

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

$$y_{ij} = \mu + s_i + e_{ij} \quad i = 1, \dots, m; \quad j = 1, \dots, n_i;$$

where y_{ij} denotes the average monthly milk production from the j^{th} daughter of sire i ; $s_1, \dots, s_m \stackrel{i.i.d.}{\sim} N(0, \sigma_s^2)$ are random sire effects; and $e_{i1}, \dots, e_{in_i} \stackrel{i.i.d.}{\sim} N(0, \sigma^2)$ are random daughter effects for the i^{th} sire ($i = 1, \dots, m$). 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 $i = 1, \dots, m$ 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.

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 .