STAT611 **Homework 6** **Solutions**

1. Prove result 9 on page 1 of the handout on the general linear model.

   Let
   $$\underline{v} = W^{-1/2}\underline{y}, \quad Z = W^{-1/2}X, \quad \underline{\delta} = W^{-1/2}\underline{\varepsilon}.$$
   Then we have $\underline{v} = Z\underline{\beta} + \underline{\delta}$, where $\text{rank}(Z_{T\times K}) = K^*$, $E(\underline{\delta}) = \underline{0}$, $E(\underline{\delta}\underline{\delta}') = D(\underline{\delta}) = \sigma^2 I$, and
   $$\text{Cov}(\underline{\delta}, \underline{u}) = \text{Cov}(W^{-1/2}\underline{\varepsilon}, \underline{u}) = W^{-1/2}\text{Cov}(\underline{\varepsilon}, \underline{u}) = \sigma^2 W^{-1/2}M.$$
   By the theorem proved in our course notes, we know that the MDLUP of $L'\underline{\beta} + \underline{u}$ is
   $$\begin{aligned} & L'(Z'Z)^{-}Z'\underline{v} + (W^{-1/2}M)'(\underline{v} - Z(Z'Z)^{-}Z'\underline{v}) \\ = \; & L'[(W^{-1/2}X)'W^{-1/2}X)]^{-}(W^{-1/2}X)'W^{-1/2}\underline{y} \\ & + M'W^{-1/2}(W^{-1/2}\underline{y} - W^{-1/2}X[(W^{-1/2}X)'W^{-1/2}X]^{-}(W^{-1/2}X)'W^{-1/2}\underline{y}) \\ = \; & L'(X'W^{-1}X)^{-}X'W^{-1}\underline{y} + M'W^{-1}(\underline{y} - X(X'W^{-1}X)^{-}X'W^{-1}\underline{y}) \\ = \; & L'\hat{\underline{\beta}} + M'W^{-1}(\underline{y} - X\hat{\underline{\beta}}) \end{aligned}$$
   where $\hat{\underline{\beta}} = (X'W^{-1}X)^{-}X'W^{-1}\underline{y}$. Because the class of linear functions of $\underline{v}$ is the same as the class of linear functions of $\underline{y}$, the result follows. $\square$

2. In animal breeding, one male can be mated to many females to produce offspring. One goal is to identify the male or males that will produce the most valuable offspring in the long run, based on a sample of offspring produced by each male. Consider, for example, a classic application in dairy cattle breeding. Each of several males (referred to as sires) produces one or more daughters. The value of the daughter is measured by the average monthly milk production over a specific time period. The goal is to rank the sires based on the milk production data of their daughters. Suppose the following model is appropriate for the milk production data.
   $$y_{ij} = \mu + s_i + e_{ij} \quad i = 1, \ldots, m; \quad j = 1, \ldots, n_i;$$
   where y_{ij} denotes the average monthly milk production from the j^{th} daughter of sire i; $s_1, \ldots, s_m \overset{i.i.d.}{\sim} N(0, \sigma_s^2)$ are random sire effects; and $e_{i1}, \ldots, e_{in_i} \overset{i.i.d.}{\sim} N(0, \sigma^2)$ are random daughter effects for the i^{th} sire $(i = 1, \ldots, m)$. All random effects are assumed to be independent. Note that, conditional on the i^{th} sire effect, the mean milk production of the daughters of sire i is $\mu + s_i$. Thus, $\mu + s_i$ represents the value of the i^{th} sire, and we seek to predict $\mu + s_i$ for all $i = 1, \ldots, m$ to rank the sires. Suppose σ_s^2/σ^2 is equal to a known constant c. Determine

an expression for the MDLUP of $\mu + s_i$ and show that it can be expressed as an intuitively appealing convex combination.

We can write this model as

$$\underline{y} = X\underline{\beta} + Z\underline{u} + \underline{e} \text{ where } X = \underline{1},\ \underline{\beta} = \mu, Z = \text{BlockDiag}(\underline{1}_{n_1}, \ldots, \underline{1}_{n_m}),\ \underline{u} = (s_1, \ldots, s_m)'.$$

With $\underline{\varepsilon} \equiv Z\underline{u} + \underline{e}$, we have

$$D(\underline{\varepsilon}) = \sigma_s^2 ZZ' + \sigma^2 I = \sigma^2 \text{BlockDiag}(I + c\underline{1}_{n_1}\underline{1}'_{n_1}, \ldots, I + c\underline{1}_{n_m}\underline{1}'_{n_m}) \equiv \sigma^2 W.$$

We need to predict $\mu + s_i$. We will obtain the result for $\mu + s_1$ and then generalize. First note that

$$\text{Cov}(\underline{\varepsilon}, s_1) = \text{Cov}(Z\underline{u} + \underline{e}, s_1) = \text{Cov}(Z\underline{u}, s_1) = \sigma_s^2[\underline{1}'_{n_1}, \underline{0}']' = \sigma^2[c\underline{1}'_{n_1}, \underline{0}']' \equiv \sigma^2 M.$$

Now we need to compute $\hat{\mu} + M'W^{-1}(\underline{y} - \hat{\mu})$. We will first compute $\hat{\mu} = (X'W^{-1}X)^{-1}X'W^{-1}\underline{y}$. Note that

$$\begin{aligned} W^{-1} &= \text{BlockDiag}[(I + c\underline{1}_{n_1}\underline{1}'_{n_1})^{-1}, \ldots, (I + c\underline{1}_{n_m}\underline{1}'_{n_m})^{-1}] \\ &= \text{BlockDiag}\left[I - \frac{c}{1 + cn_1}\underline{1}_{n_1}\underline{1}'_{n_1}, \ldots, I - \frac{c}{1 + cn_m}\underline{1}_{n_m}\underline{1}'_{n_m}\right] \text{ and} \end{aligned}$$

$$\begin{aligned} X'W^{-1}X &= \underline{1}'W^{-1}\underline{1} = \sum_{i=1}^m \underline{1}'\left[I - \frac{c}{1 + cn_i}\underline{1}_{n_i}\underline{1}'_{n_i}\right]\underline{1} = \sum_{i=1}^m \left(n_i - \frac{cn_i^2}{1 + cn_i}\right) \\ &= \sum_{i=1}^m \left(\frac{n_i}{1 + cn_i}\right) = \sigma^2 \sum_{i=1}^m \left(\frac{1}{\sigma^2/n_i + \sigma_s^2}\right). \end{aligned}$$

Similarly

$$X'W^{-1}\underline{y} = \sigma^2 \sum_{i=1}^m \left(\frac{1}{\sigma^2/n_i + \sigma_s^2}\right)\bar{y}_{i\cdot}.$$

Thus

$$\hat{\underline{\beta}} = \hat{\mu} = (X'W^{-1}X)^{-1}X'W^{-1}\underline{y} = \frac{\sum_{i=1}^m \left(\frac{1}{\sigma^2/n_i + \sigma_s^2}\right)\bar{y}_{i\cdot}}{\sum_{i=1}^m \left(\frac{1}{\sigma^2/n_i + \sigma_s^2}\right)},$$

which is a weighted average of the sire means with the inverses of the variances of the sire means as weights.

Now note that

$$\begin{aligned} M'W^{-1}\underline{y} &= c\underline{1}'\left[I - \frac{c}{1 + cn_1}\underline{1}_{n_1}\underline{1}'_{n_1}\right][y_{11}, \ldots, y_{1n_1}]' \\ &= c\sigma^2\left(\frac{1}{\sigma^2/n_1 + \sigma_s^2}\right)\bar{y}_{1\cdot} = \left(\frac{\sigma_s^2}{\sigma^2/n_1 + \sigma_s^2}\right)\bar{y}_{1\cdot}, \end{aligned}$$

and

$$\begin{aligned} M'W^{-1}X\underline{\hat{\beta}} &= c\underline{1}'\left[I - \frac{c}{1+cn_1}\underline{1}_{n_1}\underline{1}_{n_1}'\right]\underline{1}_{n_1}\hat{\mu} \\ &= c\sigma^2\left(\frac{1}{\sigma^2/n_1+\sigma_s^2}\right)\hat{\mu} = \left(\frac{\sigma_s^2}{\sigma^2/n_1+\sigma_s^2}\right)\hat{\mu}. \end{aligned}$$

Thus,

$$\begin{aligned} \hat{\mu} + M'W^{-1}(\underline{y} - \hat{\mu}) &= \hat{\mu} + \left(\frac{\sigma_s^2}{\sigma^2/n_1+\sigma_s^2}\right)\bar{y}_{1\cdot} - \left(\frac{\sigma_s^2}{\sigma^2/n_1+\sigma_s^2}\right)\hat{\mu} \\ &= \left(\frac{\sigma_s^2}{\sigma^2/n_1+\sigma_s^2}\right)\bar{y}_{1\cdot} + \left(\frac{\sigma^2/n_1}{\sigma^2/n_1+\sigma_s^2}\right)\hat{\mu}. \end{aligned}$$

This is a convex combination that blends the mean for sire 1 ($\bar{y}_{1\cdot}$) with an estimate of the population mean ($\hat{\mu}$). If the sire mean is well estimated (i.e., if σ^2/n_1 is small), most of the weight is placed on the data for sire 1. However, if σ^2/n_1 is large relative to variation among sires, the prediction will be pulled more strongly toward the estimate of the population mean. □

3. Suppose

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \sim N\left(\begin{bmatrix} \mu_1 \\ \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \sigma^2 & \sigma^2/2 & 0 \\ \sigma^2/2 & \sigma^2 & \sigma^2/2 \\ 0 & \sigma^2/2 & \sigma^2 \end{bmatrix}\right)$$

where $\mu_1 \in \mathbb{R}$, $\mu_2 \in \mathbb{R}$, and $\sigma^2 > 0$ are unknown parameters. Find the REML estimator of σ^2.

$$X = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \underline{\beta} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} \quad K^* = \text{rank}(X) = 2 \quad T = 3$$

Because $T - K^* = 3 - 2 = 1$, we need only one error contrast to estimate σ^2. Note

$$y_1 - y_2 \sim N(0, \sigma^2 + \sigma^2 - 2\sigma^2/2) \quad \Longleftrightarrow \quad y_1 - y_2 \sim N(0, \sigma^2).$$

Thus, the REML estimate of σ^2 is the MLE of σ^2 based on the one observation $y_1 - y_2$. The log likelihood is

$$-\log\sqrt{2\pi} - \frac{1}{2}\log\sigma^2 - \frac{1}{2}(y_1 - y_2)^2/\sigma^2.$$

Differentiating with respect to σ^2 and equating to 0 yields

$$-\frac{1}{2\sigma^2} + \frac{(y_1 - y_2)^2}{2\sigma^4} = 0.$$

Solving for σ^2 yields $(y_1 - y_2)^2$ as the REML estimator. □