

With respect to estimating marker effects using fixed linear effect models, Least Squares cannot be used because $N \ll k$ where N is the number of animals in the reference population (5000) and k is the number of SNP-positions.

Solutions:

1. LASSO: reduce number of SNPs and estimate marker effects.

2. Bayesian procedure

3. Mixed linear effect model

$$y = \underbrace{X\beta}_{\text{fixed}} + \underbrace{Wq + e}_{\text{random marker effects}}$$

$$\text{MLE: } \begin{bmatrix} X^T X & X^T W \\ W^T X & W^T W + I \cdot \lambda \end{bmatrix} \begin{bmatrix} \hat{\beta} \\ \hat{q} \end{bmatrix} = \begin{bmatrix} X^T y \\ W^T y \end{bmatrix}$$

$$\text{with } \lambda = \sigma_e^2 / \sigma_q^2$$

$$\text{assume: } \text{var}(q) = I \cdot \sigma_q^2$$