# Principles of mixed models and BLUP

The principle of estimation of breeding value is based on using

- Phenotypic observations as deviations to expected means (e.g. contemporary group mean)
- Weighing those deviations with a regression coefficient.

In selection index, the expected means are assumed known, and the only task of breeding value estimation is to find the appropriate weights for deviations (i.e. deviations of the observed records from those means).

In BLUP, these expected means (of a contemporary group) have to be estimated from the data.
A linear (mixed) model is used for this purpose.

> A mixed model is a linear model for fixed and random effects. Breeding values are random effects (they have variation). Random effects are estimated differently from fixed effects. Fixed effects are basically estimated as observed means (possibly corrected from some other fixed effects), whereas random effects are regressed toward a certain mean. If a herd mean is +1 above the breed average, we believe it is all herd effect and the herd is expected to be really one unit better than others. If an animal is +1 above the herd mean, we believe only part of this is due to the animals' genes, the other part is due to the environmental (or error) effect. The distinction between fixed and random effects is hard and would need a more theoretical statistical coverage

This practical will show by a simple example that both methods use the same weights for amalgamating the different pieces of information to estimate and EBV.

Given is data on 6 animals with the following information

| animal | sire | herd | observed performance |
|---|---|---|---|
| 1 | unknown | 1 | 35 |
| 2 | unknown | 2 | 50 |
| 3 | 2 | 1 | 42 |
| 4 | 2 | 1 | 38 |
| 5 | 1 | 2 | 47 |
| 6 | 1 | 2 | 43 |

Assume a heritability of 0.25

1. Calculate the EBV for animal 1, using selection index (you can use the spreadsheet STEBVAccuracy/STSELIND for this purpose to derive the weights, or, if you feel challenged, derive them yourself)

2. Calculate the EBV of all animals using a random model. *Ignore herd effects in first instance*

Work this out as follows (you can work in excel or R)

- Let the observed data be in **y** (a vector of 6 x 1)
- Express them as deviations from the mean:
  this is a new vector called $\mathbf{y^d}$.

In a random model we only estimate random effects, but we estimate them all jointly.
The random model looks like (see also the notes)

$$\mathbf{y^d} \quad = \quad Zu + e$$

- Now work out what the matrix Z looks like (it links the data up to the animals)
- Work out the additive genetic relationships between all animals, call this matrix **A**

- The breeding values can be estimated from $$[Z'Z + \lambda A^{-1}][\hat{u}] = Z'y^d$$

so that $$\hat{u} = [Z'Z + \lambda A^{-1}]^{-}Z'y^d$$

work in steps:

- Determine the relationships matrix A, and invert it → $A^{-1}$
- Determine the value for $\lambda$
  (see notes: $\lambda = (1\text{-}h^2/h^2)$)
- determine the matrix $Z'Z + \lambda.A^{-1}$
- invert this matrix
- multiply with the vector $Z'y^d$

- This gives you the EBV for each animal, they are in the vector *u-hat*

- Now look also at the elements in the inverse of $[Z'Z + \lambda.A^{-1}]$ and compare these with the index weights in 1)

# Rules for building the inverse of relationships matrix

For reference only

Some relevant inverses

$$\begin{array}{l} \text{Sire} \\ \text{animal} \end{array} \begin{bmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{bmatrix}^{-1} = \begin{bmatrix} \frac{4}{3} & \frac{-2}{3} \\ \frac{-2}{3} & \frac{4}{3} \end{bmatrix}$$

$$\begin{array}{l} \text{Sire} \\ \text{Dam} \\ \text{animal} \end{array} \begin{pmatrix} 1 & 0 & \frac{1}{2} \\ 0 & 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1\frac{1}{2} & \frac{1}{2} & -1 \\ \frac{1}{2} & 1\frac{1}{2} & -1 \\ -1 & -1 & 2 \end{pmatrix}$$

For each animal which is to have an estimate of u, add to $A^{-1}$:

| | Both parents known | One parent known | Neither parent known |
|---|---|---|---|
| Own diagonal | 2 | 4/3 | 1 |
| parent x animal | -1 | -2/3 | |
| parents' diagonals | 1/2 | 1/3 | |
| parent x parent | 1/2 | | |

*Challenge:*

1) Calculate the EBV of all animals using a mixed model.

In a <u>mixed model</u> the fixed and random effects are estimated jointly: y = Xb + Zu + e

Hint: in Excel you can build an X, the Z matrix, the y vector (note, use y rather than $y^d$!) and the relationships matrix, and you can build up the mixed model equations

$$\begin{bmatrix} \hat{b} \\ \hat{u} \end{bmatrix} = \begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z + \lambda A^{-1} \end{bmatrix}^{-1} \begin{bmatrix} X'y \\ Z'y \end{bmatrix}$$