File S3: the likelihood is constant for a kinship matrix with compound symmetry structure

Mixed model based estimation of heritability using genotypic means may become problematic when the sample size is small and the kinship matrix is close to compound symmetry, i.e. the structure where all off-diagonal elements are equal. Here we show that in the case the kinship matrix is exactly compound symmetry, the likelihood is constant in $\eta = \sigma_E^2/\sigma_A^2$. We use the notation η to avoid confusion with $\delta = \sigma_A^2/\sigma_E^2$, used in our results on genomic prediction. We write 1_n for the $n \times 1$ column vector of ones, and I_n for the n-dimensional identity matrix. Finally, let J_n be the $n \times n$ matrix of ones.

Suppose that $K = I_n + aJ_n$, for some $a > 0$. The key observation is that the covariance matrix of the data can be written as

$$\Sigma = \sigma_A^2 K + \sigma_E^2 I_n = \tilde{\sigma}_A^2 \tilde{K} + \tilde{\sigma}_E^2 I_n = \tilde{\sigma}_A^2(\tilde{K} + \tilde{\eta} I_n),$$

where $\tilde{\sigma}_A^2 = \sigma_A^2$, $\tilde{\sigma}_E^2 = \sigma_A^2 + \sigma_E^2$, $\tilde{K} = aJ_n = a1_n1_n^t$ and $\tilde{\eta} = (\sigma_A^2 + \sigma_E^2)/\sigma_A^2 > 0$. We can then directly apply the results in section 3 of [5], with (in their notation) $k = 1$, $d = 1$ and $X = 1_n$ (we only include an intercept, and no marker effect), and replacing σ_A^2, σ_E^2, η and K by respectively $\tilde{\sigma}_A^2$, $\tilde{\sigma}_E^2$, $\tilde{\eta}$ and $\tilde{K}$. In particular, we have the spectral decomposition

$$\begin{aligned}\tilde{K} = USU^t &= [U_1, U_2]\left[\begin{array}{cc} S_1 & 0 \\ 0 & S_2 \end{array}\right][U_1, U_2]^t \\ &= \begin{pmatrix} n^{-\frac{1}{2}} & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ n^{-\frac{1}{2}} & 0 & \dots & 0 \end{pmatrix} \begin{pmatrix} na & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \dots & \dots & 0 \end{pmatrix} \begin{pmatrix} n^{-\frac{1}{2}} & \dots & n^{-\frac{1}{2}} \\ 0 & \dots & 0 \\ 0 & \dots & 0 \end{pmatrix},\end{aligned}$$

i.e. the only non-zero eigenvalue of $\tilde{K}$ is an, with eigenvector $(n^{-\frac{1}{2}}, \dots, n^{-\frac{1}{2}})$.

For this choice of X and $\tilde{K}$, the expressions for the (RE)ML estimates of β and $\tilde{\sigma}_A^2$ given in sections 3.2 and 4 of [5] greatly simplify: $\hat{\beta} = \bar{y}$ and the REML-estimate of $\tilde{\sigma}_A^2$ is $\sum_{i=1}^n (y_i - \bar{y})^2/(\tilde{\eta}(n-1))$. The extra terms

$$\frac{1}{2}\left(d \log(2\pi\tilde{\sigma}_A^2) + \log|X^tX| - \log|X^t(\tilde{K} + \tilde{\eta}I_n)^{-1}X|\right)$$

in the REML-log-likelihood (see the first equation in section 4 of [5]), now equal

$$\frac{1}{2}\left(d \log(2\pi\tilde{\sigma}_A^2) + \log n - \log\left(\frac{n}{na + \tilde{\eta}}\right)\right).$$

Combining this with their equation (3.7), it follows that the REML-log-likelihood is constant in $\tilde{\eta} = (\sigma_A^2 + \sigma_E^2)/\sigma_A^2 > 0$, and hence also constant in $\eta = \tilde{\eta} - 1 = \sigma_E^2/\sigma_A^2$.