数理解析研究所講究録
第 1758 巻 2011 年 60-79

異なる球面対称分布の位置混合分布の principal points の性質について

青山学院大学・理工学部　松浦　峻 (Shun Matsuura)
College of Science and Engineering, Aoyama Gakuin University
東京大学大学院・総合文化研究科　倉田　博史 (Hiroshi Kurata)
Graduate School of Arts and Sciences, The University of Tokyo

概要

確率分布の n-principal points (主要点とも呼ばれる) とは，その確率分布に従う確率変数との平均 2 乗距離を最小にする n 個の点のことである．多次元確率分布の $n(\geq 3)$-principal points の理論的性質はほとんど知られておらず，principal points の探索や推定を困難にしている．本稿では，異なる球面対称分布の位置混合分布の principal points の性質について議論し，いくつかの条件のもと，n-principal points が各球面対称分布の位置ベクトルが張る線形部分空間上に存在することを示す．

キーワード　クラスター分析，主部分空間定理，主要点，線形部分空間，多変量混合分布．

1 はじめに

$\boldsymbol{X}$ を有限な 2 次モーメントを持つ p 次元確率変数ベクトルとする．このとき，$\boldsymbol{X}$ の n-principal points (主要点とも呼ばれる) とは，$\boldsymbol{X}$ との平均 2 乗距離 (mean squared distance) を最小にする n 個の点のことである．より正確には，

$$E[d^2(\boldsymbol{X}|\gamma_1,\ldots,\gamma_n)] \tag{1.1}$$

を最小にする R^p 上の n 点 $\{\gamma_1^*,\ldots,\gamma_n^*\}$ を $\boldsymbol{X}$ の n-principal points と呼ぶ (Flury (1990, Definition 2))．ただし，

$$d^2(\boldsymbol{x}|\gamma_1,\ldots,\gamma_n)=\min_{i=1,\ldots,n}\|\boldsymbol{x}-\gamma_i\|^2$$

である．$\boldsymbol{X}$ の n-principal points は $\boldsymbol{X}$ の 2 次モーメントが有限である限り，全ての自然数 n に対して存在することが知られている (Graf and Luschgy (2000, Theorem 4.12))．

$\boldsymbol{X}$ の 1-principal point は常に $E[\boldsymbol{X}]$ となる．また，$\boldsymbol{X}\sim N_p(\boldsymbol{0}_p,I_p)$ のとき，2-principal points は $\{\pm\boldsymbol{x}\in R^p|\|\boldsymbol{x}\|=\sqrt{2/\pi}\}$ で与えられる．

基本的な性質の一つとして，principal points は常に self-consistent points であることが知られている (Flury (1993, Lemma 1))．ただし，逆は成り立つとは限らない．$\boldsymbol{X}$ の n-self-consistent points とは，

$$E[\boldsymbol{X}|\boldsymbol{X}\in C_i]=\gamma_i,\ i=1,\ldots,n \tag{1.2}$$

を満たす R^p 上の n 点 $\{\gamma_1,\ldots,\gamma_n\}$ のことである．ただし，C_i, $i=1,\ldots,n$ は $\{\gamma_1,\ldots,\gamma_n\}$ から導かれるボロノイ領域：

$$C_i=\{\boldsymbol{x}\in R^p|\|\boldsymbol{x}-\boldsymbol{\gamma}_i\|<\|\boldsymbol{x}-\boldsymbol{\gamma}_j\|,\ j=1,\ldots,i-1,\ \|\boldsymbol{x}-\boldsymbol{\gamma}_i\|\leq\|\boldsymbol{x}-\boldsymbol{\gamma}_j\|,\\ j=i+1,\ldots,n\},\ i=1,\ldots,n$$

である．(1.2) 式より，

$$E[\boldsymbol{X}]=\sum_{i=1}^{n}E[\boldsymbol{X}|\boldsymbol{X}\in C_i]P(\boldsymbol{X}\in C_i)=\sum_{i=1}^{n}\gamma_iP(\boldsymbol{X}\in C_i)$$

となるから，$\boldsymbol{X}$ の self-consistent points および principal points の convex hull は必ず $E[\boldsymbol{X}]$ を含む（Tarpey, Li, and Flury (1995, Lemma 2.1)）．従って，$E[\boldsymbol{X}]=\mathbf{0}_p$ のとき，$\boldsymbol{X}$ の n-self-consistent points および n-principal points は必ず $\min\{n-1,p\}$ 以下の次元の線形部分空間上に存在する．また，principal points は確率分布に対して回転共変性および位置共変性を持つ．$\{\gamma_1^*,\ldots,\gamma_n^*\}$ が $\boldsymbol{X}$ の n-principal points であるとき，任意の $p\times p$ 直交行列 Γ および $p\times 1$ ベクトル $\boldsymbol{b}$ に対し，$\{\Gamma\gamma_1^*+\boldsymbol{b},\ldots,\Gamma\gamma_n^*+\boldsymbol{b}\}$ は $\Gamma\boldsymbol{X}+\boldsymbol{b}$ の n-principal points である（Tarpey, Li, and Flury (1995, Lemma 2.2)）．従って，本稿の議論では一般性を失うことなく $E[\boldsymbol{X}]=\mathbf{0}_p$ と仮定する．

principal points は確率分布の最適分割や離散分布への最適近似とみなすことができ，標本の最適分割である k-means 法によるクラスター分析と密接な関連がある．実際，確率分布からの無作為標本に k-means 法を適用することによって得られる k 個のクラスター平均は，いくつかの正則条件のもとで，その確率分布の k-principal points の推定量として強一致性を持ち，漸近的に正規分布に従うことが知られている（Pollard (1981,1982)）．その意味で，principal points の理論的性質を明らかにすることは k-means 法によるクラスター分析のふるまいに対する基礎的理論を提供することにつながると考えられる．

principal points や確率分布の最適分割の応用例としては，複数のマスクのサイズの決定問題（Flury (1993)），天気図の解析（村木・大瀧・水田 (1998)），2 種類の部品の選択的組立問題（Mease, Nair, and Sudjianto (2004), Mease and Nair (2006), Matsuura and Shinozaki (2007,2010), Matsuura (2011)）などが挙げられる．また，Tarpey and Petkova (2010) では principal points を用いた判別分析が提案され，医療データの解析に適用されている．近年では，定義を確率的に変動する関数に広げた関数主要点もよく議論されている（Tarpey and Kinateder (2003), 清水・水田 (2008), Shimizu and Mizuta (2008), Bali and Boente (2009) など）．

本稿では多次元確率分布の principal points の理論的側面について議論を行い，主部分空間定理と呼ばれる principal points が存在する範囲を陽に示す定理について先行研究および近年の著者らの論文（Matsuura and Kurata (2010,2011)）の結果を紹介し，Matsuura and Kurata (2011) の拡張を行う．次章で，principal points の理論的背景を紹介し，特に主部分空間定理に関する先行研究と Matsuura and Kurata (2010,2011) の結果を詳しく述べる．第 3 章で，Matsuura and Kurata (2011) の拡張を行う．具体的には，異なる球面対称分布の位置混合分布の principal points における主部分空間定理を導き，いくつかの例を与える．第 4 章では結論と今後の課題を述べる．

2 principal points の理論的背景

確率分布の principal points に関して理論的な観点から様々な議論がなされている．まず，基本的な問題の一つとして，既知の確率分布の principal points をどうやって求めるかという問題がある．principal points が self-consistent points でもあることを利用して，(1.2) 式を満たす n 点を繰り返しアルゴリズムを用いて求めるのが一般的であるが，self-consistent points は平均 2 乗距離 (1.1) の最小値を与える保証はなく，極小値，停留点や極大値を与えてしまうことがある．self-consistent points の一意性が成立するための条件（そのとき，self-consistent points は必ず principal points であることが保証される）や，それに関連して principal points の配置の対称性が成立する条件について，一次元確率分布の n-principal points または多次元確率分布の 2-principal points の場合には詳細な議論がなされている (Trushkin (1982,1984), Kieffer (1983), Tarpey (1994), Li and Flury (1995), Zoppe (1995,1997), 清水・水田・佐藤 (1998,1999), Yamamoto and Shinozaki (2000a,b), Gu and Mathew (2001), Mease and Nair (2006), Kurata and Qiu (2011) など)．例えば一次元確率分布の確率密度関数が log-concave であるときに n-principal points が一意に定まることはよく知られている（Trushkin (1982)）．

一方で多次元確率分布の $n(\geq 3)$-principal points に関しては，一意性の条件や principal points を求めるための明示的な式はほとんど得られていない．その結果，確率分布の次元数 p や n の値が大きい場合，p 次元空間上の n 点のあらゆる可能な配置の中から平均 2 乗距離 (1.1) を最小にするものを求めることは非常に膨大な計算量を必要とすることになる．従って，principal points が存在する範囲を陽に示すことはその計算量の大幅な減少につながる意味で重要である．

また，確率分布のパラメータが未知な場合の無作為標本からの principal points の推定問題について Pollard (1981,1982)，Flury (1993)，Tarpey (1997)，Stampfer and Stadlober (2002)，Tarpey (2007) などで様々な議論がなされてきており，Tarpey (2007) において多次元確率分布の k-principal points の推定量として

(1) ノンパラメトリック推定量：
標本に k-means 法を適用して得られる k 個のクラスター平均を用いる方法（前述したように，いくつかの正則条件のもと，強一致性を持ち，漸近的に正規分布に従う推定量である）

(2) 最尤推定量：
未知パラメータに最尤推定量をプラグインした確率分布の k-principal points を用いる方法

が議論されているが，後者の最尤推定量を求める際にも多次元確率分布の principal points の探索の困難さがネックとなっている．従って，多次元確率分布の principal points における主部分空間定理などの理論的結果を発展させることは principal points の推定問題にも貢献すると考えられる．

次節で，主部分空間定理について先行研究および Matsuura and Kurata (2010,2011) の

結果を紹介する.

2.1 多次元確率分布の principal points における主部分空間定理

Tarpey, Li, and Flury (1995) は，楕円対称分布（elliptically symmetric distribution）の n-principal points がその共分散行列の大きいほうの固有値に対応する固有ベクトルによって張られる線形部分空間上に存在することを示した．楕円対称分布とは特性関数 $\phi(\boldsymbol{t}) = E[\exp(i\boldsymbol{t}'\boldsymbol{X})]$ が

$$\phi(\boldsymbol{t}) = \exp(i\boldsymbol{t}'\boldsymbol{\mu})\psi(\boldsymbol{t}'\Psi\boldsymbol{t}) \quad \text{for some } \psi : [0,\infty) \to [0,\infty)$$

で表される確率分布のことであり，2 次モーメントが有限であるとき，期待値は $\boldsymbol{\mu}$，共分散行列は Ψ の正の定数倍で与えられる．楕円対称分布の簡単な例は多変量正規分布 $N_p(\boldsymbol{\mu}, \Sigma)$ である．

命題 1. (Tarpey, Li, and Flury (1995, Theorem 4.1))
$\boldsymbol{X}$ を p 次元楕円対称分布に従う確率変数ベクトルとし，$E[\boldsymbol{X}] = \boldsymbol{0}_p$，$V[\boldsymbol{X}] = \Sigma$ とする．$\boldsymbol{\beta}_i$ を Σ の第 i 固有値に対応する固有ベクトルとする．$\{\boldsymbol{\gamma}_1^*, \ldots, \boldsymbol{\gamma}_n^*\}$ を $\boldsymbol{X}$ の n-principal points とし，その n 点によって張られる線形部分空間 $\text{span}\{\boldsymbol{\gamma}_1^*, \ldots, \boldsymbol{\gamma}_n^*\}$ が q 次元であるとする．このとき，その線形部分空間は Σ の第 $1 \sim q$ 固有ベクトルが張る空間と等しい．すなわち，

$$\text{span}\{\boldsymbol{\gamma}_1^*, \ldots, \boldsymbol{\gamma}_n^*\} = \text{span}\{\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_q\}$$

である．

この定理は主部分空間定理（principal subspace theorem）と呼ばれ，principal points が存在する範囲を陽に示しているという点で重要な結果となっている．

一方，Li and Flury (1995) で指摘されているように，クラスター分析でしばしば想定されるような多数の群の混合分布における理論的結果を発展させることも重要であると考えられる．Yamamoto and Shinozaki (2000b) は，球面対称分布（spherically symmetric distribution）の位置混合分布の 2-principal points の 1 組以上が各球面対称分布の位置ベクトルが張る線形部分空間上に存在することを示した．

球面対称分布とは特性関数が

$$\phi(\boldsymbol{t}) = \psi(\|\boldsymbol{t}\|^2) \quad \text{for some } \psi : [0,\infty) \to [0,\infty)$$

で表される確率分布であり，直交変換に対して不変な分布である．2 次モーメントが有限であるとき，期待値は $\boldsymbol{0}$，共分散行列は単位行列の正の定数倍で与えられる．明らかに球面対称分布は楕円対称分布の特殊ケースである．球面対称分布の例としては，多変量標準正規分布 $N_p(\boldsymbol{0}_p, I_p)$ や単位超球面上 $S_p = \{\boldsymbol{x} \in R^p | \|\boldsymbol{x}\| = 1\}$ の一様分布などが挙げられる．

球面対称分布の位置混合分布は楕円対称分布の枠には収まらないため，Yamamoto and Shinozaki (2000b) が示した結果は Tarpey, Li, and Flury (1995) が導いたものとは別バージョンの主部分空間定理とみなすことができる．

近年，Kurata (2008) および Matsuura and Kurata (2010) は Yamamoto and Shinozaki (2000b) の結果を無限位置混合分布の 2-principal points に拡張し，さらに Matsuura and Kurata (2011) は $n(\geq 2)$-principal points を扱うことができるように拡張している．

命題 2. (Matsuura and Kurata (2010, Theorem 1), Matsuura and Kurata (2011, Theorems 1,2))
$\boldsymbol{Y}$ を有限な 2 次モーメントを持つ p 次元球面対称分布に従う確率変数ベクトルとし，$P(\boldsymbol{Y}=\boldsymbol{0}_p)=0$ であるとする．$\boldsymbol{U}$ を $E[\boldsymbol{U}]=\boldsymbol{0}_p$ で有限な 2 次モーメントを持つ p 次元確率変数ベクトルとし，ある $M'M=I_r$ となる $p\times r$ $(r<p)$ 行列 M が存在して

$$\boldsymbol{U}\in \mathrm{span}(M) \ \ \text{with probability 1}$$

となるとする．ただし，$\mathrm{span}(M)$ は M の r 本の列ベクトルによって張られる r 次元線形部分空間である．また，$\boldsymbol{Y}$ と $\boldsymbol{U}$ は独立であるとする．p 次元確率変数ベクトル $\boldsymbol{X}$ を

$$\boldsymbol{X}\equiv \boldsymbol{Y}+\boldsymbol{U} \tag{2.1}$$

と定義する．r 以下の次元の線形部分空間を張る $\boldsymbol{X}$ の n-principal points が 1 組以上存在すると仮定する．このとき，$\boldsymbol{X}$ の n-principal points $\{\gamma_1^*,\ldots,\gamma_n^*\}$ の 1 組以上が

$$\gamma_1^*,\ldots,\gamma_n^*\in \mathrm{span}(M) \tag{2.2}$$

を満たす．特に，$n=2$ のとき，または $\boldsymbol{Y}$ の一次元周辺分布の確率密度関数が全範囲で正の値を取るとき（例えば $\boldsymbol{Y}$ が p 変量標準正規分布に従う場合など），r 以下の次元の線形部分空間を張る $\boldsymbol{X}$ の n-principal points は必ず (2.2) を満たす．

この命題において，2-principal points に係わる部分は Matsuura and Kurata (2010, Theorem 1) によるものであり，その他の部分は Matsuura and Kurata (2011, Theorems 1,2) によるものである．

Matsuura and Kurata (2010,2011) で扱っているモデル (2.1) において，$\boldsymbol{U}$ が

$$P(\boldsymbol{U}=\boldsymbol{\mu}_i)=\frac{1}{m},\ i=1,\ldots,m \ \ \text{with} \ \sum_{i=1}^{m}\boldsymbol{\mu}_i=\boldsymbol{0}_p$$

の離散分布に従うとし，$\boldsymbol{Y}$ の確率密度関数が存在すると仮定してそれを f とおけば，$\boldsymbol{X}$ の確率密度関数 $g(\boldsymbol{x})$ は Yamamoto and Shinozaki (2000b) で想定されたモデル

$$g(\boldsymbol{x})=\frac{1}{m}\sum_{i=1}^{m}f(\boldsymbol{x}-\boldsymbol{\mu}_i)$$

になる（$\boldsymbol{Y}$ が確率密度関数を持つとき，必ず $P(\boldsymbol{Y}=\boldsymbol{0}_p)=0$ となることに注意する）．このとき，M は $\mathrm{span}\{\boldsymbol{\mu}_1,\ldots,\boldsymbol{\mu}_m\}$ の正規直交基底を並べた行列である．このモデルは球面対称分布の有限位置混合分布とみなすことができる．一方，$\boldsymbol{U}$ が連続分布に従うとすれば，$\boldsymbol{X}$ の分布は球面対称分布の無限位置混合分布と解釈することができる．特に，$\boldsymbol{\theta}$ を直交変換に対して不変な分布に従う m 次元確率変数ベクトルとして，

$$\boldsymbol{U}\equiv H\boldsymbol{\theta} \ \ \text{with} \ H=(\boldsymbol{\mu}_1,\ldots,\boldsymbol{\mu}_m):p\times m$$

とおくと，このときの $\boldsymbol{X}$ の分布は Kurata (2008) において想定された分布になる．

命題 2 では，$\boldsymbol{Y}$ と $\boldsymbol{U}$ が独立，すなわち各位置ベクトルごとの球面対称分布が同一であるという仮定をおいているが，次章において

$$g(\boldsymbol{x}) = \sum_{i=1}^{m} f_i(\boldsymbol{x} - \boldsymbol{\mu}_i)p_i \tag{2.3}$$

の形で表される確率分布を含むような異なる球面対称分布の位置混合分布への拡張を行う．また，$P(\boldsymbol{Y} = \boldsymbol{0}_p) = 0$ の仮定も外して議論する．

3 異なる球面対称分布の位置混合分布の principal points における主部分空間定理

$\boldsymbol{U}$ を $E[\boldsymbol{U}] = \boldsymbol{0}_p$ で有限な 2 次モーメントを持つ p 次元確率変数ベクトルとし，ある $M'M = I_r$ となる $p \times r$ $(r < p)$ 行列 M が存在して

$$\boldsymbol{U} \in \mathrm{span}(M) \quad \text{with probability 1} \tag{3.1}$$

となるとする．M は $\boldsymbol{U}$ のサポートが張る線形部分空間の正規直交基底を並べた行列になっている．このとき，

$$V[\boldsymbol{U}] = M\Psi M' \quad \text{for some } \Psi : r \times r$$

と表される．

p 次元確率変数ベクトル $\boldsymbol{Y}$ は $\boldsymbol{U}$ に依存し，$\boldsymbol{Y}|(\boldsymbol{U} = \boldsymbol{u})$（$\boldsymbol{U} = \boldsymbol{u}$ 条件付きの確率変数ベクトル $\boldsymbol{Y}$）が全ての $\boldsymbol{u}$ の値において有限な 2 次モーメントを持つ球面対称分布に従うとする．$E[\boldsymbol{Y}|\boldsymbol{U}] = \boldsymbol{0}_p$ であり，

$$V[\boldsymbol{Y}|\boldsymbol{U} = \boldsymbol{u}] = \sigma^2(\boldsymbol{u})I_p \quad \text{for some } \sigma^2 : R^p \to (0, \infty)$$

と表すことができる．また，$\boldsymbol{Y}$ は球面対称分布に従い，

$$V[\boldsymbol{Y}] = E[\sigma^2(\boldsymbol{U})]I_p$$

となる．

p 次元確率変数ベクトル $\boldsymbol{X}$ を

$$\boldsymbol{X} \equiv \boldsymbol{Y} + \boldsymbol{U} \tag{3.2}$$

と定義する．$\boldsymbol{X}$ の共分散行列 Σ は

$$\Sigma = E[\sigma^2(\boldsymbol{U})]I_p + M\Psi M'$$

と表されることから，Σ の第 i 固有値に対応する固有ベクトルを $\boldsymbol{\beta}_i$ とおくと，

$$\mathrm{span}\{\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_r\} = \mathrm{span}(M)$$

が成立する．

例えば，モデル (3.2) において，$\boldsymbol{U}$ が

$$P(\boldsymbol{U}=\boldsymbol{\mu}_i)=p_i,\ i=1,\ldots,m \ \text{ with } \sum_{i=1}^{m} p_i = 1 \text{ and } \sum_{i=1}^{m} \boldsymbol{\mu}_i p_i = \mathbf{0}_p \tag{3.3}$$

の離散分布に従うとし，$\boldsymbol{Y}|(\boldsymbol{U}=\boldsymbol{\mu}_i)$ の確率密度関数が存在すると仮定してそれを f_i とおけば，$\boldsymbol{X}$ の確率密度関数は (2.3) になる．

モデル (3.2) で定義される確率変数ベクトル $\boldsymbol{X}$ の n-principal points について以下の定理が導かれる．

定理 1.
$\boldsymbol{X}$ が (3.2) によって定義される p 次元確率変数ベクトルであるとする．r 以下の次元の線形部分空間を張る $\boldsymbol{X}$ の n-principal points が 1 組以上存在すると仮定する．このとき，$\boldsymbol{X}$ の n-principal points $\{\gamma_1^*,\ldots,\gamma_n^*\}$ の 1 組以上が

$$\gamma_1^*,\ldots,\gamma_n^* \in \mathrm{span}(M) \tag{3.4}$$

を満たす．特に，$\boldsymbol{Y}|\boldsymbol{U}$ の一次元周辺分布の確率密度関数が全ての $\boldsymbol{U}$ の値に対して存在し，かつ全範囲で正の値を取るとき，r 以下の次元の線形部分空間を張る $\boldsymbol{X}$ の n-principal points は必ず (3.4) を満たす．

この定理の証明は付録に記した．

期待値が $\mathbf{0}_p$ である p 次元確率分布の n-principal points は必ず $\min\{n-1,p\}$ 以下の次元の線形部分空間を張ることから，以下の定理が得られる．

定理 2.
$\boldsymbol{X}$ が (3.2) によって定義される p 次元確率変数ベクトルであるとする．また，$n \le r+1$ とする．このとき，$\boldsymbol{X}$ の n-principal points $\{\gamma_1^*,\ldots,\gamma_n^*\}$ の 1 組以上が (3.4) を満たす．特に，$\boldsymbol{Y}|\boldsymbol{U}$ の一次元周辺分布の確率密度関数が全ての $\boldsymbol{U}$ の値に対して存在し，かつ全範囲で正の値を取るとき，$\boldsymbol{X}$ の n-principal points は必ず (3.4) を満たす．

従って，異なる球面対称分布の位置混合分布の $n(\le r+1)$-principal points を求めたいとき，探索範囲を $\mathrm{span}(M)$ に限定することができることがわかる．

本稿では $E[\boldsymbol{X}]=\mathbf{0}_p$ を仮定して議論してきたが，より一般に $E[\boldsymbol{X}]=E[\boldsymbol{U}]=\boldsymbol{\mu}$ の場合，定理 2 は次のように書き直される．

定理 2'.
$\boldsymbol{U}$ を $E[\boldsymbol{U}]=\boldsymbol{\mu}$ で有限な 2 次モーメントを持つ p 次元確率変数ベクトルとし，ある $M'M=I_r$ となる $p\times r$ $(r<p)$ 行列 M が存在して

$$\boldsymbol{U}-\boldsymbol{\mu} \in \mathrm{span}(M) \ \text{ with probability 1}$$

となるとする．p 次元確率変数ベクトル $\boldsymbol{Y}$ は $\boldsymbol{U}$ に依存し，$\boldsymbol{Y}|\boldsymbol{U}$ が全ての $\boldsymbol{U}$ の値において有限な 2 次モーメントを持つ球面対称分布に従うとする．p 次元確率変数ベクトル $\boldsymbol{X}$ を $\boldsymbol{X}\equiv\boldsymbol{Y}+\boldsymbol{U}$ と定義する．$n\le r+1$ とする．このとき，$\boldsymbol{X}$ の n-principal points

$\{\gamma_1^*, \ldots, \gamma_n^*\}$ の 1 組以上が

$$\gamma_1^* - \boldsymbol{\mu}, \ldots, \gamma_n^* - \boldsymbol{\mu} \in \mathrm{span}(M) \tag{3.5}$$

を満たす．特に，$\boldsymbol{Y}|\boldsymbol{U}$ の一次元周辺分布の確率密度関数が全ての $\boldsymbol{U}$ の値に対して存在し，かつ全範囲で正の値を取るとき，$\boldsymbol{X}$ の n-principal points は必ず (3.5) を満たす．

定理 1 も同様に $E[\boldsymbol{X}] = E[\boldsymbol{U}] = \boldsymbol{\mu}$ の場合に対応して書き直すことができるが，ここでは省略する．

以下では，モデル (3.2) に含まれる多次元確率分布の例をいくつか挙げることにする．

例 1. (複数の多変量正規分布の混合分布)
$\boldsymbol{U}$ が (3.3) の離散分布に従い，各 $\boldsymbol{U} = \boldsymbol{\mu}_i$, $i = 1, \ldots, m$ ごとの $\boldsymbol{Y}$ の条件付き分布が

$$\boldsymbol{Y}|(\boldsymbol{U} = \boldsymbol{\mu}_i) \sim N_p(\boldsymbol{0}_p, \sigma_i^2 I_p)$$

であるとする．$\{\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_m\}$ によって張られる線形部分空間の次元を r とおく．$\boldsymbol{X} \equiv \boldsymbol{Y} + \boldsymbol{U}$ とおくと，これは複数の多変量正規分布の混合分布

$$N_p(\boldsymbol{\mu}_1, \sigma_1^2 I_p) \times p_1 + \cdots + N_p(\boldsymbol{\mu}_m, \sigma_m^2 I_p) \times p_m$$

に従う．このとき，定理 2 より，$\boldsymbol{X}$ の $n(\leq r+1)$-principal points は必ず $\mathrm{span}\{\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_m\}$ 上に存在することが保証される．

例 2. (多変量 t 分布の位置混合分布)
$\boldsymbol{U}$ を $E[\boldsymbol{U}] = \boldsymbol{0}_p$ で有限な 2 次モーメントを持つ p 次元確率変数ベクトルとし，ある $M'M = I_r$ となる $p \times r$ $(r < p)$ 行列 M が存在して (3.1) を満たすとする．$\boldsymbol{Z}$ を $N_p(\boldsymbol{0}, I_p)$ に従う p 次元確率変数ベクトルとし，W は $\boldsymbol{U}$ に依存する確率変数で，$W|\boldsymbol{U}$ は全ての $\boldsymbol{U}$ の値において有限な 2 次モーメントを持つ正値確率変数であるとする．$\boldsymbol{Y} \equiv W\boldsymbol{Z}$ とおくと，$\boldsymbol{Y}|\boldsymbol{U}$ は多変量正規分布の尺度混合分布に従う．$\nu(\boldsymbol{u})$ を $R^p \to \{x | x = 3, 4, \ldots\}$ の関数とし，$\sqrt{\nu(\boldsymbol{u})}/W|(\boldsymbol{U} = \boldsymbol{u})$ が自由度 $\nu(\boldsymbol{u})$ の χ 分布に従うとすると，$\boldsymbol{Y}|(\boldsymbol{U} = \boldsymbol{u})$ は自由度 $\nu(\boldsymbol{u})$ の多変量 t 分布に従う．$\boldsymbol{X} \equiv \boldsymbol{Y} + \boldsymbol{U}$ とおくと，$\boldsymbol{X}$ の分布は様々な自由度を持つ多変量 t 分布の位置混合分布となる．このとき，$\boldsymbol{X}$ の $n(\leq r+1)$-principal points は必ず $\mathrm{span}(M)$ 上に存在する．

例 3. (超球面上の一様分布の位置混合分布)
例 1 および例 2 では $\boldsymbol{Y}|\boldsymbol{U}$ の確率密度関数が存在しているが，モデル (3.2) は確率密度関数が存在しない場合も含んでいる．$\boldsymbol{U}$ を $E[\boldsymbol{U}] = \boldsymbol{0}_p$ で有限な 2 次モーメントを持つ p 次元確率変数ベクトルとし，ある $M'M = I_r$ となる $p \times r$ $(r < p)$ 行列 M が存在して (3.1) を満たすとする．$\boldsymbol{Y}$ を p 次元確率変数ベクトルとし，$\boldsymbol{Y}|(\boldsymbol{U} = \boldsymbol{u})$ は半径 $R(\boldsymbol{u}) : R^p \to (0, \infty)$ の超球面上 $S_p(\boldsymbol{u}) = \{\boldsymbol{x} \in R^p | \|\boldsymbol{x}\| = R(\boldsymbol{u})\}$ の一様分布に従うとする．このとき，$\boldsymbol{Y}|\boldsymbol{U}$ は確率密度関数を持たないが，$\boldsymbol{X} \equiv \boldsymbol{Y} + \boldsymbol{U}$ とおくと，$\boldsymbol{X}$ の $n(\leq r+1)$-principal points の 1 組以上が $\mathrm{span}(M)$ 上に存在する．

4 おわりに

本稿では，Matsuura and Kurata (2011) で示された原点を取る確率が 0 の球面対称分布の位置混合分布の principal points における主部分空間定理を異なる一般の（原点を取る確率が正の場合も含む）球面対称分布の位置混合分布を扱うことができるように拡張した．具体的には，異なる球面対称分布の位置混合分布の n-principal points が，いくつかの条件のもと，各球面対称分布の位置ベクトルが張る線形部分空間上に存在することを示した．この結果は，母集団が群構造を持ち各群が異なるばらつきを持つ混合分布であるときの principal points の探索や推定に有用であると考える．

今後の課題としては，

$$N_p(\boldsymbol{\mu}_1, \Sigma_1) \times p_1 + \cdots + N_p(\boldsymbol{\mu}_m, \Sigma_m) \times p_m$$

など，各群の共分散行列が単位行列の正の定数倍とは限らないときにも適用可能な結果を導くことが挙げられる．

付録：定理 1 の証明

Matsuura and Kurata (2011, Section 5) より，一般性を失うことなく $M = \binom{I_r}{\mathbf{0}_{(p-r)\times r}}$ の場合に限定して議論することができ，このとき，ある r 次元確率変数ベクトル $\boldsymbol{V}$ が存在して $\boldsymbol{U} \stackrel{d}{=} \binom{\boldsymbol{V}}{\mathbf{0}_{p-r}}$ が成立する．ただし，$\stackrel{d}{=}$ は確率分布として等しいの意味である．定理 1 を証明するためには以下の補題を示せば十分である．

補題 1.
$\boldsymbol{V}$ を $E[\boldsymbol{V}] = \mathbf{0}_r$ で正定符号の共分散行列を持つ r 次元確率変数ベクトルとする．p 次元確率変数ベクトル $\boldsymbol{Y}$ は $\boldsymbol{V}$ に依存し，$\boldsymbol{Y}|\boldsymbol{V}$ が全ての $\boldsymbol{V}$ の値において有限な 2 次モーメントを持つ球面対称分布に従うとする．p 次元確率変数ベクトル $\boldsymbol{X}$ を

$$\boldsymbol{X} \equiv \boldsymbol{Y} + \begin{pmatrix} \boldsymbol{V} \\ \mathbf{0}_{p-r} \end{pmatrix}$$

と定義する．$E = \binom{I_r}{\mathbf{0}_{(p-r)\times r}}$ とおく．R^p 上の n 点 $\{\gamma_1, \ldots, \gamma_n\}$ が以下の条件 (i)-(iii)：
(i) convex hull が $E[\boldsymbol{X}] = \mathbf{0}_p$ を含む
(ii) $\mathrm{span}\{\gamma_1, \ldots, \gamma_n\}$ の次元を q とおくと $q \leq r$ が成立する
(iii) $\mathrm{span}\{\gamma_1, \ldots, \gamma_n\} \not\subset \mathrm{span}(E)$
を満たすとする．このとき，ある R^p 上の n 点 $\{\boldsymbol{c}_1, \ldots, \boldsymbol{c}_n\}$ が存在し，

$$\boldsymbol{c}_1, \ldots, \boldsymbol{c}_n \in \mathrm{span}(E)$$

および

$$E[d^2(\boldsymbol{X}|\boldsymbol{c}_1, \ldots, \boldsymbol{c}_n)] \leq E[d^2(\boldsymbol{X}|\gamma_1, \ldots, \gamma_n)] \tag{4.1}$$

が成立する．特に，$\boldsymbol{Y}|(\boldsymbol{V} = \boldsymbol{v})$ の一次元周辺分布の確率密度関数 $f_1(y|\boldsymbol{v})$ が全ての $\boldsymbol{v}$ の値に対して存在し，かつ

$$f_1(y|\boldsymbol{v}) > 0 \text{ for any } y \in R \tag{4.2}$$

を満たすとき，(4.1) 式の不等式は強意となる．

証明.

R^p 上の n 点 $\{\gamma_1,\ldots,\gamma_n\}$ が条件 (i)-(iii) を満たしているとする．このとき，Matsuura and Kurata (2011, Section 5) の議論より，ある自然数 $d(\le \min\{q,p-r\})$ が存在し，d 次元ベクトル $\boldsymbol{\tau}=(\tau_1,\ldots,\tau_d)'\in[0,1)^d\equiv[0,1)\times\cdots\times[0,1)$，$R^q$ 上の n 点 $\boldsymbol{\xi}_i=(\xi_{i1},\ldots,\xi_{iq})'$，$i=1,\ldots,n$，互いに直交する q 本の長さ 1 の $r\times 1$ ベクトル $\boldsymbol{t}_{11},\ldots,\boldsymbol{t}_{1q}$，互いに直交する d 本の長さ 1 の $(p-r)\times 1$ ベクトル $\boldsymbol{t}_{21},\ldots,\boldsymbol{t}_{2d}$ を適当に選び，$\boldsymbol{a}=(a_1,\ldots,a_d)'\in[-1,1]^d\equiv[-1,1]\times\cdots\times[-1,1]$ 上の $p\times q$ 行列の関数 $T(\boldsymbol{a}):[-1,1]^d\to R^{p\times q}$ を

$$T(\boldsymbol{a})=\begin{pmatrix}T_1(\boldsymbol{a})\\T_2(\boldsymbol{a})\end{pmatrix}=\begin{pmatrix}a_1\boldsymbol{t}_{11}&\cdots&a_d\boldsymbol{t}_{1d}&\boldsymbol{t}_{1d+1}&\cdots&\boldsymbol{t}_{1q}\\\sqrt{1-a_1^2}\boldsymbol{t}_{21}&\cdots&\sqrt{1-a_d^2}\boldsymbol{t}_{2d}&\boldsymbol{0}_{p-r}&\cdots&\boldsymbol{0}_{p-r}\end{pmatrix}$$

と定義することで，

$$\{\gamma_1,\ldots,\gamma_n\}=\{T(\boldsymbol{\tau})\boldsymbol{\xi}_1,\ldots,T(\boldsymbol{\tau})\boldsymbol{\xi}_n\}$$

が成立する．$\boldsymbol{a}\in[-1,1]^d$ で常に $T(\boldsymbol{a})'T(\boldsymbol{a})=I_q$ が成立する．また，$\boldsymbol{e}\in\{-1,1\}^d\equiv\{-1,1\}\times\cdots\times\{-1,1\}$ のとき，

$$T(\boldsymbol{e})\boldsymbol{\xi}_1,\ldots,T(\boldsymbol{e})\boldsymbol{\xi}_n\in\mathrm{span}(E)$$

が成立することに注意する．

補題の証明のためには，

$$E[d^2(\boldsymbol{X}|T(\boldsymbol{e})\boldsymbol{\xi}_1,\ldots,T(\boldsymbol{e})\boldsymbol{\xi}_n)]\le E[d^2(\boldsymbol{X}|T(\boldsymbol{\tau})\boldsymbol{\xi}_1,\ldots,T(\boldsymbol{\tau})\boldsymbol{\xi}_n)]=E[d^2(\boldsymbol{X}|\gamma_1,\ldots,\gamma_n)] \tag{4.3}$$

が成立する $\boldsymbol{e}\in\{-1,1\}^d$ が存在することを示し，さらに $\boldsymbol{Y}|(\boldsymbol{V}=\boldsymbol{v})$ の一次元周辺分布の確率密度関数 $f_1(y|\boldsymbol{v})$ が全ての $\boldsymbol{v}$ の値に対して存在し，かつ (4.2) を満たすとき，(4.3) 式の不等式が強意となることを示せば十分である．

$$L(\boldsymbol{a})=E[d^2(\boldsymbol{X}|T(\boldsymbol{a})\boldsymbol{\xi}_1,\ldots,T(\boldsymbol{a})\boldsymbol{\xi}_n)]\left(=E\left[\min_i\|\boldsymbol{X}-T(\boldsymbol{a})\boldsymbol{\xi}_i\|^2\right]\right)$$

とおく．$\boldsymbol{Y}_q$ を $\boldsymbol{Y}$ の q 次元周辺分布に従う確率変数ベクトルとし，$T(\boldsymbol{a})'\boldsymbol{Y}|\boldsymbol{V}\overset{d}{=}\boldsymbol{Y}_q|\boldsymbol{V}$ が成立することに注意して $L(\boldsymbol{a})$ を展開すると，

$$\begin{aligned}L(\boldsymbol{a})&=E\left[\min_i\left\{\|\boldsymbol{\xi}_i\|^2-2\boldsymbol{\xi}_i'(\boldsymbol{Y}_q+T_1(\boldsymbol{a})'\boldsymbol{V})\right\}\right]+E[\boldsymbol{Y}'\boldsymbol{Y}]+E[\boldsymbol{V}'\boldsymbol{V}]\\&=E\left[\min_i\left\{\|\boldsymbol{\xi}_i\|^2-2\boldsymbol{\xi}_i'(\boldsymbol{Y}_q+T_1(\boldsymbol{a})'\boldsymbol{V})\right\}\,\middle|\,\boldsymbol{Y}\ne\boldsymbol{0}_p\right]P(\boldsymbol{Y}\ne\boldsymbol{0}_p)\\&\quad+E\left[\min_i\left\{\|\boldsymbol{\xi}_i\|^2-2\boldsymbol{\xi}_i'T_1(\boldsymbol{a})'\boldsymbol{V}\right\}\,\middle|\,\boldsymbol{Y}=\boldsymbol{0}_p\right]P(\boldsymbol{Y}=\boldsymbol{0}_p)+E[\boldsymbol{Y}'\boldsymbol{Y}]+E[\boldsymbol{V}'\boldsymbol{V}]\end{aligned}$$

となる．

$$\begin{aligned}L_1(\boldsymbol{a})&=E\left[\min_i\left\{\|\boldsymbol{\xi}_i\|^2-2\boldsymbol{\xi}_i'(\boldsymbol{Y}_q+T_1(\boldsymbol{a})'\boldsymbol{V})\right\}\,\middle|\,\boldsymbol{Y}\ne\boldsymbol{0}_p\right],\\L_2(\boldsymbol{a})&=E\left[\min_i\left\{\|\boldsymbol{\xi}_i\|^2-2\boldsymbol{\xi}_i'T_1(\boldsymbol{a})'\boldsymbol{V}\right\}\,\middle|\,\boldsymbol{Y}=\boldsymbol{0}_p\right]\end{aligned}$$

とおけば,

$$L(\boldsymbol{a}) = L_1(\boldsymbol{a})P(\boldsymbol{Y} \neq \boldsymbol{0}_p) + L_2(\boldsymbol{a})P(\boldsymbol{Y} = \boldsymbol{0}_p) + E[\boldsymbol{Y}'\boldsymbol{Y}] + E[\boldsymbol{V}'\boldsymbol{V}] \tag{4.4}$$

と表すことができる.

$L_1(\boldsymbol{a})$ の性質については $P(\boldsymbol{Y} = \boldsymbol{0}_p) = 0$ および $\boldsymbol{Y} \perp\!\!\!\perp \boldsymbol{V}$ の仮定をおいて議論している Matsuura and Kurata (2011, Section 5) とほぼ同様の展開を行うことができる.

$$L_1(\boldsymbol{a}) = E\left[E\left[\min_i \left\{\|\boldsymbol{\xi}_i\|^2 - 2\boldsymbol{\xi}_i'(\boldsymbol{Y}_q + T_1(\boldsymbol{a})'\boldsymbol{V})\right\} \,\middle|\, \boldsymbol{V}, (\boldsymbol{Y} \neq \boldsymbol{0}_p)\right] \,\middle|\, \boldsymbol{Y} \neq \boldsymbol{0}_p\right]$$

であり，$\boldsymbol{Y}|(\boldsymbol{V} = \boldsymbol{v}, \boldsymbol{Y} \neq \boldsymbol{0}_p)$ の任意の周辺分布は必ず確率密度関数を持つことから (Fang, Kotz, and Ng (1990, Theorem 2.10))，$\boldsymbol{Y}_q|(\boldsymbol{V} = \boldsymbol{v}, \boldsymbol{Y} \neq \boldsymbol{0}_p)$ の確率密度関数を $f_q(y_1, \ldots, y_q|\boldsymbol{v})$ とおき，さらに $\boldsymbol{V}^\dagger \equiv \boldsymbol{V}|(\boldsymbol{Y} \neq \boldsymbol{0}_p)$ とおくと,

$$\begin{aligned} L_1(\boldsymbol{a}) &= E\left[\int_{R^q} \min_i \left\{\|\boldsymbol{\xi}_i\|^2 - 2\boldsymbol{\xi}_i'(\boldsymbol{y} + T_1(\boldsymbol{a})'\boldsymbol{V})\right\} f_q(\boldsymbol{y}|\boldsymbol{V})d\boldsymbol{y} \,\middle|\, \boldsymbol{Y} \neq \boldsymbol{0}_p\right] \\ &= E\left[\int_{R^q} \min_i \left\{\|\boldsymbol{\xi}_i\|^2 - 2\boldsymbol{\xi}_i'\left(\boldsymbol{y} + T_1(\boldsymbol{a})'\boldsymbol{V}^\dagger\right)\right\} f_q(\boldsymbol{y}|\boldsymbol{V}^\dagger)d\boldsymbol{y}\right] \end{aligned}$$

となる．ここで,

$$\begin{aligned} R_i = \{&\boldsymbol{x} \in R^q | \|\boldsymbol{x} - \boldsymbol{\xi}_i\| < \|\boldsymbol{x} - \boldsymbol{\xi}_j\|,\ j = 1, \ldots, i-1,\ \|\boldsymbol{x} - \boldsymbol{\xi}_i\| \leq \|\boldsymbol{x} - \boldsymbol{\xi}_j\|, \\ &j = i+1, \ldots, n\},\ i = 1, \ldots, n, \end{aligned}$$

$\boldsymbol{y}_2 = (y_2, \ldots, y_q)'$, $\boldsymbol{\xi}_{i2} = (\xi_{i2}, \ldots, \xi_{iq})'$, $i = 1, \ldots, n$, $R_i(\boldsymbol{y}_2) = \{y_1 \mid \binom{y_1}{\boldsymbol{y}_2} \in R_i\}$, $i = 1, \ldots, n$ とおく．$R_i(\boldsymbol{y}_2)$ は空集合，1 点集合，1 つの区間のどれかに限られることに注意する．空集合または 1 点集合（すなわちルベーグ測度が 0 の集合）ではない $R_i(\boldsymbol{y}_2)$ の各 $\boldsymbol{y}_2$ ごとの個数を $n(\boldsymbol{y}_2)$ とおく．$2 \leq n(\boldsymbol{y}_2) \leq n$ であることが保証される．空集合または 1 点集合ではない $n(\boldsymbol{y}_2)$ 個の $R_i(\boldsymbol{y}_2)$ について，任意の $j < l$ で

$$a > b \text{ for any } a \in R_{c(j|\boldsymbol{y}_2)}(\boldsymbol{y}_2),\ b \in R_{c(l|\boldsymbol{y}_2)}(\boldsymbol{y}_2)$$

が成立するように関数 $c(i|\boldsymbol{y}_2) : \{1, \ldots, n(\boldsymbol{y}_2)\} \to \{1, \ldots, n\}$ を定める．明らかに

$$\xi_{c(i|\boldsymbol{y}_2)1} - \xi_{c(i+1|\boldsymbol{y}_2)1} > 0,\ i = 1, \ldots, n(\boldsymbol{y}_2) - 1$$

が成立する.

$$\begin{aligned} h(\boldsymbol{y}_2|i) &= \frac{\|\boldsymbol{y}_2 - \boldsymbol{\xi}_{c(i|\boldsymbol{y}_2)2}\|^2 - \|\boldsymbol{y}_2 - \boldsymbol{\xi}_{c(i+1|\boldsymbol{y}_2)2}\|^2 + \xi^2_{c(i|\boldsymbol{y}_2)1} - \xi^2_{c(i+1|\boldsymbol{y}_2)1}}{2(\xi_{c(i|\boldsymbol{y}_2)1} - \xi_{c(i+1|\boldsymbol{y}_2)1})}, \\ &\quad i = 1, \ldots, n(\boldsymbol{y}_2) - 1, \end{aligned} \tag{4.5}$$

$h(\boldsymbol{y}_2|0) = \infty$, $h(\boldsymbol{y}_2|n(\boldsymbol{y}_2)) = -\infty$ とおくと,

$$R_{c(i|\boldsymbol{y}_2)}(\boldsymbol{y}_2) = \{y_1 | h(\boldsymbol{y}_2|i) < y_1 \leq h(\boldsymbol{y}_2|i-1)\},\ i = 1, \ldots, n(\boldsymbol{y}_2)$$

と表わすことができる．また，$\boldsymbol{a}_2=(a_2,\ldots,a_d)'$ および $T_{12}(\boldsymbol{a}_2)=(a_2\boldsymbol{t}_{12},\cdots,a_d\boldsymbol{t}_{1d},\boldsymbol{t}_{1d+1},\cdots,\boldsymbol{t}_{1q})$ とおく．以上の表記を用いると，

$$\begin{aligned}
L_1(\boldsymbol{a}) &= E\left[\sum_{i=1}^{n}\int_{R_i}(\|\boldsymbol{\xi}_i\|^2-2\boldsymbol{\xi}_i'\boldsymbol{y})f_q(\boldsymbol{y}-T_1(\boldsymbol{a})'\boldsymbol{V}^\dagger|\boldsymbol{V}^\dagger)d\boldsymbol{y}\right]\\
&= E\left[\int_{R^{q-1}}\sum_{i=1}^{n(\boldsymbol{y}_2)}\left\{\int_{h(\boldsymbol{y}_2|i)}^{h(\boldsymbol{y}_2|i-1)}(\xi_{c(i|\boldsymbol{y}_2)1}^2-2\xi_{c(i|\boldsymbol{y}_2)1}y_1+\|\boldsymbol{\xi}_{c(i|\boldsymbol{y}_2)2}\|^2-2\boldsymbol{\xi}_{c(i|\boldsymbol{y}_2)2}'\boldsymbol{y}_2)\right.\right.\\
&\qquad\left.\left. f_q(y_1-a_1\boldsymbol{t}_{11}'\boldsymbol{V}^\dagger,\boldsymbol{y}_2-T_{12}(\boldsymbol{a}_2)'\boldsymbol{V}^\dagger|\boldsymbol{V}^\dagger)dy_1\right\}d\boldsymbol{y}_2\right]\\
&= E\left[\int_{R^{q-1}}\sum_{i=1}^{n(\boldsymbol{y}_2)}\left\{\int_{h(\boldsymbol{y}_2|i)-a_1\boldsymbol{t}_{11}'\boldsymbol{V}^\dagger}^{h(\boldsymbol{y}_2|i-1)-a_1\boldsymbol{t}_{11}'\boldsymbol{V}^\dagger}\left(\xi_{c(i|\boldsymbol{y}_2)1}^2-2\xi_{c(i|\boldsymbol{y}_2)1}\left(y_1+a_1\boldsymbol{t}_{11}'\boldsymbol{V}^\dagger\right)\right.\right.\right.\\
&\qquad\left.\left.\left.+\|\boldsymbol{\xi}_{c(i|\boldsymbol{y}_2)2}\|^2-2\boldsymbol{\xi}_{c(i|\boldsymbol{y}_2)2}'\boldsymbol{y}_2\right)f_q(y_1,\boldsymbol{y}_2-T_{12}(\boldsymbol{a}_2)'\boldsymbol{V}^\dagger|\boldsymbol{V}^\dagger)dy_1\right\}d\boldsymbol{y}_2\right]
\end{aligned}$$

となる．$h(\boldsymbol{y}_2|i)$ の定義 (4.5) より，

$$\begin{aligned}
&\xi_{c(i|\boldsymbol{y}_2)1}^2-2\xi_{c(i|\boldsymbol{y}_2)1}h(\boldsymbol{y}_2|i)+\|\boldsymbol{\xi}_{c(i|\boldsymbol{y}_2)2}\|^2-2\boldsymbol{\xi}_{c(i|\boldsymbol{y}_2)2}'\boldsymbol{y}_2\\
&-\left(\xi_{c(i+1|\boldsymbol{y}_2)1}^2-2\xi_{c(i+1|\boldsymbol{y}_2)1}h(\boldsymbol{y}_2|i)+\|\boldsymbol{\xi}_{c(i+1|\boldsymbol{y}_2)2}\|^2-2\boldsymbol{\xi}_{c(i+1|\boldsymbol{y}_2)2}'\boldsymbol{y}_2\right)\\
=\;&\|\boldsymbol{y}_2-\boldsymbol{\xi}_{c(i|\boldsymbol{y}_2)2}\|^2-\|\boldsymbol{y}_2-\boldsymbol{\xi}_{c(i+1|\boldsymbol{y}_2)2}\|^2+\xi_{c(i|\boldsymbol{y}_2)1}^2-\xi_{c(i+1|\boldsymbol{y}_2)1}^2\\
&-2(\xi_{c(i|\boldsymbol{y}_2)1}-\xi_{c(i+1|\boldsymbol{y}_2)1})h(\boldsymbol{y}_2|i)\\
=\;&0,\ i=1,\ldots,n(\boldsymbol{y}_2)-1
\end{aligned}$$

が成立することに注意して，$L_1(\boldsymbol{a})$ を a_1 で微分すると，

$$\begin{aligned}
&\frac{\partial L_1(\boldsymbol{a})}{\partial a_1}\\
&=E\left[\int_{R^{q-1}}\sum_{i=1}^{n(\boldsymbol{y}_2)}\int_{h(\boldsymbol{y}_2|i)-a_1\boldsymbol{t}_{11}'\boldsymbol{V}^\dagger}^{h(\boldsymbol{y}_2|i-1)-a_1\boldsymbol{t}_{11}'\boldsymbol{V}^\dagger}\left(-2\xi_{c(i|\boldsymbol{y}_2)1}\boldsymbol{t}_{11}'\boldsymbol{V}^\dagger\right)f_q(y_1,\boldsymbol{y}_2-T_{12}(\boldsymbol{a}_2)'\boldsymbol{V}^\dagger|\boldsymbol{V}^\dagger)dy_1d\boldsymbol{y}_2\right]
\end{aligned}$$

となる．さらに，

$$\begin{aligned}
\frac{\partial^2 L_1(\boldsymbol{a})}{\partial a_1^2}&=-2E\left[\left(\boldsymbol{t}_{11}'\boldsymbol{V}^\dagger\right)^2\int_{R^{q-1}}\sum_{i=1}^{n(\boldsymbol{y}_2)-1}\left(\xi_{c(i|\boldsymbol{y}_2)1}-\xi_{c(i+1|\boldsymbol{y}_2)1}\right)\right.\\
&\qquad\left. f_q(h(\boldsymbol{y}_2|i)-a_1\boldsymbol{t}_{11}'\boldsymbol{V}^\dagger,\boldsymbol{y}_2-T_{12}(\boldsymbol{a}_2)'\boldsymbol{V}^\dagger|\boldsymbol{V}^\dagger)d\boldsymbol{y}_2\right]\\
&\le 0
\end{aligned}$$

となる．従って，$L_1(\boldsymbol{a})$ は a_1 について凹関数であることがわかる．$a_2,\ldots,a_d$ に関しても同様の議論が成立することから，$L_1(\boldsymbol{a})$ は $a_1,\ldots,a_d$ のそれぞれについて凹関数であることが示された．

次に，$L_2(\boldsymbol{a})$ の性質について議論する．上と同様の表記を用いることで

$$L_2(\boldsymbol{a}) = E\left[\sum_{i=1}^{n} I(T_1(\boldsymbol{a})'\boldsymbol{V} \in R_i)(\|\boldsymbol{\xi}_i\|^2 - 2\boldsymbol{\xi}_i'T_1(\boldsymbol{a})'\boldsymbol{V}) \,\middle|\, \boldsymbol{Y} = \boldsymbol{0}_p\right]$$

$$= E\Bigg[\sum_{i=1}^{n(T_{12}(\boldsymbol{a}_2)'\boldsymbol{V})} I(h(T_{12}(\boldsymbol{a}_2)'\boldsymbol{V}|i) < a_1\boldsymbol{t}_{11}'\boldsymbol{V} \le h(T_{12}(\boldsymbol{a}_2)'\boldsymbol{V}|i-1))$$
$$(\xi^2_{c(i|T_{12}(\boldsymbol{a}_2)'\boldsymbol{V})1} - 2\xi_{c(i|T_{12}(\boldsymbol{a}_2)'\boldsymbol{V})1}a_1\boldsymbol{t}_{11}'\boldsymbol{V} + \|\boldsymbol{\xi}_{c(i|T_{12}(\boldsymbol{a}_2)'\boldsymbol{V})2}\|^2$$
$$-2\boldsymbol{\xi}'_{c(i|T_{12}(\boldsymbol{a}_2)'\boldsymbol{V})2}T_{12}(\boldsymbol{a}_2)'\boldsymbol{V}) \,\Bigg|\, \boldsymbol{Y} = \boldsymbol{0}_p\Bigg]$$

となる．ただし，$I(\cdot)$ は指示関数である．$\boldsymbol{V}^* \equiv \boldsymbol{V}|(\boldsymbol{Y} = \boldsymbol{0}_p)$ とおくと，

$$L_2(\boldsymbol{a}) = E\Bigg[\sum_{i=1}^{n(T_{12}(\boldsymbol{a}_2)'\boldsymbol{V}^*)} I(h(T_{12}(\boldsymbol{a}_2)'\boldsymbol{V}^*|i) < a_1\boldsymbol{t}_{11}'\boldsymbol{V}^* \le h(T_{12}(\boldsymbol{a}_2)'\boldsymbol{V}^*|i-1))$$
$$(\xi^2_{c(i|T_{12}(\boldsymbol{a}_2)'\boldsymbol{V}^*)1} - 2\xi_{c(i|T_{12}(\boldsymbol{a}_2)'\boldsymbol{V}^*)1}a_1\boldsymbol{t}_{11}'\boldsymbol{V}^* + \|\boldsymbol{\xi}_{c(i|T_{12}(\boldsymbol{a}_2)'\boldsymbol{V}^*)2}\|^2$$
$$-2\boldsymbol{\xi}'_{c(i|T_{12}(\boldsymbol{a}_2)'\boldsymbol{V}^*)2}T_{12}(\boldsymbol{a}_2)'\boldsymbol{V}^*)\Bigg]$$

と書き直すことができる．表記を簡単にするため，適宜 $Z_1 = \boldsymbol{t}_{11}'\boldsymbol{V}^*$，$\boldsymbol{Z}_2 = T_{12}(\boldsymbol{a}_2)'\boldsymbol{V}^*$，$A_i = \xi^2_{c(i|T_{12}(\boldsymbol{a}_2)'\boldsymbol{V}^*)1} - 2\xi_{c(i|T_{12}(\boldsymbol{a}_2)'\boldsymbol{V}^*)1}a_1\boldsymbol{t}_{11}'\boldsymbol{V}^* + \|\boldsymbol{\xi}_{c(i|T_{12}(\boldsymbol{a}_2)'\boldsymbol{V}^*)2}\|^2 - 2\boldsymbol{\xi}'_{c(i|T_{12}(\boldsymbol{a}_2)'\boldsymbol{V}^*)2}T_{12}(\boldsymbol{a}_2)'\boldsymbol{V}^*$ とおいて議論する．混同の危険がない限り，$\xi_{c(i|T_{12}(\boldsymbol{a}_2)'\boldsymbol{V}^*)1}$ を単に ξ_{i1}，$\boldsymbol{\xi}_{c(i|T_{12}(\boldsymbol{a}_2)'\boldsymbol{V}^*)2}$ を単に $\boldsymbol{\xi}_{i2}$ と表記する．また，$h(\boldsymbol{Z}_2|k) < 0 \le h(\boldsymbol{Z}_2|k-1)$ となる k を $k(\boldsymbol{Z}_2)$ と表記する．まず，$a_1 > 0$ の場合を議論する．

$$L_2(\boldsymbol{a}) = E\Bigg[\sum_{i=1}^{k(\boldsymbol{Z}_2)-1} A_i I\left(\frac{h(\boldsymbol{Z}_2|i)}{a_1} < Z_1 \le \frac{h(\boldsymbol{Z}_2|i-1)}{a_1}\right)$$
$$+A_{k(\boldsymbol{Z}_2)}\left\{I\left(0 < Z_1 \le \frac{h(\boldsymbol{Z}_2|k(\boldsymbol{Z}_2)-1)}{a_1}\right) + I\left(\frac{h(\boldsymbol{Z}_2|k(\boldsymbol{Z}_2))}{a_1} < Z_1 \le 0\right)\right\}$$
$$+\sum_{i=k(\boldsymbol{Z}_2)+1}^{n(\boldsymbol{Z}_2)} A_i I\left(\frac{h(\boldsymbol{Z}_2|i)}{a_1} < Z_1 \le \frac{h(\boldsymbol{Z}_2|i-1)}{a_1}\right)\Bigg]$$

となる．十分小さい $u>0$ に対し，

$$
\begin{aligned}
L_2(a_1+u,a_2,\ldots,a_d) = E\Bigg[&\sum_{i=1}^{k(\boldsymbol{Z}_2)-1}(A_i-2\xi_{i1}uZ_1)I\left(\frac{h(\boldsymbol{Z}_2|i)}{a_1+u}<Z_1\le\frac{h(\boldsymbol{Z}_2|i-1)}{a_1+u}\right)\\
&+(A_{k(\boldsymbol{Z}_2)}-2\xi_{k(\boldsymbol{Z}_2)1}uZ_1)\left\{I\left(0<Z_1\le\frac{h(\boldsymbol{Z}_2|k(\boldsymbol{Z}_2)-1)}{a_1+u}\right)\right.\\
&\left.+I\left(\frac{h(\boldsymbol{Z}_2|k(\boldsymbol{Z}_2))}{a_1+u}<Z_1\le 0\right)\right\}\\
&+\sum_{i=k(\boldsymbol{Z}_2)+1}^{n(\boldsymbol{Z}_2)}(A_i-2\xi_{i1}uZ_1)I\left(\frac{h(\boldsymbol{Z}_2|i)}{a_1+u}<Z_1\le\frac{h(\boldsymbol{Z}_2|i-1)}{a_1+u}\right)\Bigg]\\
= E\Bigg[&\sum_{i=1}^{k(\boldsymbol{Z}_2)-1}(A_i-2\xi_{i1}uZ_1)\left\{I\left(\frac{h(\boldsymbol{Z}_2|i)}{a_1+u}<Z_1\le\frac{h(\boldsymbol{Z}_2|i)}{a_1}\right)\right.\\
&\left.+I\left(\frac{h(\boldsymbol{Z}_2|i)}{a_1}<Z_1\le\frac{h(\boldsymbol{Z}_2|i-1)}{a_1+u}\right)\right\}\\
&+(A_{k(\boldsymbol{Z}_2)}-2\xi_{k(\boldsymbol{Z}_2)1}uZ_1)\left\{I\left(0<Z_1\le\frac{h(\boldsymbol{Z}_2|k(\boldsymbol{Z}_2)-1)}{a_1+u}\right)\right.\\
&\left.+I\left(\frac{h(\boldsymbol{Z}_2|k(\boldsymbol{Z}_2))}{a_1+u}<Z_1\le 0\right)\right\}\\
&+\sum_{i=k(\boldsymbol{Z}_2)+1}^{n(\boldsymbol{Z}_2)}(A_i-2\xi_{i1}uZ_1)\left\{I\left(\frac{h(\boldsymbol{Z}_2|i)}{a_1+u}<Z_1\le\frac{h(\boldsymbol{Z}_2|i-1)}{a_1}\right)\right.\\
&\left.+I\left(\frac{h(\boldsymbol{Z}_2|i-1)}{a_1}<Z_1\le\frac{h(\boldsymbol{Z}_2|i-1)}{a_1+u}\right)\right\}\Bigg]
\end{aligned}
$$

となり，合わせて $L_2(\boldsymbol{a})$ も

$$
\begin{aligned}
&L_2(a_1,a_2,\ldots,a_d)\\
&= E\Bigg[\sum_{i=1}^{k(\boldsymbol{Z}_2)-1}A_i\left\{I\left(\frac{h(\boldsymbol{Z}_2|i)}{a_1}<Z_1\le\frac{h(\boldsymbol{Z}_2|i-1)}{a_1+u}\right)\right.\\
&\left.+I\left(\frac{h(\boldsymbol{Z}_2|i-1)}{a_1+u}<Z_1\le\frac{h(\boldsymbol{Z}_2|i-1)}{a_1}\right)\right\}\\
&+A_{k(\boldsymbol{Z}_2)}\left\{I\left(0<Z_1\le\frac{h(\boldsymbol{Z}_2|k(\boldsymbol{Z}_2)-1)}{a_1+u}\right)\right.\\
&\left.+I\left(\frac{h(\boldsymbol{Z}_2|k(\boldsymbol{Z}_2)-1)}{a_1+u}<Z_1\le\frac{h(\boldsymbol{Z}_2|k(\boldsymbol{Z}_2)-1)}{a_1}\right)\right\}\\
&+A_{k(\boldsymbol{Z}_2)}\left\{I\left(\frac{h(\boldsymbol{Z}_2|k(\boldsymbol{Z}_2))}{a_1}<Z_1\le\frac{h(\boldsymbol{Z}_2|k(\boldsymbol{Z}_2))}{a_1+u}\right)+I\left(\frac{h(\boldsymbol{Z}_2|k(\boldsymbol{Z}_2))}{a_1+u}<Z_1\le 0\right)\right\}\\
&+\sum_{i=k(\boldsymbol{Z}_2)+1}^{n(\boldsymbol{Z}_2)}A_i\left\{I\left(\frac{h(\boldsymbol{Z}_2|i)}{a_1}<Z_1\le\frac{h(\boldsymbol{Z}_2|i)}{a_1+u}\right)+I\left(\frac{h(\boldsymbol{Z}_2|i)}{a_1+u}<Z_1\le\frac{h(\boldsymbol{Z}_2|i-1)}{a_1}\right)\right\}\Bigg]
\end{aligned}
$$

と書き直すことができることから,

$$
\begin{aligned}
&L_2(a_1+u, a_2, \ldots, a_d) - L_2(a_1, a_2, \ldots, a_d) \\
&= E\left[\sum_{i=1}^{k(\boldsymbol{Z}_2)-1} (-2\xi_{i1}uZ_1) I\left(\frac{h(\boldsymbol{Z}_2|i)}{a_1} < Z_1 \le \frac{h(\boldsymbol{Z}_2|i-1)}{a_1+u}\right)\right. \\
&\quad + \sum_{i=1}^{k(\boldsymbol{Z}_2)-1} (A_i - 2\xi_{i1}uZ_1 - A_{i+1}) I\left(\frac{h(\boldsymbol{Z}_2|i)}{a_1+u} < Z_1 \le \frac{h(\boldsymbol{Z}_2|i)}{a_1}\right) \\
&\quad + (-2\xi_{k(\boldsymbol{Z}_2)1}uZ_1)\left\{I\left(0 < Z_1 \le \frac{h(\boldsymbol{Z}_2|k(\boldsymbol{Z}_2)-1)}{a_1+u}\right) + I\left(\frac{h(\boldsymbol{Z}_2|k(\boldsymbol{Z}_2))}{a_1+u} < Z_1 \le 0\right)\right\} \\
&\quad + \sum_{i=k(\boldsymbol{Z}_2)+1}^{n(\boldsymbol{Z}_2)} (-2\xi_{i1}uZ_1) I\left(\frac{h(\boldsymbol{Z}_2|i)}{a_1+u} < Z_1 \le \frac{h(\boldsymbol{Z}_2|i-1)}{a_1}\right) \\
&\quad \left. + \sum_{i=k(\boldsymbol{Z}_2)+1}^{n(\boldsymbol{Z}_2)} (A_i - 2\xi_{i1}uZ_1 - A_{i-1}) I\left(\frac{h(\boldsymbol{Z}_2|i-1)}{a_1} < Z_1 \le \frac{h(\boldsymbol{Z}_2|i-1)}{a_1+u}\right)\right] \\
&= E\left[\sum_{i=1}^{n(\boldsymbol{Z}_2)} (-2\xi_{i1}uZ_1) I\left(\frac{h(\boldsymbol{Z}_2|i)}{a_1+u} < Z_1 \le \frac{h(\boldsymbol{Z}_2|i-1)}{a_1+u}\right)\right. \\
&\quad + \sum_{i=1}^{k(\boldsymbol{Z}_2)-1} (A_i - A_{i+1}) I\left(\frac{h(\boldsymbol{Z}_2|i)}{a_1+u} < Z_1 \le \frac{h(\boldsymbol{Z}_2|i)}{a_1}\right) \\
&\quad \left. - \sum_{i=k(\boldsymbol{Z}_2)}^{n(\boldsymbol{Z}_2)-1} (A_i - A_{i+1}) I\left(\frac{h(\boldsymbol{Z}_2|i)}{a_1} < Z_1 \le \frac{h(\boldsymbol{Z}_2|i)}{a_1+u}\right)\right]
\end{aligned} \tag{4.6}
$$

となる．ここで,

$$
\begin{aligned}
&A_i - A_{i+1} \\
&= (\xi_{i1}^2 - 2\xi_{i1}a_1Z_1 + \|\boldsymbol{\xi}_{i2}\|^2 - 2\boldsymbol{\xi}_{i2}'\boldsymbol{Z}_2) - (\xi_{i+11}^2 - 2\xi_{i+11}a_1Z_1 + \|\boldsymbol{\xi}_{i+12}\|^2 - 2\boldsymbol{\xi}_{i+12}'\boldsymbol{Z}_2) \\
&= -2(\xi_{i1} - \xi_{i+11})(aZ_1 - h(\boldsymbol{Z}_2|i)),\ i = 1, \ldots, n(\boldsymbol{Z}_2) - 1
\end{aligned}
$$

より，(4.6) の第 2 項と第 3 項の絶対値を取ると,

$$
\begin{aligned}
&\left| - \sum_{i=1}^{k(\boldsymbol{Z}_2)-1} 2(\xi_{i1} - \xi_{i+11})(aZ_1 - h(\boldsymbol{Z}_2|i)) I\left(\frac{h(\boldsymbol{Z}_2|i)}{a_1+u} < Z_1 \le \frac{h(\boldsymbol{Z}_2|i)}{a_1}\right)\right. \\
&\quad \left. + \sum_{i=k(\boldsymbol{Z}_2)}^{n(\boldsymbol{Z}_2)-1} 2(\xi_{i1} - \xi_{i+11})(aZ_1 - h(\boldsymbol{Z}_2|i)) I\left(\frac{h(\boldsymbol{Z}_2|i)}{a_1} < Z_1 \le \frac{h(\boldsymbol{Z}_2|i)}{a_1+u}\right)\right| \\
&\le \sum_{i=1}^{k(\boldsymbol{Z}_2)-1} 2u(\xi_{i1} - \xi_{i+11}) \left|\frac{h(\boldsymbol{Z}_2|i)}{a+u}\right| I\left(\frac{h(\boldsymbol{Z}_2|i)}{a_1+u} < Z_1 \le \frac{h(\boldsymbol{Z}_2|i)}{a_1}\right) \\
&\quad + \sum_{i=k(\boldsymbol{Z}_2)}^{n(\boldsymbol{Z}_2)-1} 2u(\xi_{i1} - \xi_{i+11}) \left|\frac{h(\boldsymbol{Z}_2|i)}{a+u}\right| I\left(\frac{h(\boldsymbol{Z}_2|i)}{a_1} < Z_1 \le \frac{h(\boldsymbol{Z}_2|i)}{a_1+u}\right)
\end{aligned} \tag{4.7}
$$

となり，(4.7) を u で割って $u \to +0$ とすると 0 になる．十分 0 に近い $u < 0$ に対しても同様に展開することができ，さらに $a_1 < 0$ および $a_1 = 0$ の場合も同様に議論することができることから，

$$L_2'(a_1, \ldots, a_d) \equiv \frac{\partial L_2(\boldsymbol{a})}{\partial a_1} = \lim_{u \to 0} \frac{L_2(a_1+u, a_2, \ldots, a_d) - L_2(a_1, a_2, \ldots, a_d)}{u}$$
$$= E\left[\sum_{i=1}^{n(\boldsymbol{Z}_2)} (-2\xi_{i1} Z_1) I\left(h(\boldsymbol{Z}_2|i) < a_1 Z_1 \le h(\boldsymbol{Z}_2|i-1)\right)\right]$$

となる．ここで再び $a_1 > 0$ として議論を進めると，

$$\begin{aligned}
&L_2'(a_1, \ldots, a_d)\\
&= E\left[\sum_{i=1}^{k(\boldsymbol{Z}_2)-1} (-2\xi_{i1} Z_1) I\left(\frac{h(\boldsymbol{Z}_2|i)}{a_1} < Z_1 \le \frac{h(\boldsymbol{Z}_2|i-1)}{a_1}\right)\right.\\
&\quad +(-2\xi_{k(\boldsymbol{Z}_2)1} Z_1)\left\{I\left(0 < Z_1 \le \frac{h(\boldsymbol{Z}_2|k(\boldsymbol{Z}_2)-1)}{a_1}\right) + I\left(\frac{h(\boldsymbol{Z}_2|k(\boldsymbol{Z}_2))}{a_1} < Z_1 \le 0\right)\right\}\\
&\quad \left.+\sum_{i=k(\boldsymbol{Z}_2)+1}^{n(\boldsymbol{Z}_2)} (-2\xi_{i1} Z_1) I\left(\frac{h(\boldsymbol{Z}_2|i)}{a_1} < Z_1 \le \frac{h(\boldsymbol{Z}_2|i-1)}{a_1}\right)\right]
\end{aligned}$$

となる．さらに，十分小さい $u > 0$ に対して，

$$\begin{aligned}
&L_2'(a_1+u, a_2, \ldots, a_d)\\
&= E\left[\sum_{i=1}^{k(\boldsymbol{Z}_2)-1} (-2\xi_{i1} Z_1) I\left(\frac{h(\boldsymbol{Z}_2|i)}{a_1+u} < Z_1 \le \frac{h(\boldsymbol{Z}_2|i-1)}{a_1+u}\right)\right.\\
&\quad +(-2\xi_{k(\boldsymbol{Z}_2)1} Z_1)\left\{I\left(0 < Z_1 \le \frac{h(\boldsymbol{Z}_2|k(\boldsymbol{Z}_2)-1)}{a_1+u}\right) + I\left(\frac{h(\boldsymbol{Z}_2|k(\boldsymbol{Z}_2))}{a_1+u} < Z_1 \le 0\right)\right\}\\
&\quad \left.+\sum_{i=k(\boldsymbol{Z}_2)+1}^{n(\boldsymbol{Z}_2)} (-2\xi_{i1} Z_1) I\left(\frac{h(\boldsymbol{Z}_2|i)}{a_1+u} < Z_1 \le \frac{h(\boldsymbol{Z}_2|i-1)}{a_1+u}\right)\right]\\
&= E\left[\sum_{i=1}^{k(\boldsymbol{Z}_2)-1} (-2\xi_{i1} Z_1)\left\{I\left(\frac{h(\boldsymbol{Z}_2|i)}{a_1+u} < Z_1 \le \frac{h(\boldsymbol{Z}_2|i)}{a_1}\right)\right.\right.\\
&\quad \left.+I\left(\frac{h(\boldsymbol{Z}_2|i)}{a_1} < Z_1 \le \frac{h(\boldsymbol{Z}_2|i-1)}{a_1+u}\right)\right\}\\
&\quad +(-2\xi_{k(\boldsymbol{Z}_2)1} Z_1)\left\{I\left(0 < Z_1 \le \frac{h(\boldsymbol{Z}_2|k(\boldsymbol{Z}_2)-1)}{a_1+u}\right) + I\left(\frac{h(\boldsymbol{Z}_2|k(\boldsymbol{Z}_2))}{a_1+u} < Z_1 \le 0\right)\right\}\\
&\quad +\sum_{i=k(\boldsymbol{Z}_2)+1}^{n(\boldsymbol{Z}_2)} (-2\xi_{i1} Z_1)\left\{I\left(\frac{h(\boldsymbol{Z}_2|i)}{a_1+u} < Z_1 \le \frac{h(\boldsymbol{Z}_2|i-1)}{a_1}\right)\right.\\
&\quad \left.\left.+I\left(\frac{h(\boldsymbol{Z}_2|i-1)}{a_1} < Z_1 \le \frac{h(\boldsymbol{Z}_2|i-1)}{a_1+u}\right)\right\}\right]
\end{aligned}$$

となり，合わせて $L_2'(a_1, a_2, \ldots, a_d)$ も

$$
\begin{aligned}
&L_2'(a_1, a_2, \ldots, a_d) \\
&= E\left[\sum_{i=1}^{k(\boldsymbol{Z}_2)-1} (-2\xi_{i1}Z_1)\left\{I\left(\frac{h(\boldsymbol{Z}_2|i)}{a_1} < Z_1 \leq \frac{h(\boldsymbol{Z}_2|i-1)}{a_1+u}\right)\right.\right. \\
&\quad \left.+I\left(\frac{h(\boldsymbol{Z}_2|i-1)}{a_1+u} < Z_1 \leq \frac{h(\boldsymbol{Z}_2|i-1)}{a_1}\right)\right\} \\
&\quad +(-2\xi_{k(\boldsymbol{Z}_2)1}Z_1)\left\{I\left(0 < Z_1 \leq \frac{h(\boldsymbol{Z}_2|k(\boldsymbol{Z}_2)-1)}{a_1+u}\right)\right. \\
&\quad +I\left(\frac{h(\boldsymbol{Z}_2|k(\boldsymbol{Z}_2)-1)}{a_1+u} < Z_1 \leq \frac{h(\boldsymbol{Z}_2|k(\boldsymbol{Z}_2)-1)}{a_1}\right) \\
&\quad \left.+I\left(\frac{h(\boldsymbol{Z}_2|k(\boldsymbol{Z}_2))}{a_1} < Z_1 \leq \frac{h(\boldsymbol{Z}_2|k(\boldsymbol{Z}_2))}{a_1+u}\right) + I\left(\frac{h(\boldsymbol{Z}_2|k(\boldsymbol{Z}_2))}{a_1+u} < Z_1 \leq 0\right)\right\} \\
&\quad + \sum_{i=k(\boldsymbol{Z}_2)+1}^{n(\boldsymbol{Z}_2)} (-2\xi_{i1}Z_1)\left\{I\left(\frac{h(\boldsymbol{Z}_2|i)}{a_1} < Z_1 \leq \frac{h(\boldsymbol{Z}_2|i)}{a_1+u}\right)\right. \\
&\quad \left.\left.+I\left(\frac{h(\boldsymbol{Z}_2|i)}{a_1+u} < Z_1 \leq \frac{h(\boldsymbol{Z}_2|i-1)}{a_1}\right)\right\}\right]
\end{aligned}
$$

と書き直すことができることから，

$$
\begin{aligned}
&L_2'(a_1+u, a_2, \ldots, a_d) - L_2'(a_1, a_2, \ldots, a_d) \\
&= E\left[\sum_{i=1}^{k(\boldsymbol{Z}_2)-1} (-2\xi_{i1}Z_1 + 2\xi_{i+11}Z_1) I\left(\frac{h(\boldsymbol{Z}_2|i)}{a_1+u} < Z_1 \leq \frac{h(\boldsymbol{Z}_2|i)}{a_1}\right)\right. \\
&\quad \left.+ \sum_{i=k(\boldsymbol{Z}_2)+1}^{n(\boldsymbol{Z}_2)} (-2\xi_{i1}Z_1 + 2\xi_{i-11}Z_1) I\left(\frac{h(\boldsymbol{Z}_2|i-1)}{a_1} < Z_1 \leq \frac{h(\boldsymbol{Z}_2|i-1)}{a_1+u}\right)\right] \\
&= E\left[-2Z_1 \sum_{i=1}^{k(\boldsymbol{Z}_2)-1} (\xi_{i1} - \xi_{i+11}) I\left(\frac{h(\boldsymbol{Z}_2|i)}{a_1+u} < Z_1 \leq \frac{h(\boldsymbol{Z}_2|i)}{a_1}\right)\right. \\
&\quad \left.+2Z_1 \sum_{i=k(\boldsymbol{Z}_2)}^{n(\boldsymbol{Z}_2)-1} (\xi_{i1} - \xi_{i+11}) I\left(\frac{h(\boldsymbol{Z}_2|i)}{a_1} < Z_1 \leq \frac{h(\boldsymbol{Z}_2|i)}{a_1+u}\right)\right]
\end{aligned}
$$

となる．ここで，簡潔にしていた表記を戻し，$h(\boldsymbol{Z}_2|k) < 0 \leq h(\boldsymbol{Z}_2|k-1)$ となる k を $k(\boldsymbol{Z}_2)$ としていたこと（すなわち，$h(T_{12}(\boldsymbol{a}_2)'\boldsymbol{V}^\star|k) < 0 \leq h(T_{12}(\boldsymbol{a}_2)'\boldsymbol{V}^\star|k-1)$ となる k

を $k(T_{12}(\boldsymbol{a}_2)'\boldsymbol{V}^\star)$ としていたこと）に注意すれば

$$\begin{aligned}
&L_2'(a_1+u, a_2, \ldots, a_d) - L_2'(a_1, a_2, \ldots, a_d) \\
&= E\Bigg[-2\boldsymbol{t}_{11}'\boldsymbol{V}^\star \sum_{i=1}^{k(T_{12}(\boldsymbol{a}_2)'\boldsymbol{V}^\star)-1} (\xi_{c(i|T_{12}(\boldsymbol{a}_2)'\boldsymbol{V}^\star)1} - \xi_{c(i+1|T_{12}(\boldsymbol{a}_2)'\boldsymbol{V}^\star)1}) \\
&\quad I\left(\frac{h(T_{12}(\boldsymbol{a}_2)'\boldsymbol{V}^\star|i)}{a_1+u} < \boldsymbol{t}_{11}'\boldsymbol{V}^\star \le \frac{h(T_{12}(\boldsymbol{a}_2)'\boldsymbol{V}^\star|i)}{a_1}\right) \\
&\quad +2\boldsymbol{t}_{11}'\boldsymbol{V}^\star \sum_{i=k(T_{12}(\boldsymbol{a}_2)'\boldsymbol{V}^\star)}^{n(T_{12}(\boldsymbol{a}_2)'\boldsymbol{V}^\star)-1} (\xi_{c(i|T_{12}(\boldsymbol{a}_2)'\boldsymbol{V}^\star)1} - \xi_{c(i+1|T_{12}(\boldsymbol{a}_2)'\boldsymbol{V}^\star)1}) \\
&\quad I\left(\frac{h(T_{12}(\boldsymbol{a}_2)'\boldsymbol{V}^\star|i)}{a_1} < \boldsymbol{t}_{11}'\boldsymbol{V}^\star \le \frac{h(T_{12}(\boldsymbol{a}_2)'\boldsymbol{V}^\star|i)}{a_1+u}\right)\Bigg] \\
&\le 0
\end{aligned}$$

となる．$a_1 < 0$ および $a_1 = 0$ の場合も同様に議論することができ，従って $L_2(\boldsymbol{a})$ が a_1 について凹関数であることがわかる．さらに，$a_2, \ldots, a_d$ の場合も同様に議論することができることから，$L_2(\boldsymbol{a})$ が $a_1, \ldots, a_d$ のそれぞれについて凹関数であることが示された．

ここで，(4.4) を思い出せば，$L(\boldsymbol{a})$ が $a_1, \ldots, a_d$ のそれぞれについて凹関数であることがわかり，従って，(4.3) 式が成立する $\boldsymbol{e} \in \{-1, 1\}^d$ が存在することが示された．

最後に，$\boldsymbol{Y}|(\boldsymbol{V} = \boldsymbol{v})$ の一次元周辺分布の確率密度関数 $f_1(y|\boldsymbol{v})$ が全ての $\boldsymbol{v}$ の値に対して存在し，かつ (4.2) が成立する場合を議論する．このとき，$P(\boldsymbol{Y} = \boldsymbol{0}_p) = 0$ であることから，$L_1(\boldsymbol{a})$ の性質のみを考えればよい．$\boldsymbol{Y}|(\boldsymbol{V} = \boldsymbol{v})$ が球面対称分布に従うことから，(4.2) は

$$f_q(y_1, \ldots, y_q|\boldsymbol{v}) > 0 \text{ for any } (y_1, \ldots, y_q)' \in R^q$$

であることを意味しており，従って $\frac{\partial^2 L_1(\boldsymbol{a})}{\partial a_1^2} < 0$ となる．これは $L(\boldsymbol{a})$ が $a_1, \ldots, a_d$ のそれぞれについて強凹関数であることを意味している．従って，このとき (4.3) 式の不等式は強意となる．

以上より，補題が示された．

参考文献

Bali, J.L., Boente, G. (2009). Principal points and elliptical distributions from the multivariate setting to the functional case. *Statistics & Probability Letters*, **79**, 1858-1865.

Fang, K.T., Kotz, S., Ng, K.W. (1990). *Symmetric Multivariate and Related Distributions.* Chapman and Hall, London.

Flury, B. (1990). Principal points. *Biometrika*, **77**, 33-41.

Flury, B. (1993). Estimation of principal points. *Applied Statistics*, **42**, 139-151.

Graf, L., Luschgy, H. (2000). *Foundations of Quantization for Probability Distributions.* Springer, Berlin.

Gu, X.N., Mathew, T. (2001). Some characterizations of symmetric two-principal points. *Journal of Statistical Planning and Inference*, **98**, 29-37.

Kieffer, J.C. (1983). Uniqueness of locally optimal quantizer for log-concave density and convex error weighting function. *IEEE Transactions on Information Theory*, **29**, 42-47.

Kurata, H. (2008). On principal points for location mixtures of multivariate spherically symmetric distributions. *Journal of Statistical Planning and Inference*, **138**, 3405-3418.

Kurata, H., Qiu, D. (2011). Linear subspace spanned by principal points of a mixture of spherically symmetric distributions. *Communications in Statistics–Theory and Methods* (in press).

Li, L., Flury, B. (1995). Uniqueness of principal points for univariate distributions. *Statistics & Probability Letters*, **25**, 323-327.

Matsuura, S., Shinozaki, N. (2007). Optimal binning strategies under squared error loss in selective assembly with measurement error. *Communications in Statistics–Theory and Methods*, **36**, 2863-2876.

Matsuura, S., Shinozaki, N. (2010). Optimal binning strategies under squared error loss in selective assembly with a tolerance constraint. *Communications in Statistics–Theory and Methods*, **39**, 592-605.

Matsuura, S., Kurata, H. (2010). A principal subspace theorem for 2-principal points of general location mixtures of spherically symmetric distributions. *Statistics & Probability Letters*, **80**, 1863-1869.

Matsuura, S., Kurata, H. (2011). Principal points of a multivariate mixture distribution. *Journal of Multivariate Analysis*, **102**, 213-224.

Matsuura, S. (2011). Optimal partitioning of probability distributions under general convex loss functions in selective assembly. *Communications in Statistics–Theory and Methods*, **40**, 1545-1560.

Mease, D., Nair, V.N., Sudjianto, A. (2004). Selective assembly in manufacturing: statistical issues and optimal binning strategies. *Technometrics*, **46**, 165-175.

Mease, D., Nair, V.N. (2006). Unique optimal partitions of distributions and connections to hazard rates and stochastic ordering. *Statistica Sinica*, **16**, 1299-1312.

村木千恵, 大瀧慈, 水田正弘 (1998). 主要点解析法による極東夏期天気図の分類. 応用統計学, **27**, 17-31.

Pollard, D. (1981). Strong consistency of k-means clustering. *Annals of Statistics*, **9**, 135-140.

Pollard, D. (1982). A central limit theorem for k-means clustering. *Annals of Probability*, **10**, 919-926.

清水信夫, 水田正弘, 佐藤義治 (1998). Principal Points の性質について. 応用統計学, **27**, 1-16.

清水信夫, 水田正弘, 佐藤義治 (1999). Principal Points の対称性に関する定理について. 計算機統計学, **12**, 45-53.

清水信夫, 水田正弘 (2008). ランダム関数の関数主要点と関数クラスタリングについて. 計算機統計学, **21**, 1-13. I

Shimizu, N., Mizuta, M. (2008). Functional principal points and functional cluster analysis. In: Jain, L.C. et al. (Eds.), *Computational Intelligence Paradigms: Innovative Applications, Studies in Computational Intelligence*, **137**, 149-165.

Stampfer, E., Stadlober, E. (2002). Methods for estimating principal points. *Communications in Statistics–Simulation and Computation*, **31**, 261-277.

Tarpey, T. (1994). Two principal points of symmetric, strongly unimodal distributions. *Statistics & Probability Letters*, **20**, 253-257.

Tarpey, T., Li, L., Flury, B. (1995). Principal points and self-consistent points of elliptical distributions. *Annals of Statistics*, **23**, 103-112.

Tarpey, T. (1997). Estimating principal points of univariate distributions. *Journal of Applied Statistics*, **24**, 499-512.

Tarpey, T., Kinateder, K. (2003). Clustering functional data. *Journal of Classification*, **20**, 93-114.

Tarpey, T. (2007). A parametric k-means algorithm. *Computational Statistics*, **22**, 71-89.

Tarpey, T., Petkova, E. (2010). Principal point classification: applications to differentiating drug and placebo responses in longitudinal studies. *Journal of Statistical Planning and Inference*, **140**, 539-550.

Trushkin, A. (1982). Sufficient conditions for uniqueness of a locally optimal quantizer for a class of convex error weighting functions. *IEEE Transactions on Information Theory*, **28**, 187-198.

Trushkin, A. (1984). Monotony of Lloyd'd method II for log-concave density and convex error weighting function. *IEEE Transactions on Information Theory*, **30**, 380-383.

Yamamoto, W., Shinozaki, N. (2000a). On uniqueness of two principal points for univariate location mixtures. *Statistics & Probability Letters*, **46**, 33-42.

Yamamoto, W., Shinozaki, N. (2000b). Two principal points for multivariate location mixtures of spherically symmetric distributions. *Journal of the Japan Statistical Society*, **30**, 53-63.

Zoppe, A. (1995). Principal points of univariate continuous distributions. *Statistics and Computing*, **5**, 127-132.

Zoppe, A. (1997). On uniqueness and symmetry of self-consistent points of univariate continuous distributions. *Journal of Classification*, **14**, 147-158.