

A Comparison of the L_2 Minimum Distance Estimator and the EM-Algorithm when Fitting k -Component Univariate Normal Mixtures

B.R. Clarke^{1*}, T. Davidson² and R. Hammarstrand¹

¹ *Mathematics and Statistics, School of Eng. and I.T., Murdoch University, Murdoch, Western Australia, 6150; B.Clarke@murdoch.edu.au, R.Hammarstrand@murdoch.edu.au.*

² *Australian Bureau of Statistics, Perth, Western Australia, 6000; tom.davidson@abs.gov.au*

**Presenting author*

Keywords. *EM algorithm; Minimum distance estimation; Robust estimation; Monte Carlo simulation.*

The method of maximum likelihood using the EM-algorithm for fitting finite mixtures of normal distributions is the accepted method of estimation ever since it has been shown to be superior to the method of moments. Recent books testify to this. There has however been criticism of the method of maximum likelihood for this problem, the main criticism being when the variances of component distributions are unequal the likelihood is in fact unbounded and there can be multiple local maxima. Another major criticism is that the maximum likelihood estimator is not robust. Several alternative minimum distance estimators have since been proposed as a way of dealing with the first problem. This paper deals with one of these estimators which is not only superior due to its robustness, but in fact can have an advantage in numerical studies even at the model distribution. Importantly, robust alternatives of the EM-algorithm, ostensibly fitting t distributions when in fact the data are mixtures of normals, are also not competitive at the normal mixture model when compared to the chosen minimum distance estimator. It is argued for instance that natural processes should lead to mixtures whose component distributions are normal as a result of the Central Limit Theorem. On the other hand data can be contaminated because of extraneous sources as are typically assumed in robustness studies. This calls for a robust estimator.

Relevant references include ?, ?, ?, ?, and ?.

References

- Clarke, B. R. (1983). Uniqueness and Fréchet differentiability of functional solutions to maximum likelihood type equations. *Annals of Statistics*, **11**, 1196-1205.
- Clarke, B. R. (1989). An unbiased minimum distance estimator of the proportion

- parameter in a mixture of two normal distributions. *Statist. Prob. Lett.*, **7**, 275-281.
- Clarke, B. R. (2000). A review of differentiability in relation to robustness with an application to seismic data analysis. *PINSA*, **66A**, 467-482.
- Clarke, B. R., Davidson, T. & Hammarstrand, R. (2016). A Comparison of the L_2 Minimum Distance Estimator and the EM-Algorithm when Fitting k -Component Univariate Normal Mixtures. *Statistical Papers*, <http://link.springer.com/article/10.1007/s00362-016-0747-x>
- Clarke, B. R. & Heathcote, C. R. (1978). Comment on "Estimating mixtures of normal distributions and switching regressions" by Quandt, R.E. and Ramsey, J.B.. *J. Am. Statist. Ass.*, **73**, 749-750.
- Clarke, B. R. & Heathcote, C. R. (1994). Robust estimation of k -component univariate normal mixtures. *Ann. Inst. Statist. Math.*, **46**, 83-93.