NIASRA
NATIONAL INSTITUTE FOR APPLIED STATISTICS RESEARCH AUSTRALIA

National Institute for Applied Statistics Research Australia

University of Wollongong, Australia

Working Paper

10-19

REML to ML

Brian Cullis and Alison Smith

National Institute for Applied Statistics Research Australia, University of Wollongong,
Wollongong NSW 2522, Australia Phone +61 2 4221 5435, Fax +61 2 42214998.
Email: karink@uow.edu.au

REML to ML

Brian Cullis and Alison Smith

June 11, 2019

In the derivation of residual maximum likelihood (REML) we consider the log-likelihood function of the transformed data $\boldsymbol{L'y}$ where $\boldsymbol{L} = [\boldsymbol{L_1}\ \boldsymbol{L_2}]$ and $\boldsymbol{L_1}'\boldsymbol{X} = \boldsymbol{I}$ and $\boldsymbol{L_2}'\boldsymbol{X} = \boldsymbol{0}$. In this document we assume $\boldsymbol{X}$ to be of full column rank. We write

$$\ell(\boldsymbol{L'y}) = \ell(\boldsymbol{y_1}|\boldsymbol{y_2}) + \ell(\boldsymbol{y_2})$$

where $\boldsymbol{y_1} = \boldsymbol{L_1}'\boldsymbol{y}$ and $\boldsymbol{y_2} = \boldsymbol{L_2}'\boldsymbol{y}$. The log-likelihood can be expressed in terms of the original data vector with the use of the Jacobian of the transformation, the logarithm of which is given by $\log|\boldsymbol{L}| = \frac{1}{2}\log|\boldsymbol{L'L}|$. Thus we can write

$$\ell(\boldsymbol{y}) = \ell(\boldsymbol{y_1}|\boldsymbol{y_2}) + \ell(\boldsymbol{y_2}) + \frac{1}{2}\log|\boldsymbol{L'L}| \tag{1}$$

It can be shown that the conditional log-likelihood (ignoring constants) is given by

$$\begin{aligned}
\ell(\boldsymbol{y_1}|\boldsymbol{y_2}) &= -\frac{1}{2}\log|(\boldsymbol{X'H}^{-1}\boldsymbol{X})^{-1}| - \frac{1}{2}(\boldsymbol{y_1} - \boldsymbol{\tau} - \boldsymbol{y_2}^*)'(\boldsymbol{X'H}^{-1}\boldsymbol{X})(\boldsymbol{y_1} - \boldsymbol{\tau} - \boldsymbol{y_2}^*) && (2)\\
&= -\frac{1}{2}\log|(\boldsymbol{X'H}^{-1}\boldsymbol{X})^{-1}| - \frac{1}{2}(\boldsymbol{y} - \boldsymbol{X\tau})'(\boldsymbol{H}^{-1} - \boldsymbol{P})(\boldsymbol{y} - \boldsymbol{X\tau})\\
&= -\frac{1}{2}\log|(\boldsymbol{X'H}^{-1}\boldsymbol{X})^{-1}| - \frac{1}{2}(\boldsymbol{y} - \boldsymbol{X\tau})'\boldsymbol{H}^{-1}(\boldsymbol{y} - \boldsymbol{X\tau}) + \frac{1}{2}\boldsymbol{y'Py} && (3)
\end{aligned}$$

where $\boldsymbol{y_2}^* = \boldsymbol{L_1}'\boldsymbol{H}\boldsymbol{L_2}(\boldsymbol{L_2}'\boldsymbol{H}\boldsymbol{L_2})^{-1}\boldsymbol{y_2}$.

The marginal log-likelihood, which is also known as the residual log-likelihood, is given by (ignoring constants)

$$\ell(\boldsymbol{y_2}) = -\frac{1}{2}\left\{\log|\boldsymbol{H}| + \log|\boldsymbol{X'H}^{-1}\boldsymbol{X}| + \boldsymbol{y'Py}\right\} - \frac{1}{2}\log|\boldsymbol{L'L}| \tag{4}$$

Finally, substituting equations (3) and (4) into equation (1) gives:

$$\ell(\boldsymbol{y}) = -\frac{1}{2}\log|\boldsymbol{H}| - \frac{1}{2}(\boldsymbol{y} - \boldsymbol{X\tau})'\boldsymbol{H}^{-1}(\boldsymbol{y} - \boldsymbol{X\tau})$$

which is the (standard and full) log-likelihood equation.

Now we consider estimation of the corresponding linear mixed model using ASReml-R. This gives rise to REML estimates of the variance parameters and we let $\hat{\boldsymbol{H}}$ denote the resultant variance matrix. We also obtain E-BLUEs of the fixed effects. We note that these can be derived from the conditional distribution of $\boldsymbol{y_1}$ given $\boldsymbol{y_2}$, leading to the form

$$\hat{\boldsymbol{\tau}} = \boldsymbol{y_1} - \boldsymbol{L_1}'\hat{\boldsymbol{H}}\boldsymbol{L_2}(\boldsymbol{L_2}'\hat{\boldsymbol{H}}\boldsymbol{L_2})^{-1}\boldsymbol{y_2}$$

Thus, at convergence of the model fitting process we have that the quadratic form in equation (2) is zero so that conditional log-likeliood is given by

$$\begin{aligned} \hat{\ell}(\boldsymbol{y_1}|\boldsymbol{y_2}) &= -\frac{1}{2}\log|(\boldsymbol{X}'\hat{\boldsymbol{H}}^{-1}\boldsymbol{X})^{-1}| \\ &= -\frac{1}{2}\log|\text{var}\,(\hat{\boldsymbol{\tau}})\,| \end{aligned} \tag{5}$$

where var $(\hat{\boldsymbol{\tau}})$ is the variance matrix of the E-BLUEs of the fixed effects.

In terms of the marginal log-likelihood we note that ASReml-R does not include $-\frac{1}{2}\log|\boldsymbol{L}'\boldsymbol{L}|$ in the calculation, so that, at convergence, the residual log-likelihood reported by ASReml-R is given by

$$\begin{aligned} \hat{\ell}_A &= -\frac{1}{2}\left\{\log|\hat{\boldsymbol{H}}| + \log|\boldsymbol{X}'\hat{\boldsymbol{H}}^{-1}\boldsymbol{X}| + \boldsymbol{y}'\hat{\boldsymbol{P}}\boldsymbol{y}\right\} \\ &= \hat{\ell}(\boldsymbol{y_2}) + \frac{1}{2}\log|\boldsymbol{L}'\boldsymbol{L}| \end{aligned} \tag{6}$$

The so-called full log-likelihood can then be computed from the residual log-likelihood by substituting using equations (5) and (6) into (1):

$$\begin{aligned} \hat{\ell}(\boldsymbol{y}) &= \hat{\ell}(\boldsymbol{y_1}|\boldsymbol{y_2}) + \hat{\ell}(\boldsymbol{y_2}) + \frac{1}{2}\log|\boldsymbol{L}'\boldsymbol{L}| \\ &= \hat{\ell}_A - \frac{1}{2}\log|\text{var}\,(\hat{\boldsymbol{\tau}})\,| \end{aligned} \tag{7}$$

This is a log-likelihood for the full data vector, but based on REML estimates of variance parameters rather than ML.

In the special case where the only fixed effect is an overall mean, so that $\boldsymbol{\tau} = \mu$, say, and $\boldsymbol{X} = \mathbf{1}_n$ then var $(\hat{\mu})$ is easily obtained in ASReml-R using various methods. When there are more fixed effects, the "Cfixed" option can be used but care is needed here since, in general, ASReml-R works with $\boldsymbol{X}$ that has less than full column rank.