Clustering via Mixture Regression Models with Random Effects

Geoffrey J. McLachlan[1], S.K. Ng[2], and Kui Wang[3]

[1] Department of Mathematics & Institute for Moelcular Bioscience,
University of Queensland, Brisbane 4072, Australia, *gjm@maths.uq.edu.au*
[2] School of Medicine, Griffith University,
University Drive, Meadowbrook QLD 4131, Australia, *s.ng@griffith.edu.au* and
Department of Mathematics, University of Queensland,
Brisbane 4072, Australia, *kwang@maths.uq.edu.au*

Abstract. In this paper, we consider the use of mixtures of linear mixed models to cluster data which may be correlated and replicated and which may have covariates. For each cluster, a regression model is adopted to incorporate the covariates, and the correlation and replication structure in the data are specified by the inclusion of random effects terms. The procedure is illustrated in its application to the clustering of gene-expression profiles.

Keywords: Correlated data, random effects, mixed linear models, gene profiles