

Applied Statistical Methods - Exercise 11

Peter von Rohr

2022-05-22

Problem 1: Marker Effects Model

Predict genomic breeding values using a marker effects model. The dataset is available from

https://charlotte-ngs.github.io/asmss2022/data/asm_genosim_data.csv

Hints

- The variance σ_q^2 of the marker effect is 3.
- The residual variance σ_e^2 is 36
- The sex of each animal can be modelled as a fixed effect

Problem 2: Breeding Value Based Model

Use the same dataset as in Problem 1 to predict genomic breeding values based on a breeding-value model. The dataset is available from

https://charlotte-ngs.github.io/asmss2022/data/asm_genosim_data.csv

Hints

- The genomic variance σ_g^2 of the marker effect is 9.
- The residual variance σ_e^2 is 36
- The sex of each animal can be modelled as a fixed effect
- Use the following function to compute the genomic relationship matrix G based on the matrix of genotypes

```
computeMatGrm <- function(pmatData) {
  matData <- pmatData
  # check the coding, if matData is -1, 0, 1 coded, then add 1 to get to 0, 1, 2 coding
  if (min(matData) < 0) matData <- matData + 1
  # Allele frequencies, column vector of P and sum of frequency products
  freq <- apply(matData, 2, mean) / 2
  P <- 2 * (freq - 0.5)
  sumpq <- sum(freq*(1-freq))
  # Changing the coding from (0,1,2) to (-1,0,1) and subtract matrix P
  Z <- matData - 1 - matrix(P, nrow = nrow(matData),
                            ncol = ncol(matData),
                            byrow = TRUE)
  # Z%*%Zt is replaced by tcrossprod(Z)
  return(tcrossprod(Z)/(2*sumpq))
}
```

- If the genomic relationship matrix G which is computed by the function above cannot be inverted, add $0.05 * I$ to G which results in G^* and use G^* as genomic relationship matrix.