Statistica Sinica Preprint No: SS-2016-0483	
Title	A CLASSICAL INVARIANCE APPROACH TO THE NORMAL MIXTURE PROBLEM
Manuscript ID	SS-2016-0483
URL	http://www.stat.sinica.edu.tw/statistica/
DOI	10.5705/ss.202016.0483
Complete List of Authors	Monia Ranalli Bruce G. Lindsay and David R. Hunter
Corresponding Author	Monia Ranalli
E-mail	mxr459@psu.edu

Statistica Sinica: Preprint
doi:10.5705/ss.202016.0483

$(1-p)$, thereby achieving the same density. However, Equation (1) is otherwise a regular model, and the mapping is locally identifiable in the parameter space. We consider a random sample $\mathbf{X}_1, \dots, \mathbf{X}_n$ from this density.

Let us define $\mathbf{U}_i = \mathbf{X}_i - \mathbf{X}_n$ for $1 \le i \le n-1$. Because $\mathbf{U}_i = (\mathbf{X}_i - \mathbf{a}) - (\mathbf{X}_n - \mathbf{a})$ for any vector $\mathbf{a}$, letting $\mathbf{a} = \boldsymbol{\mu}_1$ we notice that the distribution of $\mathbf{U}_i$ must depend on the distribution of $\mathbf{X}_1 - \boldsymbol{\mu}_1, \dots, \mathbf{X}_n - \boldsymbol{\mu}_1$, which is a sample from a mixture of normals with means 0 and $\boldsymbol{\delta} = \boldsymbol{\mu}_2 - \boldsymbol{\mu}_1$. We conclude that the distribution of $\mathbf{U} = (\mathbf{U}_1, \dots, \mathbf{U}_{n-1})$ depends only on the parameters $\boldsymbol{\tau} = (p, \boldsymbol{\delta}, \boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2)$. We therefore introduce the notation $h_{\boldsymbol{\tau}}(\mathbf{u})$ for the density of $\mathbf{U}$, and refer to $h_{\boldsymbol{\tau}}(\mathbf{u})$ as the marginal likelihood. Similarly, we call $f_{\boldsymbol{\theta}}(\mathbf{x})$ and $f_{\boldsymbol{\theta}}(\mathbf{x}|\mathbf{u})$ the full and conditional likelihoods, respectively, when they are viewed as functions of $\boldsymbol{\theta}$. We may derive an expression for $h_{\boldsymbol{\tau}}(\mathbf{u})$ by integrating out $\mathbf{x}_n$ from the joint density of $(\mathbf{U}, X_n)$:

$$h_{\boldsymbol{\tau}}(\mathbf{u}) = \int f_{\boldsymbol{\theta}}(\mathbf{u}_1 + \mathbf{x}) \cdots f_{\boldsymbol{\theta}}(\mathbf{u}_{n-1} + \mathbf{x}) f_{\boldsymbol{\theta}}(\mathbf{x})\, d\mathbf{x}. \tag{2}$$

Under the change of variables $\mathbf{t} = \mathbf{x} + \mathbf{x}_n$ and recalling that $\mathbf{x}_i = \mathbf{u}_i + \mathbf{x}_n$, we may rewrite Equation (2) as

$$h_{\boldsymbol{\tau}}(\mathbf{u}) = \int f_{\boldsymbol{\theta}}(\mathbf{x}_1 + \mathbf{t}) \cdots f_{\boldsymbol{\theta}}(\mathbf{x}_{n-1} + \mathbf{t}) f_{\boldsymbol{\theta}}(\mathbf{x}_n + \mathbf{t})\, d\mathbf{t}. \tag{3}$$

Equation (3) re-expresses the marginal likelihood in terms of the complete data, which is a computational trick we will use later in proving that the marginal likelihood is bounded in the case of univariate data. The form of Equation (3) also makes clear that the choice of $\mathbf{X}_n$ in the definition $\mathbf{U}_i = \mathbf{X}_i - \mathbf{X}_n$ is arbitrary, made here simply for convenience of notation.

This study proposes a two-stage estimation algorithm. The first stage estimates the