 \otimes A = G$$

$$R_0 = \begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{bmatrix}$$

$$R = var(e) = var\begin{bmatrix} e_1 \\ e_2 \end{bmatrix} = \begin{bmatrix} r_{11}I_n & r_{12}I_n \\ r_{21}I_n & r_{22}I_n \end{bmatrix} = R_0 \otimes I_n$$

Solutions

▶ Mixed Model Equations

$$\begin{bmatrix} X^T R^{-1} X & X^T R^{-1} Z \\ Z^T R^{-1} X & Z^T R^{-1} Z + G^{-1} \end{bmatrix} \begin{bmatrix} \hat{\beta} \\ \hat{u} \end{bmatrix} = \begin{bmatrix} X^T R^{-1} y \\ Z^T R^{-1} y \end{bmatrix}$$

Advantages

- some traits have lower heritability than others
- environmental correlations exist between traits measured on the same animal
- some traits are available only a subset of all animals
- some traits were used for a first round of selection
- accuracies are higher in multivariate analyses