IMPROVED ITERATIVE SCHEMES FOR REML ESTIMATION OF VARIANCE PARAMETERS IN LINEAR MIXED MODELS

Emma J. Knight[1], Brian R. Cullis[2]

[1]The Fenner School of Environment and Society, The Australian National University, Australia
[1]Biometrics Unit, NSW Department of Primary Industries, Australia

Linear mixed models are a powerful class of models used for the analysis of correlated data. To estimate fixed and random effects we need estimates of the variance parameters. Residual Maximum Likelihood (REML) estimation is a popular method of estimation of variance parameters. Typically REML estimation requires an iterative scheme. We will review several of these iterative schemes, including the average information (AI) algorithm, the EM algorithm and the parameter expanded EM (PXEM) algorithm. The AI algorithm is an efficient algorithm to use when the starting values are within the neighbourhood of the REML estimates of the variance parameters. However, when reasonable starting values are not available, the AI algorithm can fail to converge, particularly in models with a complex variance structure. The EM and PXEM algorithms are good alternatives in these situations, however they can be very slow to converge.

We present a series of composite algorithms where EM-type algorithms are used initially to obtain variance parameter estimates that are in the neighbourhood of the REML estimates and then the AI algorithm is used to ensure rapid convergence. We also look at non-iterative methods of obtaining starting values to use for the AI algorithm. The performance of these various schemes is investigated for a number of different linear mixed models, ranging from a simple variance components model to a model with an unstructured random effects variance matrix. The performance of the basic algorithms is compared to that of the various composite algorithms, using both uninformed and informed starting values. The composite schemes with uninformed starting values are shown to work equally as well as the AI algorithm with informed starting values.