

RECOGNITION OF A MIXTURE OF MULTIPLE GAUSSIAN PATTERNS

A.C. Mkolesia^{1 §}, C.R. Kikawa², M.Y. Shatalov³, B.M. Kalema⁴

^{1,2,3}Department of Mathematics and Statistics
Tshwane University of Technology
175 Nelson Mandela Drive, Pretoria, 0001, RSA

⁴Department of Informatics
Tshwane University of Technology
2 College Road, Block L, Pretoria North, 0116, RSA

Abstract: In this paper a methodology for the recognition of multiple Gaussian patterns by estimating sufficient parameters of a finite mixture model (FMM) is proposed. Regular methods of FMM identification require initial guess values (IGVs) that may result in high computation time, slow convergence and or even fail to converge if the provided IGVs are far from the optimal solution. The FMM is firstly decomposed into it's even and odd parts, which are linearised through differential techniques. Secondly the ordinary least squares (OLS) method is employed to estimate the unknown parameters in the linearised models. A Monte Carlo simulation is done to evaluate the performance of the proposed method (PM). It is shown that numerical results of the PM compare well with the simulated values. The study indicates that (i) the PM can be used symbiotically with the regular methods to compute IGVs; and (ii) can be used to estimate a general n-component Gaussian model.

AMS Subject Classification: 62F15, 62C10

Key Words: finite mixture models, Gaussian distributions, parameter estimation, ordinary least squares, ordinary differential equation

1. Introduction

Data measurements of real world phenomena are modeled by a mixture of con-

Received: April 4, 2016

Published: June 20, 2016

© 2016 Academic Publications, Ltd.

url: www.acadpubl.eu

[§]Correspondence author

tinuous density functions (see [1]) , these data may replicate finite mixture models (see [2]) (FMMs) that is a mixture of Gaussian distributions, and the mixture models are capable of approximating any arbitrary distribution (see [1]). The FMM provides a natural representation of the heterogeneity of a finite number of latent classes, it concerns modeling a statistical distribution by a mixture (or weighted sum) of other distributions, that is, Gaussian distributions. A FMM provides a parametric alternative that describes the unknown distribution in terms of a mixture of known distributions (see [3]). The FMMs are also known as, latent class models or unsupervised learning models (see [1], [4], [5]). A mixture model is usually probabilistic in nature and encompasses finite or infinite sub-populations referred to as components. The FMMs are closely related to intrinsic classification models, clustering (see [6]) and numerical taxonomy. Finite mixture models have a wide range of applications in the field of signal processing (see [7]), image processing (see [5], [8], [9]), pattern recognition (see [10], [8]) , machine learning such as modelling, clustering, classification (see [11], [5], [10]) and survival analysis. Mixture models of normal densities with common variance in the univariate case, may be estimated with a continuous distribution. In pattern reconstruction, FMM permit a probabilistic model based approach for unsupervised learning (see [8]). Generally from a statistical perspective FMM is a semi parametric/non parametric estimator of the density (see [7]). Experience suggests that usually only few latent classes are needed to approximate the density function well (see [12]). In practice FMM are flexible extensions to basic parametric models and can generate skewed distributions from symmetric components. Identification of heterogeneous information presents an enormous challenge and requires highly involved computations, but n -component density functions overcome the estimation problems (see [13]).

This article is organized as follows. Section 2, describes the motivation of the proposed method. In Section 3, current methods (regular guess value method) of the estimation of parameters for finite mixture distributions containing normal distributions are investigated, a theorem is stated for estimating an general n -component mixture model. Section 4, the proposed method is discussed. In Section 5, a Monte Carlo simulation is considered to assess the performance of the proposed method. In Section 6, results and discussion of the study are presented and conclusions presented in Section 7

2. Motivation

The estimation of parameters for finite mixture distributions contain normal distributions components (see [10]). In the work by (see [14], [15]), a proposed method (PM), for estimating parameters of a two component mixture with equal prior probabilities, was considered, the two component mixture model was presented as

$$f(x, \sigma_1, \sigma_2, \mu_1, \mu_2, \lambda) = \frac{\lambda}{\sigma_1\sqrt{2\pi}}e^{\left\{\frac{-(x-\epsilon\mu_1)^2}{2\sigma_1^2}\right\}} + \frac{1-\lambda}{\sigma_2\sqrt{2\pi}}e^{\left\{\frac{-(x-\epsilon\mu_2)^2}{2\sigma_2^2}\right\}}, \quad (1)$$

where x is the study variable, σ_1^2, σ_2^2 are the variances, μ_1, μ_2 are the mean and λ is the proportion of each component in the mixture model and $\lambda > 0$.

The general form of Equation (1) is given as

$$f(x, \Psi) = A_1e^{-\alpha_1(x-\epsilon\mu_1)^2} + A_2e^{-\alpha_2(x-\epsilon\mu_2)^2}, \quad (2)$$

where $\Psi = (A_1, A_2, \mu_1, \mu_2, \alpha_1, \alpha_2)$.

From Equation (2), Ψ is a complete collection of distinct parameters occurring in the mixture model and ϵ is the “small” parameter, which characterises smallness of the distance of the two means μ_1 and μ_2 from their absolute mean in the mixture. The goal here is now to estimate the parameters of Ψ .

In the case of an n -component distribution Equation (1) can be written as

$$f(x, \sigma_i, \mu_i, \lambda_i) = \sum_{i=1}^n \lambda_i \frac{1}{\sigma_i\sqrt{2\pi}}e^{\left\{\frac{-(x-\epsilon\mu_i)^2}{2\sigma_i^2}\right\}}, \quad (3)$$

where $\sum_{i=1}^n \lambda_i = 1$ for $\lambda_i \in [0, 1]$.

The general form of Equation (3) for a n -component normal mixture is given as

$$f(x, \Psi) = A_1e^{-\alpha_1(x - \epsilon\mu_1)^2} + A_2e^{-\alpha_2(x - \epsilon\mu_2)^2} + \dots + A_n e^{-\alpha_n(x - \epsilon\mu_n)^2}, \quad (4)$$

where $\Psi = (A_i, \mu_i, \alpha_i)$ for $i = 1, 2, 3, \dots, n$.

There exist several methods used to compute FMM parameters, that is, Expectation maximization (EM), Newton-Gauss, Bayesian, Markov chain Monte Carlo, Spectral method, Graphical Methods etc. However all these methods use iterative routines (see [2], [13]) where initial guess values must be provided to start the iterative process (see [5]) before computing the required optimal solutions. The challenge with initial guess value routines are, (i) slow convergence (see [10], [13]); (ii) High computational time (see [1]) and (iii) may fail to converge if the initialization values are far from the required solution.

3. Current Methods of Solution

As mentioned in Section 2, the existing methods of solution use iteration procedures to determine the optimal solutions.

Consider the three component Mixture Model given as

$$f(x) = \sum_{i=1}^n A_i e^{-\alpha_i(x-\epsilon\mu_i)^2}, \text{ for } n = 3. \quad (5)$$

The function $f(x)$ in Equation (5) is decomposed into its even and odd parts

$$f_e(x) = \frac{1}{2} [f(x) + f(-x)], \quad (6)$$

and

$$f_o(x) = \frac{1}{2} [f(x) - f(-x)]. \quad (7)$$

Considering the expansion of $f_e(x)$ and $f_o(x)$ in series with respect to ϵ (at $\epsilon = 0$).

At first approximation with error of order one, Equation (5) yields Equations (8 & 9)

$$f_e(x) \stackrel{\text{def}}{=} \sum_{i=1}^n A_i e^{-\alpha_i x^2} + O(\epsilon), \text{ for } n = 3. \quad (8)$$

$$\frac{f_o(x)}{x} \stackrel{\text{def}}{=} \sum_{i=1}^n B_i e^{-\alpha_i x^2} + O(\epsilon), \text{ for } n = 3. \quad (9)$$

From Equation (8) the following parameters can be approximated A_i and α_i . From Equation (9) the following parameters can be obtained a_i where

$$a_i = \frac{B_i}{2\alpha_i A_i}.$$

At second approximation with error of order two, Equation (5) yields Equations (10 & 11)

$$f_e(x) \stackrel{\text{def}}{=} \sum_{i=1}^n \left(\tilde{A}_i + \tilde{B}_i x^2 \right) e^{-\alpha_i x^2} + O(\epsilon^2), \tag{10}$$

where $\tilde{A}_i = A_i (1 - \alpha_i (\epsilon \mu_i)^2)$
and $\tilde{B}_i = 2A_i \alpha_i^2 (\epsilon \mu_i)^2$.

$$\frac{f_o(x)}{x} \stackrel{\text{def}}{=} \sum_{i=1}^n \left(\tilde{C}_i + \tilde{D}_i x^2 \right) e^{-\alpha_i x^2} + O(\epsilon), \tag{11}$$

where $\tilde{C}_i = A_i 2\alpha_i \epsilon \mu_i (1 - \alpha_i (\epsilon \mu_i)^2)$
and $\tilde{D}_i = A_i \frac{4}{3} \alpha_i^3 (\epsilon \mu_i)^3$.

By change of parameters, let $Z = x^2$ Equation (10 & 11) can be written as

$$f_e(x) \stackrel{\text{def}}{=} \sum_{i=1}^n \left(\tilde{A}_i + \tilde{B}_i Z \right) e^{-\alpha_i x^2} + O(\epsilon^2). \tag{12}$$

$$\frac{f_o(x)}{x} \stackrel{\text{def}}{=} \sum_{i=1}^n \left(\tilde{C}_i + \tilde{D}_i Z \right) e^{-\alpha_i x^2} + O(\epsilon). \tag{13}$$

Equation (12) represents the solution of an exact linear differential equation with real roots, having a characteristic equation of the form

$$\prod_{i=1}^n (\lambda + \alpha_i)^2 = 0. \tag{14}$$

Thus the solution of Equation (14) is $\lambda_{j,k} = -\alpha_i$ where $j = 1, 3, 5, \dots, n$ and $k = 2, 4, 6, \dots, n$ and $i = 1, 2, 3, \dots, n$.

Assuming that a dataset $\{x_k, f(x_k) = f_k\}$, for $k = 1, \dots, N$, is available, a goal function (G_n) where n is the number of mixture models can be used to estimate the parameters A_i , α_i and a_i of Equation (15). Using the first

approximation Equations (8 & 9), the goal function for a three component mixture model is given as

$$G_3 = G_3(A_i, \alpha_i, a_i) = \frac{1}{2} \sum_{k=0}^N \{A_i e^{-\alpha_i(x_k - \epsilon\mu_i)^2} - f_k\}^2, \quad (15)$$

$$G_3 \rightarrow \min, \text{ for } i = 1, 2, 3.$$

The approximated goal function of Equation (15), \tilde{G}_1 is given by

$$\begin{aligned} \tilde{G}_3 &= \tilde{G}_3(\tilde{A}_i, \tilde{B}_i, \tilde{C}_i, \alpha_i) \\ &= \frac{1}{2} \sum_{k=0}^N \left((\tilde{A}_i + \tilde{B}_i x_k + \tilde{C}_i x_k^2) e^{-\alpha_i x_k^2} - f_k \right)^2, \end{aligned} \quad (16)$$

$$\tilde{G}_3 \rightarrow \min, \text{ for } i = 1, 2, 3.$$

Deriving the partial derivatives of Equation (16) with respect to \tilde{A}_i , \tilde{B}_i , \tilde{C}_i and α_i . Which generate the following system of equations

$$\frac{\partial \tilde{G}_1}{\partial \tilde{A}_i} = \sum_{k=0}^N e^{-\alpha_i x_k^2} (\xi e^{-\alpha_i x_k^2} - f_k), \quad (17)$$

$$\frac{\partial \tilde{G}_1}{\partial \tilde{B}_i} = \sum_{k=0}^N x_k e^{-\alpha_i x_k^2} (\xi e^{-\alpha_i x_k^2} - f_k), \quad (18)$$

$$\frac{\partial \tilde{G}_1}{\partial \tilde{C}_i} = \sum_{k=0}^N x_k^2 e^{-\alpha_i x_k^2} (\xi e^{-\alpha_i x_k^2} - f_k), \quad (19)$$

$$\frac{\partial \tilde{G}_1}{\partial \alpha_i} = \sum_{k=0}^N \left(x_k^2 e^{-\alpha_i x_k^2} \xi \right) (f_k - e^{-\alpha_i x_k^2} \xi), \quad (20)$$

where $\xi = \tilde{A}_i + \tilde{B}_i x_k + \tilde{C}_i x_k^2$ for $i = 1, 2, 3$.

It can be observed that the least squares normal Equations (17, 18, 19 and , 20) can not be solved explicitly to give exact solutions. Iterative methods that require IGVs to initiate the iteration process have to be supplied. Thus the justification of developing the proposed method.

Theorem 1. *A generalised n -component Gaussian mixture model can be approximated using a series about a predetermined origin.*

Proof. Considering the general form of an n -component normal mixture model, Equation (4) can be generalised as

$$f(x) = \sum_{i=1}^n A_i e^{-\alpha_i(x-\epsilon\mu_i)^2}. \tag{21}$$

Equation (21) can be decomposed into even and odd.

$$f_e(x) = \sum_{i=1}^n \tilde{A}_i e^{-\alpha_i x^2} + O(\epsilon^2 x^2) + O(\epsilon^2), \tag{22}$$

where $\tilde{A}_i = A_i e^{-\alpha_i \mu_i^2 \epsilon^2}$
for $i = 1, 2, \dots, n$.

The n -component mixture model odd part is represented by

$$f_o(x) = \sum_{i=1}^n \tilde{B}_i x e^{-\alpha_i x^2} + O(\epsilon^2 x^3) + O(\epsilon^2 x), \tag{23}$$

where $\tilde{B}_i = 2A_i \alpha_i \mu_i e^{-\alpha_i \mu_i^2 \epsilon^2}$
for $i = 1, 2, \dots, n$.

The working error of the even and odd parts are of order two ($O(\epsilon^2 x^2) + O(\epsilon^2)$), the error indicates the closeness of the mean μ_i of the mixture model. Extending the order of the error say to order three will only increase the number of parameters to be estimated thus rendering the formulation to be less accurate, that is, the even and odd part with error of order three now becomes Equation (24) and Equation (25) respectively.

$$f_e(x) = \sum_{i=1}^n \tilde{A}_i e^{-\alpha_i x^2} (1 + 4\alpha_i^2 \mu_i^2 \epsilon^2 x^2) + O(\epsilon^4 x^2) + O(\epsilon^4). \tag{24}$$

$$f_o(x) = \sum_{i=1}^n \tilde{B}_i x e^{-\alpha_i x^2} + O(\epsilon^4 x^2) + O(\epsilon^4). \quad (25)$$

The odd part Equation (23), was divided by an x term so as to approximate it as an even part Equation (26). In this way the odd part (23), can be estimated as those of the even part Equations (22, 26 and 28) respectively.

$$\frac{f_o(x)}{x} = \sum_{i=1}^n \tilde{B}_i e^{-\alpha_i x^2} + O(\epsilon^2 x^2) + O(\epsilon^2), \quad (26)$$

where $\tilde{B}_i = 2A_i \alpha_i \mu_i e^{-\alpha_i \mu_i^2 \epsilon^2}$
for $i = 1, 2, \dots, n$.

Comment 2.

$$\frac{f_o(x)}{x} = \sum_{i=1}^n \tilde{B}_i e^{-\alpha_i x^2} + O(\epsilon^2 x^2) + O(\epsilon^2). \quad (27)$$

Hence

$$f_e(x) \equiv \frac{f_o(x)}{x}. \quad (28)$$

The parameters of the even function Equation (22), are

$$\tilde{A}_i = A_i e^{-\alpha_i \mu_i^2 \epsilon^2}, \quad (29)$$

for $i = 1, 2, \dots, n$.

The parameters of the odd function Equation (23), are

$$\tilde{B}_i = 2A_i \alpha_i \epsilon \mu_i e^{-\alpha_i \mu_i^2 \epsilon^2}, \quad (30)$$

for $i = 1, 2, \dots, n$.

Using the data for the n -component mixture model Equation (21) the parameter estimates \tilde{B}_i and \tilde{A}_i , for $i = 1, 2, \dots, n$ may be obtained using the ordinary least squares (OLS) method, via parametric ODEs.

In order to simplify the computations, let $y = x^2$, then $f(x) = f(\sqrt{y})$, for ease of presentation this can be written as $f(x) = f(\sqrt{y}) = \tilde{f}(y)$. Thus the even part can be represented as

$$f_e(x) = \tilde{f}_e(y) = \sum_{i=1}^n \tilde{A}_i e^{-\alpha_i y} + O(\epsilon^2 y) + O(\epsilon^2). \tag{31}$$

Similarly the odd part is represented by

$$\begin{aligned} \frac{f_o(x)}{x} = \tilde{f}_o(y) &= \sum_{i=1}^n \tilde{B}_i e^{-\alpha_i y} + O(\epsilon^2 y) + O(\epsilon^2), \\ \text{where } \tilde{B}_i &= 2A_i \alpha_i \epsilon \mu_i e^{-\alpha_i \mu_i^2 \epsilon^2}, \\ &\text{for } i = 1, 2, \dots, n. \end{aligned} \tag{32}$$

Let $Z_i = e^{-\alpha_i y}$ for $i = 1, 2, \dots, n$ this transforms Equations (31 & 32) to Equations (33 & 34) respectively.

$$\begin{aligned} f_e(x) = \tilde{f}_e(y) &= \tilde{f}_e(Z) \\ &= \sum_{i=1}^n \tilde{A}_i Z_i + O(\epsilon^2 y) + O(\epsilon^2). \end{aligned} \tag{33}$$

$$\begin{aligned} \frac{f_o(x)}{x} = \tilde{f}_o(y) &= \tilde{f}_o(Z) \\ &= \sum_{i=1}^n \tilde{B}_i Z_i + O(\epsilon^2 y) + O(\epsilon^2). \end{aligned} \tag{34}$$

By letting $Z_i = e^{-\alpha_i y}$ for $i = 1, 2, \dots, n$ in Equations (31 & 32) yields a linear transformation of the model, Equations (33 & 34) respectively.

Equation (31) is the solution to the exact linear ordinary differential equation with real roots of order n , Equation (35).

$$\begin{aligned} \frac{d^n}{dy^n} \tilde{f}_e + \eta_1 \frac{d^{n-1}}{dy^{n-1}} \tilde{f}_e + \eta_2 \frac{d^{n-2}}{dy^{n-2}} \tilde{f}_e \\ + \dots + \eta_{n-1} \frac{d}{dy} \tilde{f}_e + \eta_n \tilde{f}_e = 0. \end{aligned} \tag{35}$$

Using the multiple OLS method, the estimated values of η_i for $i = 1, 2, \dots, n$ may be obtained using the data set $\{y_i, \tilde{f}_e(y_i)\}$ for $i = 1, 2, 3, \dots, m$. The numeric computations of $\frac{d^i}{dy^i} \tilde{f}_e$ for $i = 1, 2, \dots, n$ are used to obtain the estimated parameters of η_i . The multiple regression is given by Equation (36).

$$\begin{aligned}
 Y_e &= \sum_{i=1}^n M_{e_i} X_{e_i} , \\
 \text{where } Y_e &= \frac{d^n \tilde{f}_e}{dy^n} , \quad X_{e_i} = \frac{d^{n-i} \tilde{f}_e}{dy^{n-i}} \\
 \text{and } M_{e_i} &= -\eta_i , \quad \text{for } i = 1, 2, \dots, n.
 \end{aligned}
 \tag{36}$$

The auxiliary equation for the linear ODE Equation (35) is given by

$$\gamma^n + \eta_i \gamma^{n-1} + \eta_{i+1} \gamma^{n-2} + \dots + \eta_n = 0.
 \tag{37}$$

Considering the linear roots of the auxiliary Equation (37) thus

$$\begin{aligned}
 \gamma^n + \eta_i \gamma^{n-1} + \eta_{i+1} \gamma^{n-2} + \dots + \eta_n &= \prod_{i=1}^n (\gamma + \tau_i), \\
 \text{for } i &= 1, 2, \dots, n.
 \end{aligned}
 \tag{38}$$

Thus solving for the roots of Equation (37) yields the parameters of $Z_i = e^{-\alpha_i y}$.

$$\prod_{i=1}^n (\gamma + \tau_i) = 0 , \quad \text{for } i = 1, 2, \dots, n.
 \tag{39}$$

The finite mixture dataset $\{x_i, f(x_i)\}$ for $i = 1, \dots, m$ a new dataset $\{y_i = x_i^2, f(y_i) = f(x_i^2)\}$ for $i = 1, 2, \dots, m$ can be constructed. Equation (31) can be solved using multiple ordinary least squares (OLS) method to obtain parameter estimates for \tilde{A}_i , for $i = 0, 1, 2, \dots, m$, since $Y_{ji} = e^{-\alpha_j y_i}$, where $i = 1, 2, \dots, n$ and $j = 1, 2, \dots, m$. Substituting for \tilde{A}_i , in to Equation (22), n equations are obtained but with $2n$ unknown parameters, at this stage, the parameters cannot be estimated. From Equation (22) $\tilde{A}_i = A_i e^{(-\alpha_i \mu_i^2 \epsilon^2)}$ for $i = 1, 2, \dots, n$.

The odd part can be estimated exactly as the even part, since the relation in Equation (28) holds. Thus the methodology of estimation are similar for the odd and even parts.

$$\frac{f_o(x)}{x} = \sum_{i=1}^n \tilde{B}_i e^{-\alpha_i x^2} + O(\epsilon^2 x^2) + O(\epsilon^2), \tag{40}$$

where $\tilde{B}_i = 2A_i \epsilon \mu_i \alpha_i e^{-\alpha_i \epsilon^2 (\mu_i)^2}$.

The even and odd components can be compared using Equation (31 & 32). Let $y = x^2$ then $f(x) = f(\sqrt{y})$ and, $\tilde{f}(y) = f(\sqrt{y})$.

$$\tilde{f}(y) = \sum_{i=1}^n \tilde{B}_i e^{-\alpha_i y} + O(\epsilon^2 y) + O(\epsilon^2). \tag{41}$$

Consider

$$Y_i = e^{-\alpha_i y}, \text{ for } i = 1, 2, \dots, n. \tag{42}$$

Let $\tilde{f}(y) = \tilde{f}(Y)$ and substitute Equation (42) into Equation (41) to obtain

$$\tilde{f}(z) = \sum_{i=1}^3 \tilde{B}_i Y_i. \tag{43}$$

Considering the dataset for the mixture model $\{x_i, f(x_i)\}$, using the multiple OLS on the dataset $\{x_i, f(x_i)\}$ for the mixture model Equation (43) the parameter estimates of \tilde{B}_i for $i = 1, 2, \dots, n$.

This completes the proof.

□

The function now has n equations with n unknowns. Combining the Even and Odd functions Equations (36 & 43) estimating the general unknown parameters A_i, μ_i , of the model Equation (21), the sufficient parameters of the Mixture Model can be obtained, that is, a system of n Equations and n unknowns.

4. Proposed Method

In this Section Theorem 1 is applied to a three component mixture model. The 3-component univariate normal density mixture model is considered and can

be formulated as follows

$$f(x, \Psi) = f(x, \sigma_1, \sigma_2, \sigma_3, \mu_1, \mu_2, \mu_3, \lambda, \epsilon), \quad (44)$$

where $\Psi = (\sigma_i, \mu_i, \epsilon)$ for $i = 1, 2, 3$.

A complete collection of distinct parameters for the mixture model is given by Ψ and ϵ is the “small” distance of the means μ_i for $i = 1, 2, 3$. Thus Equation (44) can be written as

$$f(x, \Psi) = \sum_{k=0}^N A_i e^{-\alpha_i(x-\epsilon\mu_i)^2}, \text{ where } A_i = \frac{\lambda_i}{\sigma_i\sqrt{2\pi}} \quad (45)$$

for $\sum \lambda_i = 1, \lambda_i \in [0, 1]$.

$$f_e(x) = \sum_{i=1}^3 \tilde{A}_i e^{-\alpha_i x^2} + O(\epsilon^2 x^2) + O(\epsilon^2), \quad (46)$$

where $\tilde{A}_i = A_i e^{-\alpha_i \epsilon^2 \mu_i^2}$.

$$f_o(x) = \sum_{i=1}^3 \tilde{B}_i x e^{-\alpha_i x^2} + O(\epsilon^2 x^2) + O(\epsilon^2), \quad (47)$$

where $\tilde{B}_i = 2A_i \epsilon \mu_i \alpha_i e^{-\alpha_i \epsilon^2 (\mu_i)^2}$.

The odd part of Equation (47) may be estimated as the even part as

$$\frac{f_o(x)}{x} = \sum_{i=1}^3 \tilde{B}_i e^{-\alpha_i x^2} + O(\epsilon^2 x^2) + O(\epsilon^2), \quad (48)$$

where $\tilde{B}_i = 2A_i \epsilon \mu_i \alpha_i e^{-\alpha_i \epsilon^2 (\mu_i)^2}$.

The parameters α_i for $i = 1, 2, 3$ for Equation (46) and Equation (48) can now be estimated using the ordinary least squares method (OLS). Using the

change of parameters $y = x^2$ Equation (46) is the Even Function Transformed.

$$\begin{aligned}
 f(x) &= f(\sqrt{y}) = \tilde{f}(y), \\
 f_e(x) &= \tilde{f}_e(y), \\
 \tilde{f}_e(y) &= \sum_{i=1}^3 \tilde{A}_i e^{-\alpha_i y} + O(\epsilon^2 y) + O(\epsilon^2).
 \end{aligned}
 \tag{49}$$

The transformation of the odd and even parts of the three FMM is given as

$$\begin{aligned}
 f_e(x) &= \tilde{f}_e(y) = \tilde{f}_e(z) = \sum_{i=1}^3 \tilde{A}_i Z_i, \\
 \text{where } Z_i &= e^{-\alpha_i y} \text{ for } i = 1, 2, 3.
 \end{aligned}
 \tag{50}$$

$$\begin{aligned}
 \frac{f_o(x)}{x} &= \frac{\tilde{f}_o(y)}{\sqrt{y}} = \frac{\tilde{f}_o(z)}{x} = \sum_{i=1}^3 \tilde{B}_i Y_i, \\
 \text{where } Y_i &= e^{-\alpha_i y} \text{ for } i = 1, 2, 3.
 \end{aligned}
 \tag{51}$$

Using the Ordinary Least Squares (OLS) to estimate the parameters \tilde{A}_i and \tilde{B}_i for $i = 1, 2, 3$ from the data set $\{x_i, f(x_i)\}$ for $i = 1, \dots, m$. Equations (50 & 51) are solutions to the homogeneous linear ordinary differential equation (ODE) with constant coefficients Equation (52).

$$\begin{aligned}
 \frac{d^3}{dy^3} \tilde{f}_e + \eta_1 \frac{d^2}{dy^2} \tilde{f}_e + \eta_2 \frac{d}{dy} \tilde{f}_e + \eta_3 \tilde{f}_e &= 0, \\
 \frac{d^3}{dy^3} \tilde{f}_o + \eta_1 \frac{d^2}{dy^2} \tilde{f}_o + \eta_2 \frac{d}{dy} \tilde{f}_o + \eta_3 \tilde{f}_o &= 0.
 \end{aligned}
 \tag{52}$$

Using the multiple OLS method, the estimated values of η_i for $i = 1, 2, 3$ maybe obtained using the data set $\{y_i, f_e(y_i)\}$ for $i = 1, 2, 3, \dots, m$.

The solutions η_i yields the values of α_i using the characteristic equation

$$\gamma^3 + \eta_1\gamma^2 + \eta_2\gamma + \eta_3 = \prod_{i=1}^3(\gamma + \tau_i), \quad (53)$$

$$(\gamma + \tau_1)(\gamma + \tau_2)(\gamma + \tau_3) = \prod_{i=1}^3(\gamma + \tau_i).$$

$$\eta_1 = \tau_1 + \tau_2 + \tau_3, \quad (54)$$

$$\eta_2 = \tau_1\tau_2 + \tau_1\tau_3 + \tau_2\tau_3,$$

$$\eta_3 = \tau_1\tau_2\tau_3.$$

Thus solving Equation (54) will yield:

$$\tau_1 = \alpha_1, \quad (55)$$

$$\tau_2 = \alpha_2,$$

$$\tau_3 = \alpha.$$

The Finite Mixture dataset $\{x_i, f(x_i)\}$ for $i = 1, \dots, m$ a new dataset $\{y_i = x_i^2, f(y_i) = f(x_i^2)\}$ for $i = 1, \dots, m$ can be constructed. Equation (50) can also be solved using multiple ordinary least squares (OLS) method to obtain parameter estimates for \tilde{A}_1 , \tilde{A}_2 and \tilde{A}_3 , since $\tilde{Y}_{1i} = e^{-\alpha_1 y_i}$, $\tilde{Y}_{2i} = e^{-\alpha_2 y_i}$ and $\tilde{Y}_{3i} = e^{-\alpha_3 y_i}$. Substituting for \tilde{A}_1 , \tilde{A}_2 and \tilde{A}_3 , in to Equation (46), three equations are obtained but with six unknown parameters, at this stage, the parameters cannot be estimated. From Equation (46) $\tilde{A}_i = A_i e^{(-\alpha_i \mu_i^2 \epsilon^2)}$ for $i = 1, 2, 3$.

In the same way for the odd part of the finite Mixture Model function is estimated, Equation (51), since the even by is similar to the odd part, Equation (28). The methodology of estimation are similar.

$$\frac{f_o(x)}{x} = \sum_{i=1}^3 \tilde{B}_i e^{-\alpha_i x^2} + O(\epsilon^2 x^2) + O(\epsilon^2), \quad (56)$$

$$\text{where } \tilde{B}_i = 2A_i \epsilon \mu_i \alpha_i e^{-\alpha_i \epsilon^2 (\mu_i)^2}.$$

The even and odd components can be compared using Equation (49 & 57). Let $y = x^2$ then $f(x) = f(\sqrt{y})$ and, $\tilde{f}(y) = f(\sqrt{y})$.

$$\tilde{f}(y) = \sum_{i=1}^3 \tilde{B}_i e^{-\alpha_i y} + O(\epsilon^2 y) + O(\epsilon^2). \tag{57}$$

Consider

$$Y_i = e^{-\alpha_i y}, \text{ for } i = 1, 2, 3. \tag{58}$$

Let $\tilde{f}(y) = \tilde{f}(Y)$ and substitute Equation (58) into Equation (57) to obtain

$$\tilde{f}(z) = \sum_{i=1}^3 \tilde{B}_i Y_i. \tag{59}$$

Considering the dataset for the mixture model $\{x_i, f(x_i)\}$, using the multiple OLS on the dataset $\{x_i, f(x_i)\}$ for the mixture model Equation (59) the parameter estimates of \tilde{B}_i for $i = 1, 2, 3$.

The odd function was divided by an x term so as to approximate it as an even function. In this way the equations from the odd function can be compared to those from the even component. The odd function has three equations with six unknowns. Combining the Even and odd functions Equation (50 & 51) estimating the general unknown parameters A_i, μ_i , of model (61), the sufficient parameters of the mixture model can be obtained. i.e A system of six Equations and six unknowns.

The estimated parameters \tilde{A}_i and $\tilde{\mu}_i$, for $i = 1, 2, 3$ can be used as the IGVs for the unknown parameters in the iteration methods for the Three Mixture Model.

5. A Monte Carlo Simulation Study of a Three Component Mixture Model

5.1. Three Component Mixture Model

The three component mixture model is represented by Equation (61).

$$f(x, \sigma_i, \mu_i, \lambda_i) = \sum_{i=1}^3 \lambda_i \frac{1}{\sigma_i \sqrt{2\pi}} \exp\left(\frac{-(x - \epsilon\mu_i)^2}{2(\sigma_i)^2}\right), \quad (60)$$

where $\sum_{i=1}^3 \lambda_i = 1$ and $\lambda_i \in [0, 1]$.

From Equation (60) σ_i is the standard deviation μ_i is the mean λ_i is the scaling factor.

$$f(x) = \sum_{i=1}^3 A_i e^{-\alpha_i(x - \epsilon\mu_i)^2}, \quad (61)$$

where $A_i = \frac{\lambda_i}{\sigma_i \sqrt{2\pi}}$.

A simulation of a three component mixture model Equation (61) with the following individual parameters as $\epsilon = 0.1$, $A_1 = 1$, $A_2 = 1.5$, $A_3 = 2$, $\alpha_1 = 3$, $\alpha_2 = 2.5$, $\alpha_3 = 3$, $\lambda_1 = 1.023327$, $\lambda_2 = 1.681497$, $\lambda_3 = 2.046653$, $\mu_1 = 0.1$, $\mu_2 = 1.3$, $\mu_3 = 1$, $\sigma_1 = 0.408248$, $\sigma_2 = 0.447214$, $\sigma_3 = 0.408248$. This is presented as

$$f(x) = e^{-3(x - (0.1)(0.1))^2} + 1.5 e^{-2.5(x - (0.1)(1.3))^2} + 2 e^{-3(x - (0.1)(1))^2}. \quad (62)$$

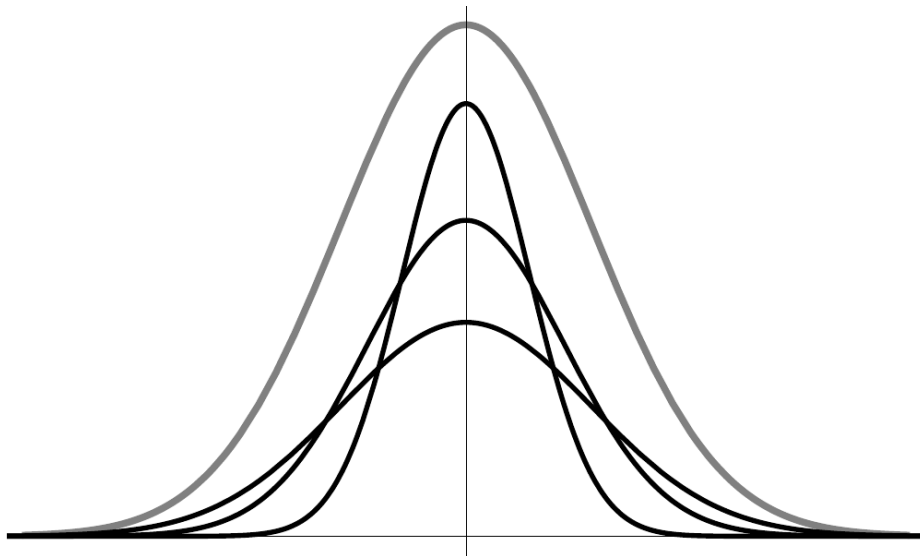


Figure 1: Three Component Mixture Model with small variance

According to Figure 1, the three component mixture model may appear as one Gaussian distribution. The main idea is to estimate the sufficient parameters that represent the three component mixture model by reducing the parameter dimensions of the original mixture model on subsequent estimation steps (see Section 4).

6. Results and Discussion

Tables 1 & 2 represent the result of the proposed method for a three component mixture model. In Table 1, the parameters of the three component mixture model are presented, and in Table 2, the sufficient parameters of the three component mixture model are presented, which are the main focus of the estimation problem. The exact values represent the Monte Carlo simulation of the model and the estimate values represent the values obtained using the proposed method.

Considering the results in Section 6, according to Table 1, the values of the exact simulated model of the three mixture model can be estimated using the proposed method (Theorem 1).

The sufficient parameters of the three mixture model can also be computed using the proposed method, according to Table 2, the sufficient parameters are accurately estimated with minimal errors. The proposed method thus gives the

Table 1: Parameter values for the Three Mixture Model

Parameter	Values	
	Exact	Estimated
ε	0.1	0.101581
A_1	1	0.987103
A_2	1.5	1.310132
A_3	2	1.871012
α_1	3	2.713215
α_2	2.5	2.468563
α_3	3	3.012349
λ_1	1.023327	1.062173
λ_2	1.681497	1.477978
λ_3	2.046653	1.910728

Table 2: Sufficient Parameter values for the Three Mixture Model

Parameter	Values	
	Exact	Estimated
μ_1	0.1	0.057918
μ_2	1.3	1.301013
μ_3	1	0.871463
σ_1	0.408248	0.429282
σ_2	0.447214	0.450052
σ_3	0.408248	0.407411

sufficient parameters for the mixture model that can be used to approximate the mixture model.

7. Conclusion

In this paper the problem of estimation of the finite mixture model was considered. The problem was solved using the proposed method that does not require

initialization or iterative methods for estimating the sufficient parameters of the FMM. It was also shown in the Monte Carlo simulation how the proposed method can be applied to a FMM and generalised for an n -component FMM. All simulations and computations were performed in Mathematica[®]. Considering the estimates, of the proposed method, it is reasonable to conclude that, the PM produced acceptable results. Nevertheless, it has to be noted that the results of the regular guess value method largely depend on the initial approximations. Hence the PM can be used as an algorithmic approach to computing initial approximations for the iteration methods.

8. Acknowledgments

The authors wish to thank the Department of Mathematics and Statistics, Faculty of Science of Tshwane University of Technology for funding the research.

References

- [1] G. J. McLachlan and D. Peel. *Finite mixture models*. Wiley Series in Probability and Statistics. WILEY & SONS., New York., 2000.
- [2] B. Van Dijk. *Essays on Finite Mixture Models*. Tinbergen Institute., 2009.
- [3] D. Kessler and A. McDowell. Introducing the fmm procedure for finite mixture models. 2012.
- [4] G. J. McLachlan and K. E. Basford. *Mixture models: Inference and applications to clustering*. MARCEL DEKKER, INC., New York, Basel, 1988.
- [5] T. R. Hastie, R. Tibshirani, and J. H. Friedman. *The elements of statistical learning*. SPRINGER, 2001.
- [6] A.K. Jain and R. Dubes. *Algorithms for Clustering Data*. PRENTICE HALL., Eaglewood Cliffs. New Jersey., 1988.
- [7] B.G. Lindsay. Mixture models: Theory, geometry and applications. *NSF-CBMS Regional Conference Series in Probability and Statistics*, 5:1–3, 5–9, 1–163., 1995.
- [8] A.K. Jain, R. Duin, and J. Mao. Statistical pattern recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(1):4–38, 2000.
- [9] Z. Liang, R.J. Jaszczak, and R.E. Coleman. Parameter estimation of finite mixtures using the em algorithm and information criteria with application to medical image processing, *iee transactions on nuclear science*, 1992. <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=159772>.
- [10] T. Hastie and R. Tibshirani. Discriminant analysis by gaussian mixtures. *Journal of the Royal Statistical Society (B)*, 58:155–176, 1996.
- [11] I.G.C. Filho. *Mixture Models for the Analysis of Gene Expression: Integration of Multiple Experiments and Cluster Validation*. University of Berlin., University of Berlin., 2007.

- [12] D. Titterington, A. Smith, and U. Makov. *Statistical Analysis of Finite Mixture Distributions*. JOHN WILEY & SONS., Chichester United Kingdom., 1985.
- [13] R. Duda, P. Hart, and D. Stork. *Pattern Classification*. JOHN WILEY & SONS, 2012.
- [14] C.R. Kikawa. *Methods to Solve Transcendental Least-Squares Problems and Their Statistical Inferences*. PhD thesis, Tshwane University of Technology, 2013.
- [15] C.R. Kikawa, M.Y. Shatalov, P.H. Kloppers, and A. Mkolesia. Parameter estimation for a mixture of two univariate gaussian distributions: A comparative analysis of the proposed and maximum likelihood methods. *British Journal of Mathematics and Computer Science*, 12(1):1–8, 2016.