Multivariate Logistic Mixtures

Xiao Liu

TUM School of Education, and Centre for International Student Assessment, Arcisstrasse 21, 80333 Munich, Germany

Copyright ©2015 by authors, all rights reserved. Authors agree that this article remains permanently open access under the terms of the Creative Commons Attribution License 4.0 International License

Abstract Logistic mixtures, unlike normal mixtures, have not been studied for their topography. In this paper we discuss analogs of some of the multivariate normal mixture results for the multivariate logistic distribution. We focus on graphical techniques that are based on displaying the elevation of the density on the ridgeline. These techniques are quite elementary, and carry full information about the location and relative heights of the modes and saddle points. Moreover, we turn to a technique that names II-Plot which denotes that the first differentiation of the second component density ratios the difference between the first differentiations of the second component density and the first component density.

Keywords Logistic Mixture, Multivariate Mode, Ridgeline, Pi-Plot

1 Introduction

There is the work by Ray and Lindsay ([6]) on the key features of multivariate normal mixtures, including the determination of the number of modes and general modality theorems. For the logistic distribution, on the other hand, such information seems to be lacking. The logistic distribution plays an important role in psychometrics for instance, for modeling item response functions (see [7]).

In this paper, we propose analogs of the multivariate normal mixture results for the multivariate logistic distribution (see [4]). The literature on determination of the number of modes in logistic mixture models has focused primarily on univariate mixtures. In fact, there is a simple description of modality when one is mixing two univariate components. Unlike the mixture of multivariate normal distributions (see [6]), for the logistic case it seems infeasible to express the ridgeline function explicitly. However, applying the implicit function theorem, we can prove that a unique explicit formula is possible locally. Moreover, we focus on displaying the elevation of the logistic mixture density on the ridgeline and address a technique called the Π -plot, both of which carry important information about modality properties of the mixture.

We conclude with remarks about the similarities and differences between the multivariate normal and logistic distributions in regards to their mixture properties and conclusions thereof.

2 The Ridgeline Manifold

2.1 The $(K - 1)$ -dimensional ridgeline manifold

A K -component mixture of D -dimensional logistic distributions can be represented by the probability density function

$$g(\mathbf{x}) = \sum_{i=1}^K \pi_i \phi(\mathbf{x}; \boldsymbol{\mu}_i, \mathbf{s}_i), \quad \mathbf{x} \in \mathbb{R}^D \quad (2.1)$$

where π_i is the mixing proportion of component i , $\pi_i \in [0, 1]$, $\sum_{i=1}^K \pi_i = 1$, and $\phi(\mathbf{x}; \boldsymbol{\mu}, \mathbf{s})$ is the density of a multivariate logistic distribution with mean $\boldsymbol{\mu}$ and standard deviation $\mathbf{s}$. We will sometimes use $\phi_i(\mathbf{x})$ as shorthand notation for $\phi(\mathbf{x}; \boldsymbol{\mu}_i, \mathbf{s}_i)$, and call ϕ_i the i th component density, where (following [4])

$$\phi_i(\mathbf{x}) = \frac{D!}{\prod_{j=1}^D s_{ij}} \exp \left[- \sum_{j=1}^D (x_j - \mu_{ij})/s_{ij} \right] \left\{ 1 + \sum_{j=1}^D \exp \left[- (x_j - \mu_{ij})/s_{ij} \right] \right\}^{-(D+1)}. \quad (2.2)$$

Definition 1. The $(K - 1)$ -dimensional set of points

$$\mathfrak{s}_K = \left\{ \boldsymbol{\alpha} \in \mathbb{R}^K : \alpha_i \in [0, 1], \sum_{i=1}^K \alpha_i = 1 \right\} \quad (2.3)$$

will be called the unit simplex. The function $\mathbf{x}^*(\boldsymbol{\alpha})$ from $\mathfrak{s}_K$ into $\mathbb{R}^D$ will be called the ridgeline function, which satisfies the following condition

$$\sum_{i=1}^K \alpha_i \sum_{k=1}^D s_{ik}^{-1} \left(1 - \frac{D+1}{E_{ik}} \right) = 0 \quad (2.4)$$

where $E_{ik} = 1 + e^{\frac{x_k - \mu_{ik}}{s_{ik}}} + e^{\frac{x_k - \mu_{ik}}{s_{ik}}} \sum_{j \neq k}^D e^{-\frac{x_j - \mu_{ij}}{s_{ij}}}$.

The image of this map will be denoted by $\mathfrak{M}$ and called the ridgeline surface or manifold. According to Lemma 2 (see Appendix A), we can get that Definition 1 is well-defined.

Theorem 1. Let $g(\mathbf{x})$ be the density of a K -component multivariate logistic density as given by formula (2.1). Then all of $g(\mathbf{x})$'s critical values, i.e. the values of $\mathbf{x}$ such that $\nabla g(\mathbf{x}) = 0$, are points in $\mathfrak{M}$.

Proof. Suppose that $\nabla g(\mathbf{x}^*) = 0$, so $\mathbf{x}^*$ is a critical point.

Then we have

$$\sum_{i=1}^K \pi_i \phi_i(\mathbf{x}^*) \frac{\nabla \phi_i(\mathbf{x}^*)}{\phi_i(\mathbf{x}^*)} = 0. \quad (2.5)$$

If we let

$$\alpha_i = \frac{\pi_i \phi_i(\mathbf{x}^*)}{\sum_{i=1}^K \pi_i \phi_i(\mathbf{x}^*)}, \quad (2.6)$$

then obviously, $0 \leq \alpha_i \leq 1$ and $\sum_{i=1}^K \alpha_i = 1$.

Further, note that

$$\frac{\nabla \phi_i(\mathbf{x}^*; \boldsymbol{\mu}, \mathbf{s})}{\phi_i(\mathbf{x}^*; \boldsymbol{\mu}, \mathbf{s})} = - \sum_{k=1}^D s_{ik}^{-1} \left(1 - \frac{D+1}{E_{ik}} \right). \quad (2.7)$$

Thus from equation (2.4), we have that for every critical value $\mathbf{x}^*$ there exists an $\boldsymbol{\alpha}$ such that

$$\sum_{i=1}^K \alpha_i \sum_{k=1}^D s_{ik}^{-1} \left(1 - \frac{D+1}{E_{ik}} \right) = 0.$$

The above formula gives the theorem. □

Remark 1. Due to the Taylor expansion (see Appendix B) until first order at $a = -1$ from the right side of formula (2.2), omitting the remainder, we can get that

$$\phi_i(\mathbf{y}) = \frac{D!}{\prod_{j=1}^D s_{ij}} \cdot \frac{\prod_{j=1}^D (-ey_j)}{(1 - \sum_{j=1}^D ey_j)^{D+1}}. \quad (2.8)$$

where $y_j = (x_j - \mu_{ij})/s_{ij}$.

Following formulae (2.5) and (2.6), we can get

$$\sum_{i=1}^K \alpha_i \frac{\nabla \phi_i(\mathbf{x}^*)}{\phi_i(\mathbf{x}^*)} = 0. \quad (2.9)$$

When $D = 2$, according to

$$\frac{\nabla \phi(\mathbf{y})}{\phi(\mathbf{y})} = \frac{y_1 + y_2 + 6e[1 - e(y_1 + y_2)]^2}{[1 - e(y_1 + y_2)]y_1y_2},$$

we obtain

$$\sum_{i=1}^K \alpha_i \left\{ y_1 + y_2 + 6e[1 - e(y_1 + y_2)]^2 \right\} = 0 \quad (2.10)$$

Corollary 1. If $s_{ik} = s_{i\cdot}$, let $D_i = \frac{\prod_{j=1}^K E_j}{E_i}$, and

$$F_i = \sum_{k'=1}^D F_{ik'} = \sum_{k'=1}^D \frac{\prod_{k=1}^D E_{ik}}{E_{ik'}},$$

then the convex hull of $(D+1)D_i F_i$ contains the value of $\prod_{i=1}^K E_i$ at all critical points of the density g , where $E_i = \prod_{k=1}^D E_{ik}$.

Proof. Following formula (2.4), we can get

$$\sum_{i=1}^K \alpha_i s_{i\cdot}^{-1} = \sum_{i=1}^K \alpha_i s_{i\cdot}^{-1} \sum_{k=1}^D \frac{D+1}{E_{ik}} = \sum_{i=1}^K \alpha_i s_{i\cdot}^{-1} \frac{(D+1)F_i}{E_i} = \frac{(D+1) \sum_{i=1}^K \alpha_i s_{i\cdot}^{-1} D_i F_i}{\prod_{i=1}^K E_i}.$$

Then

$$\prod_{i=1}^K E_i = \frac{(D+1) \sum_{i=1}^K \alpha_i s_{i\cdot}^{-1} D_i F_i}{\sum_{i=1}^K \alpha_i s_{i\cdot}^{-1}}. \quad (2.11)$$

Thus the above formula gives the corollary. $\square$

2.2 The ridgeline elevation plot

The next step in our analysis is to consider the diagnostic properties of the *elevation plot*, which is a plot of the *ridgeline elevation function* defined by

$$h(\alpha) = g(\mathbf{x}^*(\alpha)).$$

Remark 2. The positive and bounded density function $\phi_i(\mathbf{x})$ characterized by the parameters $(\mu_i, \mathbf{s}_i)$ in its second-order Taylor expansion depends on $\mathbf{x} \in \mathbb{R}^D$ as a decreasing function of the Mahalanobis distance $(\mathbf{x} - \mu_i)^T \text{diag}(s_{i1}^2, \dots, s_{iD}^2)^{-1} (\mathbf{x} - \mu_i)$. This ties in with Remark 4 of Ray and Lindsay (2005), where these authors generalize the theorem to other families of multivariate distributions.

Due to second-order Taylor series approximation (see Appendix B), l_{ϕ_i} , of ϕ_i about $a = -1$, omitting the remainder,

$$\begin{aligned} l_{\phi_i}(\mathbf{y}) &= \frac{D!}{\prod_{j=1}^D s_{ij}} \cdot \frac{\left(\frac{e}{2}\right)^D \prod_{j=1}^D (y_j^2 + 1)}{\left[1 + \frac{e}{2} \sum_{j=1}^D (y_j^2 + 1)\right]^{D+1}} \\ &\leq \frac{D!}{\prod_{j=1}^D s_{ij}} \cdot \frac{\left(\frac{e}{2}\right)^D (\max_j (y_j^2 + 1))^D}{\left(\frac{e}{2}\right)^D (\max_j (y_j^2 + 1))^D} \cdot \frac{1}{1 + \frac{e}{2} \sum_{j=1}^D (y_j^2 + 1)} \\ &= \frac{D!}{\left(\prod_{j=1}^D s_{ij}\right) \left[1 + \frac{eD}{2} + \frac{e}{2} \sum_{j=1}^D y_j^2\right]}, \end{aligned}$$

where $y_j = (x_j - \mu_{ij})/s_{ij}$.

Remark 3. The density function of t distribution with ν degrees of freedom

$$f(\mathbf{x}; \mu, \Sigma, \nu) = \frac{\Gamma\left(\frac{\nu+D}{2}\right) |\Sigma|^{-1/2}}{(\pi\nu)^{\frac{D}{2}} \Gamma\left(\frac{\nu}{2}\right) [1 + (\mathbf{x} - \mu)' \Sigma^{-1} (\mathbf{x} - \mu) / \nu]^{\frac{1}{2}(\nu+D)}} \quad (2.12)$$

where

$$\Gamma\left(\frac{\nu}{2}\right) := \int_0^\infty x^{\frac{\nu}{2}-1} e^{-x} dx,$$

which follows [5], characterized by parameters (μ, Σ) depends on $\mathbf{x} \in \mathbb{R}^D$ as a decreasing function of the Mahalanobis distance $(\mathbf{x} - \mu)' \Sigma^{-1} (\mathbf{x} - \mu)$, then we can get the result for t distribution from Remark 4 of [6].

Example 1. Consider the logistic mixture with $D = 2$ and $K = 2$, and with the parameters

$$\begin{aligned}\boldsymbol{\mu}_1 &= \begin{pmatrix} 0 \\ 0 \end{pmatrix}, & \mathbf{s}_1 &= \begin{pmatrix} 1 \\ 0.5 \end{pmatrix}, \\ \boldsymbol{\mu}_2 &= \begin{pmatrix} 1 \\ 1 \end{pmatrix}, & \mathbf{s}_2 &= \begin{pmatrix} 1 \\ 0.5 \end{pmatrix}, & \pi_1 &= \pi_2 = \frac{1}{2}.\end{aligned}$$

The corresponding ridgeline elevation function $h(\boldsymbol{\alpha})$ is shown in Figure 1. Following formula (2.4), we know

$$G_{\boldsymbol{\alpha}}(x_1, x_2) := \alpha_1 \left[\left(1 - \frac{3}{E_{11}}\right) + 2 \left(1 - \frac{3}{E_{12}}\right) \right] + \alpha_2 \left[\left(1 - \frac{3}{E_{21}}\right) + 2 \left(1 - \frac{3}{E_{22}}\right) \right] = 0$$

where

$$\begin{aligned}E_{11} &= 1 + e^{x_1} + e^{x_1-2x_2}, & E_{12} &= 1 + e^{2x_2} + e^{2x_2-x_1}; \\ E_{21} &= 1 + e^{x_1-1} + e^{x_1-2x_2+1}, & E_{22} &= 1 + e^{2(x_2-1)} + e^{2x_2-x_1-1}.\end{aligned}$$

Definition 2. Let

$$\mathbf{d}_i = S_{\boldsymbol{\alpha}}^{-1}(\mathbf{v}_i - \mathbf{v}_K),$$

where for fixed $\boldsymbol{\alpha}$ defined the matrix $S_{\boldsymbol{\alpha}} = \sum \alpha_i s_i^{-1}$ and vectors $\mathbf{v}_i = s_i^{-1} \sum_k \left(1 - \frac{D+1}{E_{ik}}\right)$. We can define a linear subspace of vectors that are orthogonal to the surface's direction vectors $\mathbf{d}_i$ in an appropriate sense,

$$\mathbf{W} = \{\mathbf{w} : \mathbf{w}' S_{\boldsymbol{\alpha}} \mathbf{d}_i = 0 \forall i = 1, \dots, K-1\}. \quad (2.13)$$

Theorem 2. If $\mathbf{w} \in \mathbf{W}$, then along the path $\{\mathbf{x}(\boldsymbol{\alpha}) + \delta \mathbf{w} : \delta \in \mathbb{R}\}$ the function $g(\mathbf{x})$ takes its maximum value at $\delta = 0$.

Proof.

We know that the point $\mathbf{x}(\boldsymbol{\alpha})$ lies in one of the elliptical contours of the density ϕ_i . According to formula (2.7), at this point the gradient in $\mathbf{x}$ of the density $\phi_i(\mathbf{x})$ is proportional to $\mathbf{v}_i$, and so $\mathbf{v}_i$ is orthogonal to the contour. Thus, if we were to start at $\mathbf{x}(\boldsymbol{\alpha})$ and travel in any direction $\mathbf{w}$ orthogonal to $\mathbf{v}_i$, our path is in the support hyperplane to the elliptically shaped upper set $\{\mathbf{x} : \phi_i(\mathbf{x}) \geq \phi_i(\mathbf{x}^*(\boldsymbol{\alpha}))\}$. Using the fact that the ellipse of ϕ_i is convex, our path lies outside the ellipse, and so in the set $\{\mathbf{x} : \phi_i(\mathbf{x}) < \phi_i(\mathbf{x}(\boldsymbol{\alpha}))\}$, except for equality at $\mathbf{x} = \mathbf{x}^*(\boldsymbol{\alpha})$. This means, the point of $\mathbf{x}^*(\boldsymbol{\alpha})$ is a local maximum to $\phi_i(\mathbf{x})$ along any path orthogonal to $\mathbf{v}_i$.

Now, assume that $\mathbf{w} \in \mathbf{W}$. It follows from the form of $\mathbf{d}_i$ that

$$\mathbf{w}'(\mathbf{v}_i - \mathbf{v}_K) = 0. \quad (2.14)$$

However, from the fomula (2.4), we have $\sum \alpha_i \mathbf{v}_i = 0$, and so $\sum \alpha_i (\mathbf{v}_i - \mathbf{v}_K) = -\mathbf{v}_K$. Due to (2.13), we know that $\mathbf{w}' \mathbf{v}_K = 0$. Putting this together with (2.14) shows that $\mathbf{w}' \mathbf{v}_i = 0$ for $i = 1, \dots, K$. That is, $\mathbf{w}'$ is orthogonal to every $\mathbf{v}_i$, and hence, by the above paragraph, every component of the mixture density $g(\mathbf{x})$ is locally maximized along the given line $\{\mathbf{x}(\boldsymbol{\alpha}) + \delta \mathbf{w} : \delta \in \mathbb{R}\}$ at $\delta = 0$, and thus, so is $g(\mathbf{x})$. □

Corollary 2. If $D \geq K - 1$, then at a critical point of $h(\boldsymbol{\alpha})$ whose second derivative matrix has $K - 1$ negative eigenvalues the function $g(\mathbf{x})$ will have a critical point whose second derivative matrix has an additional $D - K + 1$ negative eigenvalues corresponding to the dimension of the orthogonal directions $\mathbf{w}$. Specially, for $D > K - 1$, so that the $h(\boldsymbol{\alpha})$ plot is a true dimension reduction, then $g(\mathbf{x})$ has no local minima, only saddlepoints and local maxima.

Proof.

The directional vectors, together with their orthogonal complement $\mathbf{W}$, span the space, and following Theorem 2, we know that the $\mathbf{W}$ vectors are all directions of local maximization. □

Figure 1. The first one shows that ridgeline elevation for the bivariate logistic mixture of Example 1 along the ridgeline path $\mathbf{x}^*(\alpha)$, expressed as a function of parameter α ; The local maximum representing the mode of the mixture density is visible near $\alpha = 0.5$. The second one shows that density contour plot for the two component mixture density of Example 1.

Figure 2. Ridgeline elevation for the bivariate logistic mixture of Example 2 along the ridgeline path $\mathbf{x}^*(\alpha)$, expressed as a function of parameter α . The local maximum representing the mode of the density is visible near $\alpha = 0.5$.

3 Some Illustrative Examples

Example 2. The mixture logistic density with $D = 3$ and $K = 2$, and the parameters

$$\begin{aligned} \mu_1 &= \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, & s_1 &= \begin{pmatrix} 1 \\ 1 \\ 0.05 \end{pmatrix}, \\ \mu_2 &= \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, & s_2 &= \begin{pmatrix} 1 \\ 1 \\ 0.05 \end{pmatrix}, & \pi_1 &= \pi_2 = \frac{1}{2}. \end{aligned}$$

Example 3. The mixture logistic density with $D = 2$ and $K = 3$, and the parameters

$$\begin{aligned} \mu_1 &= \begin{pmatrix} 0 \\ 0 \end{pmatrix}, & s_1 &= \begin{pmatrix} 1 \\ 0.05 \end{pmatrix}, \\ \mu_2 &= \begin{pmatrix} 1 \\ 1 \end{pmatrix}, & s_2 &= \begin{pmatrix} 0.05 \\ 1 \end{pmatrix}, \\ \mu_3 &= \begin{pmatrix} 2 \\ 2 \end{pmatrix}, & s_3 &= \begin{pmatrix} 1 \\ 0.05 \end{pmatrix}, & \pi_1 &= \pi_2 = \pi_3 = \frac{1}{3}. \end{aligned}$$

Figure 3 shows the contours of the density given in Example 3.

Remark 4. More expressive detail for the two-dimensional plots of this section could have been obtained by displaying the critical net of the density using, for example, the approximation techniques of Danovaro et al. [1]. This would show the maxima, saddlepoints and separatrices over the manifold region based on evaluating the elevation at a finite network of points.

Proof. First to prove that the logistic density is a Morse function. When $D = 2$, following formula (2.2),

$$\phi = \frac{2}{s_1 s_2} e^{-\frac{x_1 - \mu_1}{s_1} - \frac{x_2 - \mu_2}{s_2}} \left(1 + e^{-\frac{x_1 - \mu_1}{s_1}} + e^{-\frac{x_2 - \mu_2}{s_2}} \right)^{-3}.$$

Figure 3. Density contour plot for the three component mixture density of Example 3.

After calculation, we know that

$$\begin{aligned} \frac{\partial^2 \phi}{\partial x_1^2} &= \frac{2}{s_1^3 s_2} e^{-\frac{x_1 - \mu_1}{s_1} - \frac{x_2 - \mu_2}{s_2}} \left(1 + e^{-\frac{x_1 - \mu_1}{s_1}} + e^{-\frac{x_2 - \mu_2}{s_2}} \right)^{-3} \\ &\quad - \frac{18}{s_1^3 s_2} e^{-\frac{x_1 - \mu_1}{s_1} - \frac{x_2 - \mu_2}{s_2}} \left(1 + e^{-\frac{x_1 - \mu_1}{s_1}} + e^{-\frac{x_2 - \mu_2}{s_2}} \right)^{-4} e^{-\frac{x_1 - \mu_1}{s_1}} \\ &\quad + \frac{24}{s_1^3 s_2} e^{-\frac{x_1 - \mu_1}{s_1} - \frac{x_2 - \mu_2}{s_2}} \left(1 + e^{-\frac{x_1 - \mu_1}{s_1}} + e^{-\frac{x_2 - \mu_2}{s_2}} \right)^{-5} e^{-2\frac{x_1 - \mu_1}{s_1}}, \end{aligned} \quad (3.1)$$

$$\begin{aligned} \frac{\partial^2 \phi}{\partial x_1 \partial x_2} &= \frac{2}{s_1^2 s_2^2} e^{-\frac{x_1 - \mu_1}{s_1} - \frac{x_2 - \mu_2}{s_2}} \left(1 + e^{-\frac{x_1 - \mu_1}{s_1}} + e^{-\frac{x_2 - \mu_2}{s_2}} \right)^{-3} \\ &\quad - \frac{6}{s_1^2 s_2^2} e^{-\frac{x_1 - \mu_1}{s_1} - \frac{x_2 - \mu_2}{s_2}} \left(1 + e^{-\frac{x_1 - \mu_1}{s_1}} + e^{-\frac{x_2 - \mu_2}{s_2}} \right)^{-4} e^{-\frac{x_1 - \mu_1}{s_1}} \\ &\quad - \frac{6}{s_1^2 s_2^2} e^{-\frac{x_1 - \mu_1}{s_1} - \frac{x_2 - \mu_2}{s_2}} \left(1 + e^{-\frac{x_1 - \mu_1}{s_1}} + e^{-\frac{x_2 - \mu_2}{s_2}} \right)^{-4} e^{-\frac{x_2 - \mu_2}{s_2}} \\ &\quad + \frac{24}{s_1^2 s_2^2} e^{-\frac{x_1 - \mu_1}{s_1} - \frac{x_2 - \mu_2}{s_2}} \left(1 + e^{-\frac{x_1 - \mu_1}{s_1}} + e^{-\frac{x_2 - \mu_2}{s_2}} \right)^{-5} e^{-\frac{x_1 - \mu_1}{s_1} - \frac{x_2 - \mu_2}{s_2}}, \end{aligned} \quad (3.2)$$

$$\begin{aligned} \frac{\partial^2 \phi}{\partial x_2^2} &= \frac{2}{s_1 s_2^3} e^{-\frac{x_1 - \mu_1}{s_1} - \frac{x_2 - \mu_2}{s_2}} \left(1 + e^{-\frac{x_1 - \mu_1}{s_1}} + e^{-\frac{x_2 - \mu_2}{s_2}} \right)^{-3} \\ &\quad - \frac{18}{s_1 s_2^3} e^{-\frac{x_1 - \mu_1}{s_1} - \frac{x_2 - \mu_2}{s_2}} \left(1 + e^{-\frac{x_1 - \mu_1}{s_1}} + e^{-\frac{x_2 - \mu_2}{s_2}} \right)^{-4} e^{-\frac{x_2 - \mu_2}{s_2}} \\ &\quad + \frac{24}{s_1 s_2^3} e^{-\frac{x_1 - \mu_1}{s_1} - \frac{x_2 - \mu_2}{s_2}} \left(1 + e^{-\frac{x_1 - \mu_1}{s_1}} + e^{-\frac{x_2 - \mu_2}{s_2}} \right)^{-5} e^{-2\frac{x_2 - \mu_2}{s_2}}. \end{aligned} \quad (3.3)$$

Then the Hessian matrix is non-degenerate (invertible) at critical points.

Thus, following [1], we get that this remark is given. $\square$

Until this point we have focused on a terrain morphological technique that is based on Morse Theory. This technique is quite elementary, and carry full information about the location and relative heights of the modes and saddlepoints. Someone may be also interested in another technique that focuses on the Multi-Triangulation in [1] corresponding to a refinement modification and a coarsening modification. By doing so, we can gain compact data structures for encoding an Multi-Triangulation, and algorithm for extracting Triangulated Irregular Networks according to user-defined resolution requirements, following [2].

Figure 4. Π -plot of Example 1.

4 The Π -PLOT

For $K = 2$, considering the ridgeline curve $\mathbf{x}^*(\alpha)$ defined in condition (2.4). When $\mathbf{x}^*(\alpha)$ is a critical value of $h(\alpha)$, it satisfies

$$h'(\alpha) = \pi \phi_1'(\mathbf{x}^*(\alpha))' + (1 - \pi) \phi_2'(\mathbf{x}^*(\alpha))' = 0,$$

where $'$ denotes differentiation with respect to α .

Solving the above displayed equation for π , and turning it into a function of α , we get

$$\Pi(\alpha) = \frac{\phi_2'(\alpha)}{\phi_2'(\alpha) - \phi_1'(\alpha)}. \quad (4.1)$$

We can also derive the following simple calculation formula for $\Pi(\alpha)$:

$$\frac{1}{\Pi(\alpha)} = \frac{\phi_2'(\alpha) - \phi_1'(\alpha)}{\phi_2'(\alpha)} = 1 + \frac{1 - \alpha}{\alpha} \frac{\phi_1(\alpha)}{\phi_2(\alpha)}, \quad (4.2)$$

which can be verified by formulae (2.5) and (2.6).

As an example, let us examine the Π -plot (graph of $\Pi(\alpha)$) of the two-component bivariate logistic mixture with two modes given in Example 1. As the mixing proportion in Example 1 is $\pi = 0.5$, we would draw a horizontal line across the $(\alpha, \Pi(\alpha))$ plot (Figure 4) at height $\pi = 0.5$. This line crosses the curve once. Among these, near $\alpha = 0.5$ correspond to the one mode, as was verified by the ridgeline elevation plot (see Figure 1).

For Example 2, similar with Example 1 we can find (see Figure 5), at height $\pi = 0.5$, near $\alpha = 0.5$ corresponds to one mode, as was verified by the ridgeline elevation plot (see Figure 2).

5 Analytic tools for detecting modality

5.1 The curvature function

In this section, we look more deeply into the properties of the Π function.

Considering back to the above section, the first differentiation of Π function with respect to α is

$$\Pi'(\alpha) = -\frac{\phi_2''(\alpha)\phi_1'(\alpha) - \phi_1''(\alpha)\phi_2'(\alpha)}{(\phi_2'(\alpha) - \phi_1'(\alpha))^2}. \quad (5.1)$$

In general, we can use any function of α with the same numerator $\phi_2''(\alpha)\phi_1'(\alpha) - \phi_1''(\alpha)\phi_2'(\alpha)$ to determine the sign changes of Π' because the denominator is a positive function of α .

Figure 5. Π -plot of Example 2.

As following we will use the curvature function $\kappa(\alpha)$ defined by

$$\kappa(\alpha) = \frac{\phi_2''(\alpha)\phi_1'(\alpha) - \phi_1''(\alpha)\phi_2'(\alpha)}{\phi_2(\alpha)\phi_1(\alpha)}. \quad (5.2)$$

5.2 Properties of the curvature function $\kappa(\alpha)$

Now we turn to give two properties of $\kappa(\alpha)$.

Lemma 1. (Two components, two dimensions, equal variance). In the equal variance case ($s_1 = s_2$), $\kappa(\alpha)$ reduces to the following expression

$$q(\mathbf{x}) = \phi_2''(\mathbf{x})\phi_1'(\mathbf{x}) - \phi_1''(\mathbf{x})\phi_2'(\mathbf{x}). \quad (5.3)$$

The logistic mixture g will be bimodal if and only if $\pi \in (\pi_1, \pi_2)$, where

$$\frac{1}{\pi_i} = 1 + \frac{1 - \alpha_i}{\alpha_i} \frac{\phi_1(\alpha_i)}{\phi_2(\alpha_i)} \quad (5.4)$$

and the α_i are the solutions in $[0, 1]$ of $q(\mathbf{x}) = 0$.

Proof. According to that the case is in two dimensions,

$$\Pi'(\alpha) = 0$$

$$\Rightarrow \phi_2''(\alpha)\phi_1'(\alpha) - \phi_1''(\alpha)\phi_2'(\alpha) = 0 \quad (5.5)$$

where

$$\phi'(\alpha) = \left(\frac{\partial \phi}{\partial x_1} + \frac{\partial \phi}{\partial x_2} \right) \dot{\mathbf{x}}(\alpha);$$

and

$$\phi''(\alpha) = \left(\frac{\partial^2 \phi}{\partial x_1^2} + 2 \frac{\partial^2 \phi}{\partial x_1 \partial x_2} + \frac{\partial^2 \phi}{\partial x_2^2} \right) \dot{\mathbf{x}}(\alpha) + \left(\frac{\partial \phi}{\partial x_1} + \frac{\partial \phi}{\partial x_2} \right) \ddot{\mathbf{x}}(\alpha).$$

Then

$$\begin{aligned}
& \phi_2''(\alpha)\phi_1'(\alpha) - \phi_1''(\alpha)\phi_2'(\alpha) \\
&= \left(\frac{\partial\phi_1}{\partial x_1} + \frac{\partial\phi_1}{\partial x_2} \right) \dot{\mathbf{x}}(\alpha) \left[\left(\frac{\partial^2\phi_2}{\partial x_1^2} + 2\frac{\partial^2\phi_2}{\partial x_1\partial x_2} + \frac{\partial^2\phi_2}{\partial x_2^2} \right) \dot{\mathbf{x}}(\alpha) + \left(\frac{\partial\phi_2}{\partial x_1} + \frac{\partial\phi_2}{\partial x_2} \right) \ddot{\mathbf{x}}(\alpha) \right] \\
&\quad - \left(\frac{\partial\phi_2}{\partial x_1} + \frac{\partial\phi_2}{\partial x_2} \right) \dot{\mathbf{x}}(\alpha) \left[\left(\frac{\partial^2\phi_1}{\partial x_1^2} + 2\frac{\partial^2\phi_1}{\partial x_1\partial x_2} + \frac{\partial^2\phi_1}{\partial x_2^2} \right) \dot{\mathbf{x}}(\alpha) + \left(\frac{\partial\phi_1}{\partial x_1} + \frac{\partial\phi_1}{\partial x_2} \right) \ddot{\mathbf{x}}(\alpha) \right] \\
&= \phi_2''(\mathbf{x})\phi_1'(\mathbf{x})\dot{\mathbf{x}}^2(\alpha) - \phi_1''(\mathbf{x})\phi_2'(\mathbf{x})\dot{\mathbf{x}}^2(\alpha)
\end{aligned} \tag{5.6}$$

Thus

$$\phi_2''(\mathbf{x})\phi_1'(\mathbf{x}) - \phi_1''(\mathbf{x})\phi_2'(\mathbf{x}) = 0. \tag{5.7}$$

□

Corollary 3. Let g be the mixture of two logistic densities with means μ_1 and μ_2 and standard deviation s_1 and $s_2 = \sigma s_1$. g is bimodal if and only if $\pi \in (\pi_1, \pi_2)$, where

$$\frac{1}{\pi_i} = 1 + \frac{1 - \alpha_i}{\alpha_i} \frac{\phi_1(\alpha_i)}{\phi_2(\alpha_i)}$$

and correspondingly the α_i are the solutions in $[0, 1]$ of

$$\phi_2''(\mathbf{x})\phi_1'(\mathbf{x}) - \phi_1''(\mathbf{x})\phi_2'(\mathbf{x}) = 0. \tag{5.8}$$

Specially, the one dimensional case of logistic mixture is in the preprint paper "A Note On The Convex Combination Of Two 1-, 2-, 3-, and 4-Parameter Logistic Item Response Functions", which is according to [8].

6 Conclusion

In this paper, we have proposed a technique for the topography of multivariate logistic mixtures. This is performed by the ridgeline function $\mathbf{x}^*(\alpha)$ and the Π -Plot. Unlike with the case of multivariate normal mixtures, it is difficult to express the ridgeline function $\mathbf{x}^*(\alpha)$ explicitly. However we can prove that we can get a unique explicit formula of $\mathbf{x}^*(\alpha)$ through the implicit function theorem. Thus we can obtain the Π -Plot even if we can not get the ridgeline contour plot when $K \geq 3$.

Future developments of the work described here consists of improving over the technique for displaying the contour plot when $K = 3$ following the Taylor expansion to take into account of the solution of ridgeline equation. The mathematical expression of the constructive implicit function theorem in logistic case is also very interesting. Otherwise, an application of this work can be used in PISA (Programme for International Student Assessment) analysis.

A Definition 1 is well-defined

Lemma 2. The left side of formula (2.4) is satisfied implicit function theorem.

Proof. Let

$$\psi(\mathbf{x}) := \sum_{i=1}^K \alpha_i \sum_{k=1}^D s_{ik}^{-1} \left(1 - \frac{D+1}{E_{ik}} \right)$$

for any fixed k , we can get

$$\begin{aligned}
\frac{\partial\psi}{\partial x_k} &= \sum_{i=1}^K \alpha_i \left(s_{ik}^{-1} \frac{D+1}{E_{ik}^2} \cdot \frac{dE_{ik}}{dx_k} + \sum_{k' \neq k}^D s_{ik'}^{-1} \frac{D+1}{E_{ik'}^2} \cdot \frac{dE_{ik'}}{dx_k} \right) \\
&= \sum_{i=1}^K \alpha_i \left[s_{ik}^{-2} e^{y_k} \frac{D+1}{(1 + e^{y_k} + e^{y_k} \sum_{j \neq k}^D e^{-y_j})^2} \left(1 + \sum_{j \neq k}^D e^{-y_j} \right) \right. \\
&\quad \left. - \sum_{k' \neq k}^D s_{ik'}^{-1} s_{ik}^{-1} \frac{D+1}{(1 + e^{y_{k'}} + e^{y_{k'}} \sum_{j \neq k'}^D e^{-y_j})^2} e^{y_{k'}} e^{-y_k} \right] \\
&= \sum_{i=1}^K \alpha_i (D+1) s_{ik}^{-1} \left[\frac{s_{ik}^{-1} e^{y_k}}{(1 + e^{y_k} + e^{y_k} \sum_{j \neq k}^D e^{-y_j})^2} \left(1 + \sum_{j \neq k}^D e^{-y_j} \right) \right. \\
&\quad \left. - \sum_{k' \neq k}^D \frac{s_{ik'}^{-1} e^{y_{k'}}}{(1 + e^{y_{k'}} + e^{y_{k'}} \sum_{j \neq k'}^D e^{-y_j})^2} e^{-y_k} \right].
\end{aligned} \tag{A.1}$$

According to the arbitrariness of μ_{ik} and $\mu_{ik'}$, we know that for $\forall x_k$ and $x_{k'}$, we can get $y_k = y_{k'}$. So formula (A.1) are not permanently equal to 0, for $k = 1, \dots, D$.

Following implicit function theorem (see page 89 of [3]), formula (2.4) can overlap locally with the graph of $\mathbf{x} = \mathbf{x}^*(\alpha)$, where $\mathbf{x}^*(\alpha)$ is an explicit function. $\square$

B Taylor Expansion

The Taylor expansion of a real-valued function $f(x)$ that is infinitely differentiable at a real number a is the power series

$$f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \frac{f^{(3)}(a)}{3!}(x-a)^3 + \dots$$

which can be written in the more compact sigma notation as

$$\sum_{n=0}^N \frac{f^{(n)}(a)}{n!}(x-a)^n + o((x-a)^N)$$

where $n!$ denotes the factorial of n , $o((x-a)^N)$ denotes that the remainder is higher order infinitesimal of $(x-a)^N$ and $f^{(n)}(a)$ denotes the n th derivative of f evaluated at the point a .

C Implicit function theorem

Let $f : \mathbb{R}^{n+m} \rightarrow \mathbb{R}^m$ be a continuously differentiable function, and let $\mathbb{R}^{n+m}$ have coordinates $(\mathbf{x}, \mathbf{y})$. Fix a point $(\mathbf{a}, \mathbf{b}) = (a_1, \dots, a_n, b_1, \dots, b_m)$ with $f(\mathbf{a}, \mathbf{b}) = \mathbf{c}$, where $\mathbf{c} \in \mathbb{R}^m$. If the matrix

$$\mathbf{Y} = \begin{pmatrix} \frac{\partial f_1}{\partial y_1}(\mathbf{a}, \mathbf{b}) & \dots & \frac{\partial f_1}{\partial y_m}(\mathbf{a}, \mathbf{b}) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial y_1}(\mathbf{a}, \mathbf{b}) & \dots & \frac{\partial f_m}{\partial y_m}(\mathbf{a}, \mathbf{b}) \end{pmatrix}$$

is invertible, then there exists an open set U containing $\mathbf{a}$, an open set V containing $\mathbf{b}$, and a unique continuously differentiable function $g : U \rightarrow V$ such that

$$\{(\mathbf{x}, g(\mathbf{x})) | \mathbf{x} \in U\} = \{(\mathbf{x}, \mathbf{y}) \in U \times V | f(\mathbf{x}, \mathbf{y}) = \mathbf{c}\}.$$

REFERENCES

- [1] E. Danovaro, L. De Floriani, P. Magillo, M. M. Mesmoudi and B. Abraham. Morphology-driven simplification and multiresolution modeling of terrains. *In Proc. Eleventh ACM International Symposium on Advances in Geographic Information Systems*, 63–70, ACM Press, New York, 2003.
- [2] L. De Floriani and P. Magillo. Multiresolution mesh representation: Models and data structures. *Tutorials on Multiresolution in Geometric Modelling*, 363–418, Springer-Verlag, 2002.
- [3] O. Forster. Analysis 2. *Vieweg Tuebner*, 8th. Version, 2008.
- [4] H. J. Malik and B. Abraham. Multivariate logistic distributions. *The Annals of Statistics*, 1(3):588–590, 1973.
- [5] D. Peel and G. J. McLachlan. Robust mixture modelling using the t Distribution. *Statistics and Computing*, 10:339–348, 2000.
- [6] S. Ray and B. G. Lindsay. The topography of multivariate normal mixtures. *The Annals of Statistics*, 33(5):2042–2065, 2005.
- [7] M. D. Reckase. Multidimensional Item Response Theory. *Springer*, New York, 2009.
- [8] C. A. Robertson and J. G. Fryer. Some descriptive properties of normal mixtures. *Skand. Aktuarietidskr.*, 52:137–146, 1969.