

The Constructive Implicit Function Theorem and Proof in Logistic Mixtures

Xiao Liu

Methods in Empirical Educational Research, TUM School of Education and Centre for International Student Assessment (ZIB),
TU München, Arcisstr. 21, 80333 Munich, Germany

Copyright ©2016 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 There is the work by Bridges et al (1999) on the key features of a constructive proof of the implicit function theorem, including some applications to physics and mechanics. For mixtures of logistic distributions such information is lacking, although a special instance of the implicit function theorem prevails therein. The theorem is needed to see that the ridgeline function, which carries information about the topography and critical points of a general logistic mixture problem, is well-defined [2]. In this paper, we express the implicit function theorem and related constructive techniques in their multivariate extension and propose analogs of Bridges and colleagues' results for the multivariate logistic mixture setting. In particular, the techniques such as the inverse of Lagrange's mean value theorem [4] allow to prove that the key concept of a logistic ridgeline function is well-defined in proper vicinities of its arguments.

Keywords Constructive Implicit Function Theorem, Logistic Mixture, Lagrange Mean Value Theorem, Ridgeline

1 Introduction

In this paper, we focus on constructive techniques that are based on expressing the proof of implicit function theorem in their multivariate extension and propose analogs of Bridges and colleagues' results for the multivariate logistic mixture setting.

According to [2], applying the implicit function theorem, we can prove that a unique explicit formula for the ridgeline function is possible locally in Theorem 1; In this paper, we propose analogs of Bridges et al.'s results (see [1]) for the multivariate logistic distribution.

Applying Lagrange's mean value theorem (see [4]), we can get the first lemma. Moreover, we are interested in uniform differentiability of 2 variables on Lemma 3 due to [5], which carries important information about the proof of our goal.

2 Applications to Topography

Theorem 1 later frequently allows us to show some ridgeline and contour plots, where the ridgeline function satisfies that the left side of formula (1) is equal to null (see [3]). The following example is in the case of two dimensions and three components.

Example 1. The mixture logistic density with $D = 2$ and $K = 3$, and the parameters

$$\begin{aligned}\mu_1 &= \begin{pmatrix} 0 \\ 0 \end{pmatrix}, & s_1 &= \begin{pmatrix} 1 \\ 0.07 \end{pmatrix}, \\ \mu_2 &= \begin{pmatrix} 1 \\ 1 \end{pmatrix}, & s_2 &= \begin{pmatrix} 0.07 \\ 1 \end{pmatrix}, \\ \mu_3 &= \begin{pmatrix} 2 \\ 2 \end{pmatrix}, & s_3 &= \begin{pmatrix} 1 \\ 0.07 \end{pmatrix}, & \pi_1 &= \pi_2 = \pi_3 = \frac{1}{3}.\end{aligned}$$

Figure 1 shows the contours of the density given in Example 1.

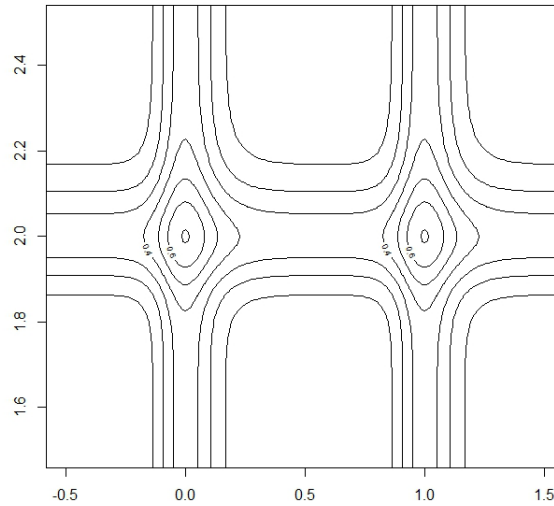


Figure 1. Density contour plot for the three component mixture density of Example 1.

3 The Constructive Implicit Function Theorem

For the most part we confine our attention to the following special case of the Implicit Function Theorem.

Theorem 1. *Let*

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

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

be a differentiable mapping of a neighbourhood of $(\alpha_0, \mathbf{x}_0) \in \mathbb{R}^K \times \mathbb{R}^D$ into \mathbb{R}^p , let $\psi(\alpha_0, \mathbf{x}_0) = 0$, and let $\det(\mathfrak{D}_2\psi(\alpha_0, \mathbf{x}_0)) \neq 0$. Then there exists $r > 0$ and a differentiable function $f : \bar{B}(\alpha_0, r) \subset \mathbb{R}^K \rightarrow \mathbb{R}^D$ such that for each $\alpha \in \bar{B}(\alpha_0, r)$, $(\alpha, f(\alpha))$ is the unique solution \mathbf{x} of the equation $\psi(\alpha, \mathbf{x}) = 0$ in some neighbourhood of (α_0, \mathbf{x}_0) .

Where we use standard modern notations for derivatives, such as \mathfrak{D} for the derivative itself, and \mathfrak{D}_l for the l th partial derivative ($l = 1, 2$), of a mapping from a subset of $\mathbb{R}^K \times \mathbb{R}^D$ to \mathbb{R}^p .

This result, given here for the logistic mixture case, has obvious generalizations to any general implicit function as following corollary. To keep the presentation short we refrain from presenting them here, as their treatment follows the same strategy analogously to [1].

Corollary 1. *Let ψ be a differentiable mapping of a neighbourhood of $(x_0, y_0) \in \mathbb{R}^K \times \mathbb{R}^D$ into \mathbb{R}^p , let $\psi(x_0, y_0) = 0$, and let $\det(\mathfrak{D}_2\psi(x_0, y_0)) \neq 0$. Then there exists $r > 0$ and a differentiable function $f : \bar{B}(x_0, r) \subset \mathbb{R}^K \rightarrow \mathbb{R}^D$ such that for each $x \in \bar{B}(x_0, r)$, $(x, f(x))$ is the unique solution y of the equation $\psi(x, y) = 0$ in some neighbourhood of (x_0, y_0) .*

In next section, we will give the proof of Theorem 1. At last, we will show some necessary lemmata for preparing the proof of our main theorem in appendix.

4 A New Proof of The Constructive Implicit Function Theorem in The Case of Logistic Mixtures

4.1 Proof of Theorem 1

Proof.

(i) Suppose that the assumption of Theorem 1 are satisfied, and choose $r, s > 0$ as in Lemma 2. Assume

$$J := B \times C = \{(\alpha, \mathbf{x}) : |\alpha - \alpha_0| \leq r, |\mathbf{x} - \mathbf{x}_0| \leq s\},$$

which is a compact set.

Fix ξ with $|\xi - \alpha_0| \leq r$, and there $\exists \varepsilon > 0$ such that

$$0 < \varepsilon < m \equiv \inf_{|\mathbf{h}| \leq r, |\mathbf{g}| \leq s} |\mathfrak{D}_2\psi(\alpha_0 + \mathbf{h}, \mathbf{x}_0 + \mathbf{g})|. \tag{2}$$

Consider \mathbf{x}, \mathbf{x}^* such that $|\mathbf{x} - \mathbf{x}_0| \leq s$, $|\mathbf{x}^* - \mathbf{x}_0| \leq s$, and $|\psi(\boldsymbol{\xi}, \mathbf{x}) - \psi(\boldsymbol{\xi}, \mathbf{x}^*)| < \varepsilon^2$.
If $|\mathbf{x} - \mathbf{x}^*| \geq \varepsilon$, then due to Lemma 1, $\exists \boldsymbol{\eta}$ between \mathbf{x} and \mathbf{x}^* such that

$$|\mathcal{D}_2\psi(\boldsymbol{\xi}, \boldsymbol{\eta})| \leq |\mathbf{x} - \mathbf{x}^*|^{-1}\varepsilon^2 \leq \varepsilon < m \leq |\mathcal{D}_2\psi(\boldsymbol{\xi}, \boldsymbol{\eta})|, \quad (3)$$

a contradiction.

Thus $|\mathbf{x} - \mathbf{x}^*| < \varepsilon$.

In particular, if $|\psi(\boldsymbol{\xi}, \mathbf{x})| < \varepsilon^2$ and $|\psi(\boldsymbol{\xi}, \mathbf{x}^*)| < \varepsilon^2$, then $|\mathbf{x} - \mathbf{x}^*| < 2\varepsilon$.

(ii) Next give a hypothesis of that

$$0 < \gamma := \inf_{|\mathbf{x} - \mathbf{x}_0| \leq s} |\psi(\boldsymbol{\xi}, \mathbf{x})|. \quad (4)$$

If $|\mathbf{x} - \mathbf{x}_0| \leq s$, then

$$|\mathcal{D}_2\psi^2(\boldsymbol{\xi}, \mathbf{x})| = 2|\psi(\boldsymbol{\xi}, \mathbf{x})||\mathcal{D}_2\psi(\boldsymbol{\xi}, \mathbf{x})| \geq 2\gamma m > 0. \quad (5)$$

According to the function $\mathcal{D}_2\psi^2(\boldsymbol{\xi}, \cdot)$ is continuous on the segment

$C := [\mathbf{x}_0 - \mathbf{m}, \mathbf{x}_0 + \mathbf{m}]$, without loss of generality, the Intermediate Value Theorem (see [4], Ch. 2) allows us to suppose that $\mathcal{D}_2\psi^2(\boldsymbol{\xi}, \mathbf{x}) \geq 2\gamma m$ for $\forall \mathbf{x} \in C$.

This implies that $\psi^2(\boldsymbol{\xi}, \cdot)$ is strictly increasing on C , and hence $\psi^2(\boldsymbol{\xi}, \mathbf{x}_0) > \psi^2(\boldsymbol{\xi}, \mathbf{x}_0 - \mathbf{m})$; this is contradictory due to (12) in view of the choice of s in Lemma 2.

Therefore $\gamma = 0$, then we can choose \mathbf{x}_n for each n such that $|\mathbf{x}_n - \mathbf{x}_0| \leq s$ and $|\psi(\boldsymbol{\xi}, \mathbf{x}_n)| < \frac{1}{n^2}$. Now the work in part (i)

of this proof shows that $\exists N$, such that when $n_j, n_l > N$, we can get $|\mathbf{x}_{n_j} - \mathbf{x}_{n_l}| < \frac{2}{n}$; then the sequence $\{\mathbf{x}_n\}$ is a Cauchy sequence, and hence converges to a limit \mathbf{x}_∞ on the segment $C = [\mathbf{x}_0 - \mathbf{m}, \mathbf{x}_0 + \mathbf{m}]$. The same argument shows that \mathbf{x}_∞ is also the unique solution \mathbf{x} on C of the equation $\psi(\boldsymbol{\xi}, \mathbf{x}) = 0$, thus we can define a function $f : [\boldsymbol{\alpha}_0 - \mathbf{n}, \boldsymbol{\alpha}_0 + \mathbf{n}] \rightarrow C$ due to $f(\boldsymbol{\xi}) = \mathbf{x}_\infty$.

On account of Lemma 3, we complete the proof. □

A Definition of Uniform Differentiability in 2 Variables

Definition 1. Let $f : \mathbb{R}^m \times \mathbb{R}^n \rightarrow \mathbb{R}$ be differentiable and such that ∇f is uniformly continuous. We define that f is uniformly differentiable, i.e., for any $\varepsilon > 0$, there is a $\delta > 0$ such that for all $a, x \in \mathbb{R}^m$ and $b, y \in \mathbb{R}^n$, we have

$$\frac{|f(a, b) - f(x, y) - \mathcal{D}_1f(x, y)(a - x) - \mathcal{D}_2f(x, y)(b - y)|}{\|(a, b) - (x, y)\|} < \varepsilon$$

whenever $\|(a, b) - (x, y)\| := \sqrt{\|a - x\|^2 + \|b - y\|^2} < \delta$.

B Three Lemmata

Although, logically unnecessary for the inverse of Lagrange's Mean Value Theorem, it is, for pedagogical reasons, interesting to see that we show the following lemma. Moreover, the result will be also useful for the proof of Theorem 1.

Lemma 1. For every function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ with continuous partial derivatives f_{x_1}, \dots, f_{x_n} and for all distinct pairs \mathbf{a} and \mathbf{b} in \mathbb{R}^n , there exists an intermediate point \mathbf{x} on the line segment joining \mathbf{a} and \mathbf{b} which we denote as $\mathbf{x} \in [\mathbf{a}, \mathbf{b}]$ such that

$$\inf_{\mathbf{x} \in [\mathbf{a}, \mathbf{b}]} |\mathcal{D}f(\mathbf{x})| \leq |\mathbf{b} - \mathbf{a}|^{-1} |f(\mathbf{b}) - f(\mathbf{a})|. \quad (6)$$

Proof. According to [4], we can get

$$f(\mathbf{b}) - f(\mathbf{a}) = \langle \mathbf{b} - \mathbf{a}, \mathcal{D}f(\mathbf{x}) \rangle \quad (7)$$

then

$$\left\langle \mathbf{b} - \mathbf{a}, \inf_{\mathbf{x} \in [\mathbf{a}, \mathbf{b}]} |\mathcal{D}f(\mathbf{x})| \right\rangle \leq |f(\mathbf{b}) - f(\mathbf{a})|. \quad (8)$$

Thus

$$\inf_{\mathbf{x} \in [\mathbf{a}, \mathbf{b}]} |\mathcal{D}f(\mathbf{x})| \leq |\mathbf{b} - \mathbf{a}|^{-1} |f(\mathbf{b}) - f(\mathbf{a})|. \quad (9)$$

□

The following example of Lemma 1 can be applied for calculating the average velocity of non-uniform motion in kinematics.

Example 2. Let

$$f(\mathbf{x}) = a\mathbf{x}^2 + b\mathbf{x} \quad (a \neq 0)$$

then

$$\mathfrak{D}f(\mathbf{x}) = 2a\mathbf{x} + b.$$

So we have

$$\begin{aligned} & a(\mathbf{a} + \mathbf{b} - \mathbf{a})^2 + b(\mathbf{a} + \mathbf{b} - \mathbf{a}) - a\mathbf{a}^2 - b\mathbf{a} \\ = & (\mathbf{b} - \mathbf{a}) \left\{ 2a\left[\mathbf{a} + \frac{1}{2}(\mathbf{b} - \mathbf{a})\right] + b \right\}. \end{aligned}$$

The core to our mathematical expression of the existence of an implicit function in the logistic mixtures case is provided by the following lemma.

Lemma 2. Under the hypotheses of Theorem 1, there exist $\mathbf{m} \in \mathbb{R}^D$, $|\mathbf{m}| = s$, $\mathbf{n} \in \mathbb{R}^K$, $|\mathbf{n}| = r$, and $r, s > 0$ such that

$$|\psi(\boldsymbol{\alpha}_0, \mathbf{x}_0 \pm \mathbf{m})| \geq \frac{2}{3}s|\mathfrak{D}_2\psi(\boldsymbol{\alpha}_0, \mathbf{x}_0)|, \tag{10}$$

for $\forall \mathbf{h} \in \mathbb{R}^K, \mathbf{g} \in \mathbb{R}^D$

$$\inf_{|\mathbf{h}| \leq r, |\mathbf{g}| \leq s} |\mathfrak{D}_2\psi(\boldsymbol{\alpha}_0 + \mathbf{h}, \mathbf{x}_0 + \mathbf{g})| > 0, \tag{11}$$

and

$$\inf_{|\mathbf{h}| \leq r} |\psi(\boldsymbol{\alpha}_0 + \mathbf{h}, \mathbf{x}_0 \pm \mathbf{m})| \geq \frac{1}{2}s|\mathfrak{D}_2\psi(\boldsymbol{\alpha}_0, \mathbf{x}_0)| > \sup_{|\mathbf{h}| \leq r} |\psi(\boldsymbol{\alpha}_0 + \mathbf{h}, \mathbf{x}_0)|. \tag{12}$$

Proof. Choose an open ball B , with centre $(\boldsymbol{\alpha}_0, \mathbf{x}_0)$ and radius R . Following formula (1), obviously $\mathfrak{D}_2\psi(\boldsymbol{\alpha}, \mathbf{x}) \in \mathcal{C}^0(B)$. Then we can get

$$|\mathfrak{D}_2\psi(\boldsymbol{\alpha}, \mathbf{x})| > \frac{1}{2}|\mathfrak{D}_2\psi(\boldsymbol{\alpha}_0, \mathbf{x}_0)| \tag{13}$$

for all $(\boldsymbol{\alpha}, \mathbf{x}) \in B$.

According to that ψ is differentiable at $(\boldsymbol{\alpha}_0, \mathbf{x}_0)$, there exists $s \in (0, R)$ such that if $|\mathbf{x} - \mathbf{x}_0| \leq s$, we obtain

$$\left| \frac{\psi(\boldsymbol{\alpha}_0, \mathbf{x}) - \psi(\boldsymbol{\alpha}_0, \mathbf{x}_0)}{\mathbf{x} - \mathbf{x}_0} \right| \leq \frac{4}{3}|\mathfrak{D}_2\psi(\boldsymbol{\alpha}_0, \mathbf{x}_0)|,$$

then

$$|\psi(\boldsymbol{\alpha}_0, \mathbf{x}) - \mathfrak{D}_2\psi(\boldsymbol{\alpha}_0, \mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)| \leq \frac{1}{3}|\mathfrak{D}_2\psi(\boldsymbol{\alpha}_0, \mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)| \tag{14}$$

and therefore

$$|\psi(\boldsymbol{\alpha}_0, \mathbf{x})| \geq \frac{2}{3}|\mathfrak{D}_2\psi(\boldsymbol{\alpha}_0, \mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)|. \tag{15}$$

In particular, choose $\mathbf{x} = \mathbf{x}_0 \pm \mathbf{m}$, we obtain inequality (10). Since $\psi(\boldsymbol{\alpha}_0, \mathbf{x}_0) = 0$ and ψ is continuous, we can now choose $r \in (0, R)$ such that inequality in (12) hold. According to $r < R$, our choice of R ensures that we can obtain inequality (11). \square

Next we separate out the proof of the differentiability of the implicit function. It will be convenient to establish the existence of Theorem 1 before.

Lemma 3.

Let B be a compact ball in \mathbb{R}^K , C a compact domain in \mathbb{R}^D , and $\psi : B \times C \rightarrow \mathbb{R}^P$ be a uniformly differentiable function such that

$$0 < m := \inf_{B \times C} |\mathfrak{D}_2\psi|. \tag{16}$$

Suppose that there exists a function $f : B \rightarrow C$ such that $\psi(\boldsymbol{\alpha}, \mathbf{x}) = 0$ for $\forall \boldsymbol{\alpha} \in B, \mathbf{x} := f(\boldsymbol{\alpha}) \in C$. Then f is uniformly differentiable on B , and

$$f'(\boldsymbol{\xi}) = -\frac{\mathfrak{D}_1\psi(\boldsymbol{\xi}, f(\boldsymbol{\xi}))}{\mathfrak{D}_2\psi(\boldsymbol{\xi}, f(\boldsymbol{\xi}))} \tag{17}$$

for any $\boldsymbol{\xi} \in B$.

Proof. Let $0 < \varepsilon < \frac{1}{7}m$, and let $\alpha^{(1)}, \alpha^{(2)}$ be points of B , we define

$$\|\alpha^{(1)} - \alpha^{(2)}\| := \sqrt{\sum_{i=1}^K (\alpha_i^{(1)} - \alpha_i^{(2)})^2}. \quad (18)$$

According to Definition 1, we have the definition of uniform differentiability in 2 variables (also see [5]), so we can get

$$\begin{aligned} & \left| \frac{\psi(\alpha^{(1)}, f(\alpha^{(1)})) - \psi(\alpha^{(2)}, f(\alpha^{(2)})) - \mathfrak{D}_1\psi(\alpha^{(2)}, f(\alpha^{(2)}))(\alpha^{(1)} - \alpha^{(2)})}{\sqrt{\|\alpha^{(1)} - \alpha^{(2)}\|^2 + \|f(\alpha^{(1)}) - f(\alpha^{(2)})\|^2}} \right| \\ & \leq \varepsilon. \end{aligned} \quad (19)$$

Then

$$\begin{aligned} & \left| \psi(\alpha^{(1)}, f(\alpha^{(1)})) - \psi(\alpha^{(2)}, f(\alpha^{(2)})) \right. \\ & \quad \left. - \mathfrak{D}_1\psi(\alpha^{(2)}, f(\alpha^{(2)}))(\alpha^{(1)} - \alpha^{(2)}) \right. \\ & \quad \left. - \mathfrak{D}_2\psi(\alpha^{(2)}, f(\alpha^{(2)}))(f(\alpha^{(1)}) - f(\alpha^{(2)})) \right| \\ & \leq \varepsilon \sqrt{\|\alpha^{(1)} - \alpha^{(2)}\|^2 + \|f(\alpha^{(1)}) - f(\alpha^{(2)})\|^2} \\ & \leq \varepsilon \left(\|\alpha^{(1)} - \alpha^{(2)}\| + \|f(\alpha^{(1)}) - f(\alpha^{(2)})\| \right). \end{aligned} \quad (20)$$

Otherwise $\psi(\alpha^{(1)}, f(\alpha^{(1)})) = \psi(\alpha^{(2)}, f(\alpha^{(2)})) = 0$ and $|\mathfrak{D}_2\psi(\alpha^{(2)}, f(\alpha^{(2)}))| \geq m$ due to formula (16), so

$$\begin{aligned} & \left| \frac{\mathfrak{D}_1\psi(\alpha^{(2)}, f(\alpha^{(2)}))}{\mathfrak{D}_2\psi(\alpha^{(2)}, f(\alpha^{(2)}))} (\alpha^{(1)} - \alpha^{(2)}) + f(\alpha^{(1)}) - f(\alpha^{(2)}) \right| \\ & \leq m^{-1}\varepsilon \left(\|\alpha^{(1)} - \alpha^{(2)}\| + \|f(\alpha^{(1)}) - f(\alpha^{(2)})\| \right) \\ & \leq \frac{1}{7} \left(\|\alpha^{(1)} - \alpha^{(2)}\| + \|f(\alpha^{(1)}) - f(\alpha^{(2)})\| \right). \end{aligned} \quad (21)$$

Therefore

$$\|f(\alpha^{(1)}) - f(\alpha^{(2)})\| \leq 7 \left| \frac{\mathfrak{D}_1\psi(\alpha^{(2)}, f(\alpha^{(2)}))}{\mathfrak{D}_2\psi(\alpha^{(2)}, f(\alpha^{(2)}))} (\alpha^{(1)} - \alpha^{(2)}) \right| + \|\alpha^{(1)} - \alpha^{(2)}\|.$$

Choosing a bound M for $|\mathfrak{D}_1\psi|$ on the compact set $B \times C$, we see that

$$\|f(\alpha^{(1)}) - f(\alpha^{(2)})\| \leq (7Mm^{-1} + 1) \|\alpha^{(1)} - \alpha^{(2)}\|. \quad (22)$$

It follows from formula (21) that

$$\begin{aligned} & \left| \frac{\mathfrak{D}_1\psi(\alpha^{(2)}, f(\alpha^{(2)}))}{\mathfrak{D}_2\psi(\alpha^{(2)}, f(\alpha^{(2)}))} (\alpha^{(1)} - \alpha^{(2)}) + f(\alpha^{(1)}) - f(\alpha^{(2)}) \right| \\ & \leq m^{-1} (7Mm^{-1} + 2) \varepsilon \|\alpha^{(1)} - \alpha^{(2)}\| \rightarrow 0 \end{aligned} \quad (23)$$

when $\varepsilon \rightarrow 0$.

Thus f is uniformly differentiable on B , with

$$f'(\xi) = -\frac{\mathfrak{D}_1\psi(\xi, f(\xi))}{\mathfrak{D}_2\psi(\xi, f(\xi))}$$

for $\forall \xi \in B$.

□

REFERENCES

- [1] D. Bridges, C. Calude, B. Pavlov, D. Stefanescu. The Constructive Implicit Function Theorem and Applications in Mechanics, Chaos Solitons and Fractals, Vol.10, 927-934, 1999.
- [2] X. Liu, A. Ünlü. Multivariate Logistic Mixtures, European Conference on Data Analysis (ECDA) , The University of Bremen, 105, 2014.
- [3] X. Liu. Multivariate Logistic Mixtures, Universal Journal of Applied Mathematics, Vol.3, No.4, 77-87, 2015.
- [4] P. K. Sahoo, T. Riedel. Mean Value Theorems and Functional Equations, World Scientific Press, New Jersey, 1998.
- [5] E. M. Stein. Singular Integrals and Differentiability Properties of Functions, Princeton University Press, New Jersey, 1970.