

高次元小標本における固有値の推定とその応用

筑波大学大学院・数理物質科学研究科 矢田 和善 (Kazuyoshi Yata)
Graduate School of Pure and Applied Sciences
University of Tsukuba

筑波大学・数学系 青嶋 誠 (Makoto Aoshima)
Institute of Mathematics
University of Tsukuba

1 はじめに

マイクロアレイデータや MRI データに見られるように、情報化の進展に伴い、データの次元数 d が標本数 n よりも遥かに大きな高次元小標本 (HDLSS) データが、解析対象になる場面が増えてきている。このような HDLSS データに対して従来の統計手法を用いると、次元の呪いによって解析が上手くいかない。そこで、HDLSS データのための新たな統計手法の開発が必要になってきている。

高次元データ解析を行う際、データは真には高次元でなく、むしろ高次元空間に埋め込まれていて、実際は、ずっと小さな次元をもった空間において要約できる、というコンセンサスがある。そこでは、できるだけ情報を損なうことなく、低次元空間への次元縮約を行うべく、様々な方法論が提案されている。最も一般的な手法の一つは、固有値の大きさに基づいて次元縮約する主成分分析 (PCA) であろう。実際、HDLSS データに対する PCA の研究が多数ある。次元数 d が標本数 n と同じ速さで増加する場合に、Baik et al. (2005), Johnstone (2001), Paul (2007) は正規性のもとで、Baik and Silverstein (2006) は非正規性のもとで、標本共分散行列の固有値の漸近的な挙動を研究している。これらの多くは、Johnstone (2001) によって導入された *spiked covariance model* について論じたものである。次元数 d が標本数 n を固定したままで増加する場合、これを *d-asymptotics* もしくは *HDLSS asymptotics* と呼び、Hall et al. (2005), Ahn et al. (2007), Yata and Aoshima (2008) 等によって、HDLSS データの幾何学的構造が研究されている。

Ahn et al. (2007) の主要結果には致命的な誤りがあるが、HDLSS データに対して（従来型の）PCA の使用には限界があることを、理論的に示した意味で興味深い。実際、Muller et al. (2008) は数値的な検証から新たな方法論の必要性を唱えており、また、Jung and Marron (2008) は *spiked covariance model* の場合にのみ（従来型の）PCA の使用が意味をもつことを理論的に示している。一方、Yata and Aoshima (2008) は HDLSS データのための新たなアプローチを考え、Ahn 等とは異なる方法で固有値の推定を行い、HDLSS データの有効次元数の推定を与え、*HDLSS asymptotics* に基づいて推定量の性質を論じ、さらには最大固有値の推定についても言及した。Yata and Aoshima (2008) が提案する推定

法を使えば、HDLSS データに対して *spiked covariance model* に限定される PCA の適用範囲を、上手く広げることができる。

本研究では、Ahn et al. (2007) や Jung and Marron (2008) が理論的に示した、HDLSS データに対して（従来型の）PCA では推定が困難になるパラメータ領域において、固有値と固有ベクトルの推定にチャレンジする。Yata and Aoshima (2008) の方法に基づいて、HDLSS データのための固有値と固有ベクトルの新しい推定法を提案し、推定の良さを理論とシミュレーションとで確認する。併せて、主成分スコアの推定についても論じたい。

2 問題の設定

平均が 0 の d 次元分布をもつ母集団から、 n 個のデータベクトルを無作為に抽出して、データ行列 $\mathbf{X}_{(d)} : d \times n = [\mathbf{x}_{1(d)}, \dots, \mathbf{x}_{n(d)}]$ を定義する。母共分散行列 Σ_d の固有値を $\lambda_{1(d)} \geq \dots \geq \lambda_{d(d)} \geq 0$ とし、適当な直交行列 $\mathbf{H}_d = [\mathbf{h}_{1(d)}, \dots, \mathbf{h}_{d(d)}]$ で $\Sigma_d = \mathbf{H}_d \Lambda_d \mathbf{H}_d^T$, $\Lambda_d = \text{diag}(\lambda_{1(d)}, \dots, \lambda_{d(d)})$ と分解する。ここで $\mathbf{Z}_{(d)} = \Lambda_d^{-1/2} \mathbf{H}_d^T \mathbf{X}_{(d)}$ を定義すると、これの $d \times n$ 個の成分は分散が 1 で互いに無相関になる。さらに、 $\mathbf{Z}_{(d)}^T = [\mathbf{z}_{1(d)}, \dots, \mathbf{z}_{d(d)}]$, $\mathbf{z}_{i(d)} = (z_{i1(d)}, \dots, z_{in(d)})^T$, $i = 1, \dots, d$ と書く。ここで、 $\mathbf{Z}_{(d)}$ の成分は、4 次モーメントが一様有界になることを仮定する。（これ以降、簡単のため、添え字の d は省略して記述することにする。）なお、Hall et al. (2005), Ahn et al. (2007), Muller et al. (2008), Jung and Marron (2008) 等は、母集団分布に正規性を仮定するか、もしくは、成分間の依存度に ρ -mixing 条件を仮定して理論を構築している。我々は、Yata and Aoshima (2008) の方法に基づいて、こういった制約には縛られないようにする。

いま、 Σ の固有値に次のモデルを仮定する。

$$\lambda_i = a_i d^{\alpha_i} \quad (i = 1, \dots, m), \quad \lambda_j = c_j \quad (j = m+1, \dots, d). \quad (2.1)$$

ここで、 $a_i (> 0)$, $c_j (\geq 0)$, $\alpha_i (\alpha_1 \geq \dots \geq \alpha_m > 1/4)$ は未知の実数、 m は未知の自然数とする。ただし、 $\lambda_1 > \dots > \lambda_m$ と仮定する。モデル (2.1) は、Ahn et al. (2007), Jung and Marron (2008) 等が仮定したモデルと同様であるが、あくまで我々の目的は、彼らが理論的に示した HDLSS データに対する PCA の適用条件（モデルの制約条件）を緩めることにある。つまり、先行研究で推定が困難とされたパラメータの状況下で新たな推定法を提案し、それが、先行研究で推定がなされた状況下にも有効である（推定に要する標本数がオーダーの意味で少なくなる）ような方法論の構築を目指している。

Ahn et al. (2007), Jung and Marron (2008) 等は、標本共分散行列 $\mathbf{S} = n^{-1} \mathbf{X} \mathbf{X}^T$ と同じ固有値を有する Dual な標本共分散行列 $\mathbf{S}_D = n^{-1} \mathbf{X}^T \mathbf{X}$ を用いて、高次元小標本における標本固有値と標本固有ベクトルの挙動を理論的に示している。こ

こで,

$$nS_D = Z^T \Lambda Z = \sum_{i=1}^d \lambda_i z_i z_i^T$$

と書ける. 彼等の理論には, 母集団分布に正規性の仮定か, もしくは, Hall et al. (2005) と同様の ρ -mixing 条件が必要になる. Sphericity に関する

$$\frac{\sum_{i=1}^d \lambda_i^2}{(\sum_{i=1}^d \lambda_i)^2} \rightarrow 0 \quad \text{as } d \rightarrow \infty \quad (2.2)$$

なる条件のもと, 彼等は

$$\left(\frac{n}{\sum_{i=1}^d \lambda_i} \right) S_D \xrightarrow{p} I_n \quad \text{as } d \rightarrow \infty$$

が成り立つとことを, *HDLSS asymptotics* に基づいて理論的に示している. このことは, 高次元小標本の下では, 標本固有値の間に差異が出てこないことを意味している. Yata and Aoshima (2008) は, 母集団分布に関する正規性の仮定や ρ -mixing 条件を使うことなしに, (2.2) におけるパラメータの状況下では標本固有値の間に十分な差異が検出できなくなることを, *HDLSS asymptotics* で理論的に示している.

実は, (2.1) のモデルは $\alpha_1 < 1$ かつ $\lambda_d > 0$ のとき, (2.2) を満たすモデルになり, そのとき, Ahn 等の S_D に基づく *HDLSS asymptotics* では適切に固有値を推定することが困難になる. 実際, 母集団分布が正規性を有する HDLSS データに対して, Muller et al. (2008) が膨大なシミュレーションによる考察を行い, そういった状況下での固有値推定の困難さを主張している. また, Jung and Marron (2008) は, ρ -mixing 条件を仮定した HDLSS データに対して, (2.2) の状況下での固有ベクトルの推定の強不一致性を理論的に示している.

次節では, まず, Ahn 等の S_D に基づく *HDLSS asymptotics* について, Jung and Marron (2008) では明らかにされなかった (2.2) の状況下での固有値の推定について理論的に言及する. その後で, Yata and Aoshima (2008) が提案したアプローチを用いて, (2.2) の状況下での固有値の推定を改良する. 4 節では, 3 節で提案した固有値の推定の良さを, シミュレーションによって数値的に評価する. 5 節では, 推定が困難とされた固有ベクトルについて, (2.2) の状況下での理論的な結果を与える. 最後に 6 節では, 主成分スコアの推定量に関して, 理論的な結果を与える. なお, 母集団分布に関する正規性の仮定や ρ -mixing 条件は, 特に課さないものとする.

3 固有値の推定

S_D の固有値について、次の定理が成り立つ.

定理 1 S_D の固有値を $\hat{\lambda}_1 \geq \dots \geq \hat{\lambda}_n$ とする. そのとき, $\hat{\lambda}_i$ ($i = 1, \dots, m$) について, 次の条件

- (i) $\alpha_i > 1$ ならば, $d \rightarrow \infty, n \rightarrow \infty$,
 - (ii) $\alpha_i \in (1/4, 1]$ ならば, $d \rightarrow \infty, d^{2-2\alpha_i}/n \rightarrow 0$
- のもと,

$$\frac{\hat{\lambda}_i}{\lambda_i} = 1 + o_p(1) \quad (3.1)$$

が成り立つ.

注意 1 母集団分布が正規性を有するとき, $\hat{\lambda}_i$ ($i = 1, \dots, m$) について, 定理 1 の (i) と次の条件

- (ii) $\alpha_i \in (1/3, 1]$ ならば, $d \rightarrow \infty, d^{1-\alpha_i}/n \rightarrow 0$,
 - (iii) $\alpha_i \in (1/4, 1/3]$ ならば, $d \rightarrow \infty, d^{2-4\alpha_i}/n \rightarrow 0$
- のもとで

$$\frac{\hat{\lambda}_i}{\lambda_i} \Rightarrow \frac{\chi_n^2}{n} \quad (3.2)$$

が主張できる. ここで, χ_n^2 は自由度 n の χ^2 分布に従う確率変数, $\Rightarrow$ は分布収束を意味している. さらに, (i)–(iii) のもと, (3.1) も主張できる.

注意 2 母集団分布が正規性を有するとき, Jung and Marron (2008) は, $\alpha_1 > \dots > \alpha_s > 1$ ($s < n$) なる λ_i ($i = 1, \dots, s$) に対して (3.2) を主張した. しかしながら, $\alpha_i \leq 1$ なる場合については, (3.2) を言及することはできていなかった.

注意 3 定理 1 に見られるように, S_D に基づく条件では, $\alpha_i \leq 1$ の場合には固有値の推定に多くの標本が必要になる. 実際, 推定誤差に関して基準を設ければ, Aoshima and Yata (2008) あるいは Yata (2008) で言及する漸近一致性を有する標本数を, Yata and Aoshima (2009) の方法を使って算出することも可能であろう. しかしながら, S_D に基づく理論には, これ以上, 深入りしない.

次に, Yata and Aoshima (2008) のアプローチを用いて, 固有値の推定を試みる. 標本を 2 つに分割して, 2 つの独立な $d \times n$ データ行列 X_1, X_2 を使って, $S^2 = n^{-2} X_1 X_1^T X_2 X_2^T$ を定義する. (ここでの n は, $n' = n/2$ を意味している.) いま, S^2 に対して Dual な 2 乗行列 $S_D^2 = n^{-2} X_1^T X_2 X_2^T X_1$ を考える. そのとき, S^2 と S_D^2 は同じ固有値を有する. ここで $Z_i = \Lambda^{-1/2} H^T X_i$, $i = 1, 2$ を定義する. いま, $Z_i^T = [z_{i1}, \dots, z_{id}]$, $z_{ij} = (z_{ij1}, \dots, z_{ijn})^T$, $i = 1, 2$; $j = 1, \dots, d$ と書

けば,

$$n^2 S_D^2 = \mathbf{Z}_1^T \Lambda \mathbf{Z}_2 \mathbf{Z}_2^T \Lambda \mathbf{Z}_1 = \left(\sum_{i=1}^d \lambda_i \mathbf{z}_{1i} \mathbf{z}_{2i}^T \right) \left(\sum_{i=1}^d \lambda_i \mathbf{z}_{2i} \mathbf{z}_{1i}^T \right)$$

と表される. Yata and Aoshima (2008) では, S_D^2 に基づく *HDLSS asymptotics* を考え, その理論を HDLSS データの有効次元数の推定に使い, さらには, 最大固有値の推定についても言及した.

ここでは, S_D^2 に基づく *HDLSS asymptotics* の理論を, 最大固有値に続くその他の固有値推定にも発展させる. そのとき, S_D^2 の固有値について, 次の定理が成り立つ.

定理 2 S_D^2 の固有値を $\tilde{\lambda}_1^2 \geq \dots \geq \tilde{\lambda}_n^2$ とする. ここで, $\tilde{\lambda}_i = \sqrt{\tilde{\lambda}_i^2}$ ($i = 1, \dots, m$) について, 次の条件

- (i) $\alpha_i > 1/2$ ならば, $d \rightarrow \infty, n \rightarrow \infty,$
- (ii) $\alpha_i \in (1/4, 1/2]$ ならば, $d \rightarrow \infty, d^{2-2\alpha_i}/n \rightarrow 0$

のもと,

$$\frac{\tilde{\lambda}_i}{\lambda_i} = 1 + o_p(1) \quad (3.3)$$

が成り立つ.

注意 4 母集団分布が正規性を有するとき, $\tilde{\lambda}_i^2$ ($i = 1, \dots, m$) について, 定理 2 の (i) と次の条件

- (ii) $\alpha_i \in (1/4, 1/2]$ ならば, $d \rightarrow \infty, d^{2-4\alpha_i}/n \rightarrow 0$

のもとで

$$\frac{\tilde{\lambda}_i^2}{\lambda_i^2} \Rightarrow \frac{\chi_{1n}^2 \chi_{2n}^2}{n^2}$$

が主張できる. ここで, χ_{1n}^2, χ_{2n}^2 は, 自由度 n の χ^2 分布にそれぞれ独立に従う確率変数である. さらに, (i)-(ii) のもと, (3.3) も主張できる.

4 シミュレーション

定理 2 に見られるように, S_D^2 に基づく条件では, $\alpha_i > 1/2$ の場合に, 固有値の推定に必要な標本数 n は (3.3) を主張する限りにおいて d に依存しない. これは, S_D に基づく定理 1 の条件が, $\alpha_i \in (1/2, 1]$ の場合に, 固有値を推定するための標本数 n が d のオーダーに依存したことで大きく性質を異にする. この理論上の結果を実際に応用すれば, S_D を使うよりも S_D^2 を使うことで, d と比べて n のオーダーが極端に小さな HDLSS データにおいても, 実用的な程度に固有値を推定できる可能性がある. 本節では, これをモンテカルロ・シミュレーションで確認する.

下の図 1-1 (第 1 固有値), 図 1-2 (第 2 固有値), 図 1-3 (第 3 固有値) は, $d = 1600$ 次元の正規乱数 $N_d(0, \Sigma)$ を生成して, 標本数 $n \in [40, 100]$ における A (定理 1) : $\hat{\lambda}_i/\lambda_i$, B (定理 2) : $\tilde{\lambda}_i/\lambda_i$ の値について, それぞれ 500 回のシミュレーション実験を行い, その平均値をプロットしたものである. ここでは, (2.1) のモデルにおいて, パラメータを $\lambda_1 = d^{5/6}$, $\lambda_2 = d^{2/3}$, $\lambda_3 = d^{1/2}$, $\lambda_4 = \dots = \lambda_d = 1$ と設定した. なお, B (定理 2) においては, A (定理 1) で用いた標本を 2 分割して, $d \times (n/2)$ の 2 つの行列 X_1, X_2 を使っている. A (定理 1) と B (定理 2) は, 推定に要する標本数を同じに設定して, その上で比較がなされている.

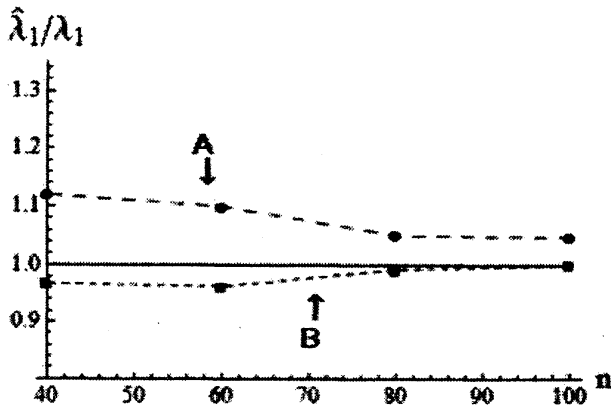


図 1-1. 第 1 固有値

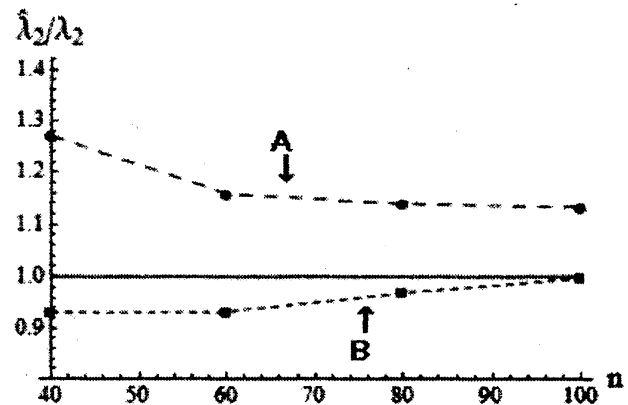


図 1-2. 第 2 固有値

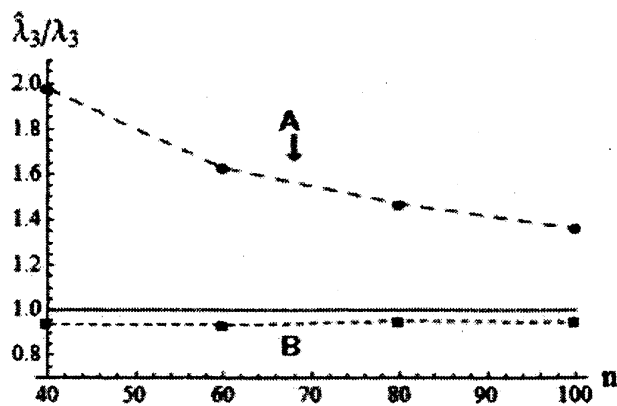


図 1-3. 第 3 固有値

これらの図から分かるように, 固有値の推定について, B (定理 2) で与えた S_D^2 に基づく手法の方が, 良い推定になっている. 特に, $\alpha_3 = 1/2$ なる第 3 固有値の推定に至っては, もはや A (定理 1) は全く推定できていないが, B (定理 2) では良く推定できている. さらに, 下の図 2-1 (第 1 固有値の分散), 図 2-2 (第 2 固有値の分散), 図 2-3 (第 3 固有値の分散) は, A (定理 1) : $\hat{\lambda}_i/\lambda_i$, B (定理 2) : $\tilde{\lambda}_i/\lambda_i$ の値について, 先ほどのシミュレーション実験における分散の値をプロットしたものである.

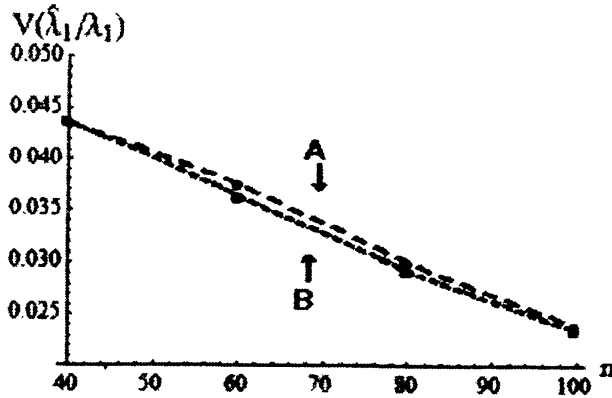


図 2-1. 第 1 固有値の分散

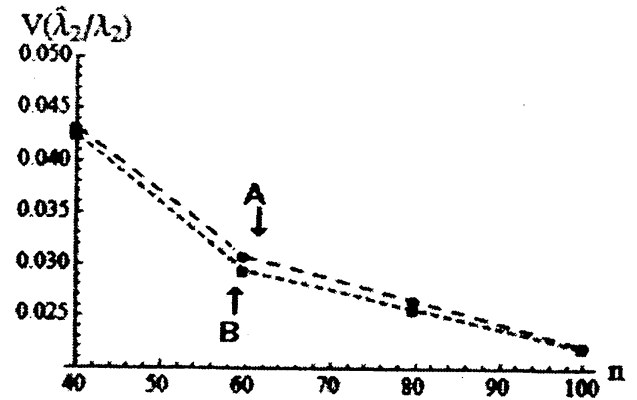


図 2-2. 第 2 固有値の分散

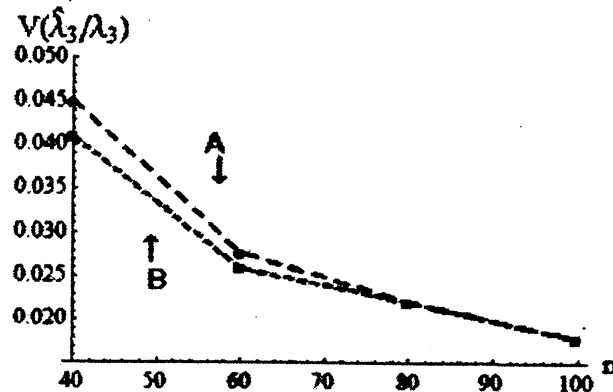


図 2-3. 第 3 固有値の分散

B (定理 2) で与えた S_D^2 に基づく推定では、標本を 2 分割して推定量を定義するので、 λ_i の推定が A (定理 1) と比べて不安定になるのではと危惧される。しかしながら、これらの図から分かるように、第 1 固有値から第 3 固有値の何れも、A (定理 1) と B (定理 2) による推定の分散は、ほぼ等しくなっている。注意 1 から、A (定理 1) における $\hat{\lambda}_i/\lambda_i$ の極限分布は χ_n^2/n であり、この分散は $V(\chi_n^2/n) = 2/n$ になる。一方、注意 4 から、 $\tilde{\lambda}_i^2/\lambda_i^2$ の極限分布は $\chi_{1n}^2\chi_{2n}^2/n^2$ であり、このことから、 $\tilde{\lambda}_i/\lambda_i$ の極限分布の分散は、おおよそ $1/n$ である。B (定理 2) では、標本を 2 分割して推定量を定義していたことを考慮すると、 $\tilde{\lambda}_i/\lambda_i$ の極限分布の分散は、おおよそ $2/n$ となり、これは、A (定理 1) における推定の極限分散と一致する。したがって、推定量の分散が、A (定理 1) と B (定理 2) とで漸近的に等しくなることは理論的に正しい結果であり、その安定した挙動がシミュレーションで確認されたということである。今回の結果を含め多くの実験結果から、HDLSS データの固有値に対して、 S_D^2 に基づく定理 2 の推定法を、本研究において推奨する。

5 固有ベクトルの推定

PCA による次元縮約を考える上で、固有値と共に固有ベクトルの推定を構築することは重要である。ここでは、Jung and Marron (2008) による固有ベクトル

推定の HDLSS asymptotics において、推定に強不一致性が結論付けられた状況を扱う。本研究で得られたアプローチを用いれば、標本数 n を d の適当なオーダーで定めることで、HDLSS データに対して固有ベクトルを有効に推定できることが期待される。

いま, Jung and Marron (2008) と同様に, S に対して $\hat{H}^T S \hat{H} = \hat{\Lambda}$, $\hat{\Lambda} = \text{diag}(\hat{\lambda}_1, \dots, \hat{\lambda}_d)$ となる直交行列を, $\hat{H} = [\hat{h}_1, \dots, \hat{h}_d]$ とする. Jung and Marron (2008) は ρ -mixing 条件を仮定して, (2.1) のモデルにおいて $\alpha_1 > \dots > \alpha_s > 1$ ($s < n$) ならば, $\hat{h}_i$ ($i = 1, \dots, s$) について

$$\text{Angle}(\hat{h}_i, h_i) \xrightarrow{p} 0 \quad \text{as } d \rightarrow \infty \quad (\text{一貫性}) \quad (5.1)$$

が主張でき, $\hat{h}_i$ ($i = s+1, \dots, n$) については, さらに

$$(a) \frac{\sum_{i=s+1}^d \lambda_i^2}{(\sum_{i=s+1}^d \lambda_i)^2} \rightarrow 0 \quad \text{as } d \rightarrow \infty,$$

$$(b) \sum_{i=s+1}^d \lambda_i = O(d)$$

なる条件を満足するとき

$$\text{Angle}(\hat{h}_i, h_i) \xrightarrow{p} \frac{\pi}{2} \quad \text{as } d \rightarrow \infty \quad (\text{強不一致性}) \quad (5.2)$$

が主張できることを示した. さらに, (2.1) のモデルにおいて $\alpha_i > 1$ を有する κ 個の固有値について, $\lambda_1 > \dots > \lambda_\kappa$ ならば, $d \rightarrow \infty, n \rightarrow \infty$ のもと, (5.1) を主張した.

(2.1) のモデルは, $\lambda_d > 0$ かつ $\alpha_{s+1} < 1$ のとき条件 (a)-(b) を満たすモデルとなるので, 標本数 n を固定したままでは, $\alpha_i < 1$ をもつ固有値 λ_i に対する固有ベクトル h_i は, 適切に推定できない. そこで, HDLSS データにおける標本数 n を, d の適当なオーダーで定めることを考える. そのとき, ρ -mixing 条件を課すことなしに, 次の結果を得る.

定理 3 $\hat{h}_i$ ($i = 1, \dots, m$) について, 次の条件

(i) $\alpha_i > 1$ ならば, $d \rightarrow \infty, n \rightarrow \infty$,

(ii) $\alpha_i \in (1/4, 1]$ ならば, $d \rightarrow \infty, d^{2-2\alpha_i}/n \rightarrow 0$

のもと,

$$\text{Angle}(\hat{h}_i, h_i) \xrightarrow{p} 0 \quad \text{as } d \rightarrow \infty \quad (\text{一貫性})$$

が成り立つ.

注意 5 いま, $S_D = \sum_{i=1}^n \hat{\lambda}_i \hat{u}_i \hat{u}_i^T$ と書ける. ただし, $\hat{u}_i$ ($i = 1, \dots, n$) は $\|\hat{u}_i\| = 1$ の S_D の固有ベクトルである. そのとき,

$$\hat{h}_i = (n\hat{\lambda}_i)^{-1/2} \mathbf{X} \hat{u}_i, \quad i = 1, \dots, n \quad (5.3)$$

が主張できる。つまり、 S の固有ベクトル $\hat{h}_i$ は S_D の固有値と固有ベクトルを使って求めることができる。よって、高次元小標本 ($d > n$) のもとでは、 $d \times d$ の行列 S よりも、ずっとサイズの小さい $n \times n$ の行列 S_D に基づいて $\hat{h}_i$ を求めることができることは本質的である。

6 主成分スコアの推定

前節で、HDLSS データに対する固有ベクトルの推定について論じた。さらに PCA を考えるうえで、主成分スコアの推定量を構築することとても重要である。いま、データ x_j の第 i 主成分スコアを $h_i^T x_j = \sqrt{\lambda_i} z_{ij}$ ($= s_{ij}$) と書く。固有ベクトル h_i は未知なので、(5.3) による固有ベクトル $\hat{h}_i$ を使い、主成分スコアを $\hat{h}_i^T x_j = \sqrt{n\hat{\lambda}_i} \hat{u}_{ij}$ ($= \hat{s}_{ij}$) で推定する。ここで、 $\hat{u}_i = (\hat{u}_{i1}, \dots, \hat{u}_{in})^T$ である。このとき、 $V(\hat{s}_i) = n^{-1} \sum_{j=1}^n (\hat{s}_{ij} - s_{ij})^2$ ($i = 1, \dots, m$) とおき、推定した主成分スコア $\hat{s}_{ij}$ と真の主成分スコア s_{ij} の標本誤差 $V(\hat{s}_i)$ について、次の定理が成り立つ。

定理 4 $V(\hat{s}_i)$ ($i = 1, \dots, m$) について、次の条件

- (i) $\alpha_i > 1$ ならば、 $d \rightarrow \infty, n \rightarrow \infty,$
- (ii) $\alpha_i \in (1/4, 1]$ ならば、 $d \rightarrow \infty, d^{2-2\alpha_i}/n \rightarrow 0$

のもと、

$$\frac{V(\hat{s}_i)}{\lambda_i} = o_p(1)$$

が成り立つ。

Appendix

Appendix を通して、

$$R_n = \{e_n \in R^n : \|e_n\| = 1\}$$

とし、 e_{1n}, e_{2n} を任意の元 ($e_{jn} \in R_n, j = 1, 2$) とする。また、

$$U_1 = n^{-1} \sum_{i=1}^m \lambda_i z_i z_i^T, \quad U_2 = n^{-1} \sum_{i=m+1}^d \lambda_i z_i z_i^T,$$

$$V_1 = n^{-1} \sum_{i=1}^m \lambda_i z_{1i} z_{2i}^T, \quad V_2 = n^{-1} \sum_{i=m+1}^d \lambda_i z_{1i} z_{2i}^T$$

とおく。いま、(2.1) における d の冪 α_i ($i = 1, \dots, m$) について、 $\alpha_1 = \dots = \alpha_{s_1} > \alpha_{s_1+1} = \dots = \alpha_{s_2} > \dots > \alpha_{s_{l-1}+1} = \dots = \alpha_{s_l} (= \alpha_m)$ とする。ただし、 $l \leq m$ で

ある。そのとき,

$$U_{1i} = n^{-1} \sum_{j=s_{i-1}+1}^{s_i} \lambda_j z_j z_j^T, \quad V_{1i} = n^{-1} \sum_{j=s_{i-1}+1}^{s_i} \lambda_j z_{1j} z_{2j}^T, \quad i = 1, \dots, l$$

とする。ただし, $s_0 = 0$ である。

補題 1 U_2, V_2 において, α_i ($i = 1, \dots, m$) に関して

(i) $\alpha_i > 1/2$ ならば, $d \rightarrow \infty, n$: 固定,

(ii) $\alpha_i \in (1/4, 1/2]$ ならば, $d \rightarrow \infty, d^{2-2\alpha_i}/n \rightarrow 0$

なる条件を満たすとき,

$$d^{-\alpha_i} e_{1n}^T U_2 e_{2n} = d^{-\alpha_i} e_{1n}^T \text{diag}(u_{11}, \dots, u_{nn}) e_{2n} + o_p(1), \quad d^{-\alpha_i} e_{1n}^T V_2 e_{2n} = o_p(1)$$

が成り立つ。ここで, u_{ii} は U_2 の対角成分である。

証明 まず, U_2 に関して, (i, j) 成分を u_{ij} とすれば, $u_{ij} = n^{-1} \sum_{i'=m+1}^d \lambda_{i'} z_{i'i} z_{i'j}$ ($i = 1, \dots, n; j = 1, \dots, n$) と書ける。いま, $U_{21} = U_2 - \text{diag}(u_{11}, \dots, u_{nn}) = [u_{21}, \dots, u_{2n}]$ とおく。ここで, $u_{2i} = (u_{1i}, \dots, u_{i-1i}, 0, u_{i+1i}, \dots, u_{ni})^T$ である。このとき, 条件 (i)-(ii) のもとで $\|d^{-\alpha_i} e_{1n}^T U_{21}\|^2 = o_p(1)$ を示す。いま, $e_{1n} = (e_{11}, \dots, e_{1n})^T, \sum_{i=1}^n e_{1i}^2 = 1$ とおく。そのとき,

$$\|d^{-\alpha_i} e_{1n}^T U_{21}\|^2 = d^{-2\alpha_i} \sum_{i=1}^n \sum_{i'=1(\setminus i)}^n e_{1i'}^2 u_{i'i}^2 + d^{-2\alpha_i} \sum_{i=1}^n \sum_{i' \neq j'(\setminus i)} e_{1i'} e_{1j'} u_{i'i} u_{j'i} \quad (\text{A.1})$$

と書ける。ただし, “ $\setminus i$ ” は添え字 i を除くことの意味である。

まず, (A.1) の第 2 項について, z_{ij} の 4 次のモーメントの上界を M とすると

$$\begin{aligned} E(u_{i'i} u_{j'i} u_{i'l} u_{j'l}) &= n^{-4} E \left(\sum_{j=m+1}^d \lambda_j^4 z_{ji}^2 z_{ji'}^2 z_{jj'}^2 z_{jl}^2 \right) \\ &= n^{-4} \sum_{j=m+1}^d \lambda_j^4 \quad (i \neq i'; i \neq j'; l \neq i'; l \neq j'; i \neq l; i' \neq j'), \end{aligned}$$

$$\begin{aligned} E(u_{i'i}^2 u_{j'i}^2) &= n^{-4} E \left\{ \left(\sum_{j=m+1}^d \lambda_j^2 z_{ji}^2 z_{ji'}^2 \right) \left(\sum_{j=m+1}^d \lambda_j^2 z_{ji}^2 z_{jj'}^2 \right) \right\} \\ &\leq n^{-4} M \left(\sum_{j=m+1}^d \lambda_j^2 \right)^2 \quad (i \neq i'; i \neq j'; i' \neq j') \end{aligned}$$

となるので, 条件 (i)-(ii) から, $u_{i'j'*} = d^{-2\alpha_i} \sum_{i=1(\setminus i', j')}^n u_{i'i} u_{j'i}$ について次を得る。

$$\begin{aligned} E \{ n^2 (u_{i'j'*})^2 \} &\leq d^{-4\alpha_i} \left(n^{-1} M \left(\sum_{j=m+1}^d \lambda_j^2 \right)^2 + \sum_{j=m+1}^d \lambda_j^4 \right) \\ &= O(d^{2-4\alpha_i}/n) + O(d^{1-4\alpha_i}) = o(1) \quad (i' \neq j'). \end{aligned}$$

よって、任意の $\tau > 0$ に対してマルコフの不等式より

$$P\left(\sum_{i' \neq j'} (u_{i'j'}^*)^2 > \tau\right) \leq \tau^{-1} E\{n^2 (u_{i'j'}^*)^2\} = o(1)$$

を得る。従って、

$$\|(0, u_{12*}, \dots, u_{1n*}, u_{21*}, 0, \dots, u_{2n*}, \dots, u_{n1*}, \dots, u_{nn-1*}, 0)\| = o_p(1)$$

が主張できる。ここで、 $\sum_{i' \neq j'} e_{1i'}^2 e_{1j'}^2 = O(1)$ から

$$\|(0, e_{11}e_{12}, \dots, e_{11}e_{1n}, e_{12}e_{11}, 0, \dots, e_{12}e_{1n}, \dots, e_{1n}e_{11}, \dots, e_{1n}e_{1n-1}, 0)\| = O(1)$$

なることに注意して、

$$d^{-2\alpha_i} \sum_{i=1}^n \sum_{i' \neq j'(\setminus i)} e_{1i'} e_{1j'} u_{i'i} u_{j'i} = \sum_{i' \neq j'} e_{1i'} e_{1j'} u_{i'j'}^* = o_p(1)$$

を得る。一方、(A.1) の第1項については、

$$\sum_{i=1}^n \sum_{i'=1(\setminus i)}^n e_{1i'}^2 u_{i'i}^2 = \sum_{i'=1}^n e_{1i'}^2 \sum_{i=1(\setminus i')}^n u_{i'i}^2$$

が主張できる。いま、任意の $\tau > 0$ に対してマルコフの不等式を用いる。条件 (i) のもとで

$$\sum_{i'=1}^n P\left(d^{-2\alpha_i} \sum_{i=1(\setminus i')}^n u_{i'i}^2 > \tau\right) \leq \tau^{-1} n^2 d^{-2\alpha_i} E(u_{i'i}^2) = O(d^{1-2\alpha_i}) = o(1)$$

となり、条件 (ii) のもとでは $E(u_{i'i}^4) \leq M(\sum_{s=m+1} \lambda_i)^4$ より

$$\begin{aligned} \sum_{i'=1}^n P\left(d^{-2\alpha_i} \sum_{i=1(\setminus i')}^n u_{i'i}^2 > \tau\right) &\leq \tau^{-2} n d^{-4\alpha_i} E\left\{\left(\sum_{i=1}^n u_{i'i}^2\right)^2\right\} \\ &\leq \tau^{-2} (d^{4-4\alpha_i}/n^2 + d^{2-4\alpha_i}/n) = o(1) \end{aligned} \quad (\text{A.2})$$

となる。よって、条件 (i)–(ii) のもとで $\sum_{i=1}^n \sum_{i'=1(\setminus i)}^n e_{1i'}^2 u_{i'i}^2 = o_p(1)$ を得る。それゆえ、(A.1) において $\|d^{-\alpha_i} e_{1n}^T U_{21}\|^2 = o_p(1)$ が成り立つ。従って、 $d^{-\alpha_i} e_{1n}^T U_{21} e_{2n} = o_p(1)$ が主張でき、 $d^{-\alpha_i} e_{1n}^T U_2 e_{2n} = d^{-\alpha_i} e_{1n}^T \text{diag}(u_{11}, \dots, u_{nn}) e_{2n} + o_p(1)$ が成り立つ。

次に V_2 について、 (i, j) 成分を v_{ij} とすると、 $v_{ij} = n^{-1} \sum_{i'=m+1}^d \lambda_{i'} z_{1i'i} z_{2i'j}$ ($i = 1, \dots, n; j = 1, \dots, n$) と書ける。このとき、 U_2 と同様にして、条件 (i)–(ii) のもとで $d^{-\alpha_i} e_{1n}^T V_2 e_{2n} = d^{-\alpha_i} e_{1n}^T \text{diag}(v_{11}, \dots, v_{nn}) e_{2n} + o_p(1)$ が成り立つ。ここで、任意の $\tau > 0$ に対してチェビシェフの不等式より

$$\begin{aligned} \sum_{i=1}^n P(d^{-\alpha_i} |v_{ii}| > \tau) &= n P\left(n^{-1} d^{-\alpha_i} \left| \sum_{i'=m+1}^d \lambda_{i'} z_{1i'i} z_{2i'i} \right| > \tau\right) \\ &\leq \tau^{-2} n^{-1} d^{-2\alpha_i} \left(\sum_{j=m+1}^d \lambda_i^2 \right) = O(d^{1-2\alpha_i}/n) = o(1) \end{aligned}$$

を得る. 従って, すべての $i (= 1, \dots, n)$ で $d^{-\alpha_i} v_{ii} = o_p(1)$ が主張でき, $d^{-\alpha_i} e_{1n}^T V_2 e_{2n} = o_p(1)$ が成立する. $\square$

注意 6 母集団分布が正規性を有するとき, $E(u_{i'i}^4) \leq 3(\sum_{s=m+1}^s \lambda_s^2)^2$ より, 注意 4 の条件 (ii) のもとで, (A.2) が主張できる. それゆえ, 補題 1 は注意 4 の条件 (ii) のもとで主張される.

補題 2 各 $i (= 1, \dots, l)$ について, U_{1i} のスペクトル分解を $U_{1i} = \sum_{j=1}^{s_i-s_{i-1}} \hat{\lambda}_{ij} \hat{u}_{ij} \hat{u}_{ij}^T$ とする. ここで, $\hat{\lambda}_{i1} \geq \dots \geq \hat{\lambda}_{is_i-s_{i-1}}$ は U_{1i} の固有値, $\hat{u}_{ij}$ ($j = 1, \dots, s_i - s_{i-1}$) $\in R_n$ は U_{1i} の固有ベクトルである. そのとき, $n \rightarrow \infty$ のもとで

$$\frac{\hat{\lambda}_{ij}}{\lambda_{s_{i-1}+j}} = 1 + o_p(1), \quad \hat{u}_{ij}^T n^{-1/2} z_{s_{i-1}+j} = 1 + o_p(1) \quad (j = 1, \dots, s_i - s_{i-1})$$

が成立する.

証明 まず, U_{11} について考察する. いま, $\lambda_1 > \dots > \lambda_{s_1}$ より, $a_1 > \dots > a_{s_1} > 0$ である. ここで, $n \rightarrow \infty$ のとき $\|n^{-1/2} z_i\| = 1 + o_p(1)$, $n^{-1} z_i^T z_j = o_p(1)$ ($i \neq j$) が主張できる. そのとき,

$$\frac{\hat{\lambda}_{11}}{\lambda_1} = \hat{u}_{11}^T \frac{U_{11}}{\lambda_1} \hat{u}_{11} = \hat{u}_{11}^T \left(\sum_{i=1}^{s_1} \frac{a_i}{a_1 n} z_i z_i^T \right) \hat{u}_{11} = 1 + o_p(1)$$

となり, $\hat{u}_{11}^T n^{-1/2} z_1 = 1 + o_p(1)$ を得る. いま, $\hat{u}_{12}^T n^{-1/2} z_1 = o_p(1)$ に注意すれば,

$$\frac{\hat{\lambda}_{12}}{\lambda_2} = \hat{u}_{12}^T \frac{U_{11}}{\lambda_2} \hat{u}_{12} = \hat{u}_{12}^T \left(\sum_{i=2}^{s_1} \frac{a_i}{a_2 n} z_i z_i^T \right) \hat{u}_{12} + o_p(1) = 1 + o_p(1)$$

となり, $\hat{u}_{12}^T n^{-1/2} z_2 = 1 + o_p(1)$ を得る. 以下同様にして,

$$\frac{\hat{\lambda}_{1j}}{\lambda_j} = 1 + o_p(1), \quad \hat{u}_{1j}^T n^{-1/2} z_j = 1 + o_p(1) \quad (j = 1, \dots, s_1)$$

を得る.

次に, U_{1i} ($i = 2, \dots, l$) に関しても同様に, $n \rightarrow \infty$ のとき

$$\frac{\hat{\lambda}_{ij}}{\lambda_{s_{i-1}+j}} = 1 + o_p(1), \quad \hat{u}_{ij}^T n^{-1/2} z_{s_{i-1}+j} = 1 + o_p(1) \quad (j = 1, \dots, s_i - s_{i-1})$$

を得ることができ, 結果が導かれる. $\square$

補題 3 各 $i (= 1, \dots, l)$ について, V_{1i} の特異値分解を $V_{1i} = \sum_{j=1}^{s_i-s_{i-1}} \tilde{\lambda}_{ij} \tilde{u}_{ij(1)} \tilde{u}_{ij(2)}^T$ とする. ここで, $\tilde{\lambda}_{i1} \geq \dots \geq \tilde{\lambda}_{is_i-s_{i-1}} > 0$ は特異値, $\tilde{u}_{ij(1)}$ ($j = 1, \dots, s_i - s_{i-1}$) $\in$

$\mathbf{R}_n$ は $\mathbf{V}_{1i} \mathbf{V}_{1i}^T$ の固有ベクトル, $\tilde{\mathbf{u}}_{ij(2)} (j = 1, \dots, s_i - s_{i-1}) \in \mathbf{R}_n$ は $\mathbf{V}_{1i}^T \mathbf{V}_{1i}$ の固有ベクトルである. そのとき, $n \rightarrow \infty$ のもとで

$$\begin{aligned} \frac{\tilde{\lambda}_{ij}}{\lambda_{s_{i-1}+j}} &= 1 + o_p(1), \\ \tilde{\mathbf{u}}_{ij(1)}^T n^{-1/2} \mathbf{z}_{1s_{i-1}+j} &= 1 + o_p(1), \quad \tilde{\mathbf{u}}_{ij(2)}^T n^{-1/2} \mathbf{z}_{2s_{i-1}+j} = 1 + o_p(1) \\ (j &= 1, \dots, s_i - s_{i-1}) \end{aligned}$$

が成立する.

証明 $\mathbf{V}_{1i} (i = 1, \dots, l)$ について, 任意の $\mathbf{e}_{1n}, \mathbf{e}_{2n} \in \mathbf{R}_n$ に対して

$$\max(\mathbf{e}_{1n}^T \mathbf{V}_{1i} \mathbf{e}_{2n}) = \tilde{\mathbf{u}}_{i1(1)}^T \mathbf{V}_{1i} \tilde{\mathbf{u}}_{i1(2)} = \tilde{\lambda}_{i1}$$

となり, さらに $\tilde{\mathbf{u}}_{i1(1)}^T \mathbf{e}_{1n} = 0, \tilde{\mathbf{u}}_{i1(2)}^T \mathbf{e}_{2n} = 0$ なる任意の $\mathbf{e}_{1n}, \mathbf{e}_{2n} \in \mathbf{R}_n$ に対して

$$\max(\mathbf{e}_{1n}^T \mathbf{V}_{1i} \mathbf{e}_{2n}) = \tilde{\mathbf{u}}_{i2(1)}^T \mathbf{V}_{1i} \tilde{\mathbf{u}}_{i2(2)} = \tilde{\lambda}_{i2}$$

となる. 以下同様にして, $\tilde{\lambda}_{ij}, j = 1, \dots, s_i - s_{i-1}$ を得る. それゆえ, 補題 2 の証明と同様にして, 各 $i (= 1, \dots, l)$ について $n \rightarrow \infty$ のもとで

$$\begin{aligned} \frac{\tilde{\lambda}_{ij}}{\lambda_{s_{i-1}+j}} &= 1 + o_p(1), \\ \tilde{\mathbf{u}}_{ij(1)}^T n^{-1/2} \mathbf{z}_{1s_{i-1}+j} &= 1 + o_p(1), \quad \tilde{\mathbf{u}}_{ij(2)}^T n^{-1/2} \mathbf{z}_{2s_{i-1}+j} = 1 + o_p(1) \\ (j &= 1, \dots, s_i - s_{i-1}) \end{aligned}$$

を得る. □

定理 1 の証明 いま, $S_D = \mathbf{U}_1 + \mathbf{U}_2$ に関して, まず $\mathbf{U}_2$ について考える. 対角成分 $u_{ii} (= n^{-1} \sum_{i'=m+1}^d \lambda_{i'} z_{i'i}^2)$ について, 条件

(i') $\alpha_i > 1$ ならば, $d \rightarrow \infty, n$: 固定,

(ii') $\alpha_i \in (0, 1]$ ならば, $d \rightarrow \infty, d^{2-2\alpha_i}/n \rightarrow 0$

のもと, 任意の $\tau > 0$ に対してチェビシエフの不等式より

$$\begin{aligned} \sum_{i=1}^n P\left((nd^{\alpha_i})^{-1} \left| \sum_{i'=m+1}^d \lambda_{i'} z_{i'i}^2 \right| > \tau\right) &\leq (\tau n^{1/2} d^{\alpha_i})^{-2} M\left(\sum_{i'=m+1}^d \lambda_{i'}\right)^2 \\ &= O(d^{2-2\alpha_i}/n) = o(1) \end{aligned}$$

を得る. ここで, M は 4 次モーメントの上界である. この結果を補題 1 と合わせると, 条件

(i'') $\alpha_i > 1$ ならば, $d \rightarrow \infty, n$: 固定,

(ii) $\alpha_i \in (1/4, 1]$ ならば, $d \rightarrow \infty$, $d^{2-2\alpha_i}/n \rightarrow 0$ のもと,

$$d^{-\alpha_i} e_{1n}^T U_2 e_{2n} = o_p(1) \quad (i = 1, \dots, m) \quad (\text{A.3})$$

を得る.

次に, U_1 について考える. いま, $d \rightarrow \infty$ のとき

$$d^{-\alpha_{s_1}} e_{1n}^T U_1 e_{2n} = d^{-\alpha_{s_1}} e_{1n}^T U_{11} e_{2n} + O_p(d^{\alpha_{s_2}-\alpha_{s_1}}) \quad (\text{A.4})$$

に注意する. そのとき, d の冪が α_{s_1} を有する λ_i ($i = 1, \dots, s_1$) について, (A.3) と (A.4) から

$$\frac{\hat{\lambda}_i}{\lambda_i} = \hat{u}_i^T \frac{S_D}{\lambda_i} \hat{u}_i = \hat{u}_i^T \frac{U_1}{\lambda_i} \hat{u}_i + o_p(1) = \hat{u}_i^T \frac{U_{11}}{\lambda_i} \hat{u}_i + o_p(1) \quad (\text{A.5})$$

を得る. (A.5) の第1項について補題2を適用する. 条件 (i'') を

(i) $\alpha_i > 1$ ならば, $d \rightarrow \infty$, $n \rightarrow \infty$ に置き換えて, 条件 (i)-(ii) のもとで

$$\frac{\hat{\lambda}_i}{\lambda_i} = 1 + o_p(1), \quad \hat{u}_i^T n^{-1/2} z_i = 1 + o_p(1) \quad (i = 1, \dots, s_1) \quad (\text{A.6})$$

を得る.

また, d の冪が α_{s_2} を有する λ_i ($i = s_1 + 1, \dots, s_2$) については, まず (A.3) から

$$e_{1n}^T \frac{U_2}{d^{\alpha_{s_2}}} e_{2n} = o_p(1) \quad (\text{A.7})$$

なること, 補題3から $n \rightarrow \infty$ のもとで $\hat{\lambda}_{1i}/\lambda_i = 1 + o_p(1)$, $i = 1, \dots, s_1$ なること, 及び, (A.4) と (A.7) から (i)-(ii) のもとで

$$\begin{aligned} \hat{u}_{1i}^T \frac{S_D}{d^{\alpha_{s_1}}} \hat{u}_j &= \frac{\hat{\lambda}_j}{d^{\alpha_{s_1}}} \hat{u}_{1i}^T \hat{u}_j \\ \hat{u}_{1i}^T \frac{S_D}{d^{\alpha_{s_1}}} \hat{u}_j &= \hat{u}_{1i}^T \frac{U_{11}}{d^{\alpha_{s_1}}} \hat{u}_j + O_p(d^{\alpha_{s_2}-\alpha_{s_1}}) = \frac{\hat{\lambda}_{1i}}{d^{\alpha_{s_1}}} \hat{u}_{1i}^T \hat{u}_j + O_p(d^{\alpha_{s_2}-\alpha_{s_1}}) \end{aligned} \quad (\text{A.8})$$

($i = 1, \dots, s_1$; $j = s_1 + 1, \dots, s_2$)

を得る. ここで, (A.5) から $d^{-\alpha_{s_1}} \hat{\lambda}_i = o_p(1)$, $i = s_1 + 1, \dots, n$ に注意すると, (A.8) は

$$\left(\frac{\hat{\lambda}_{1i}}{d^{\alpha_{s_1}}} + o_p(1) \right) \hat{u}_{1i}^T \hat{u}_j = O_p(d^{\alpha_{s_2}-\alpha_{s_1}}), \quad \text{すなわち} \quad \hat{u}_{1i}^T \hat{u}_j = O_p(d^{\alpha_{s_2}-\alpha_{s_1}})$$

($i = 1, \dots, s_1$; $j = s_1 + 1, \dots, s_2$)

と纏められる. この結果から,

$$\hat{\mathbf{u}}_j^T \frac{U_{11}}{d^{\alpha_{s_2}}} \hat{\mathbf{u}}_j = \sum_{i=1}^{s_1} \frac{\hat{\lambda}_{1i}}{d^{\alpha_{s_2}}} \hat{\mathbf{u}}_j^T \hat{\mathbf{u}}_{1i} \hat{\mathbf{u}}_{1i}^T \hat{\mathbf{u}}_j = O_p(d^{\alpha_{s_2}-\alpha_{s_1}}) \quad (j = s_1 + 1, \dots, s_2) \quad (\text{A.9})$$

を得る. それゆえ, (A.3) と (A.9) から,

$$\hat{\mathbf{u}}_j^T \frac{S_D}{d^{\alpha_{s_2}}} \hat{\mathbf{u}}_j = \hat{\mathbf{u}}_j^T \frac{U_{12}}{d^{\alpha_{s_2}}} \hat{\mathbf{u}}_j + o_p(1) \quad (j = s_1 + 1, \dots, s_2) \quad (\text{A.10})$$

となる. よって, 補題 3 より, (A.5)–(A.6) と同様に条件 (i)–(ii) のもとで,

$$\frac{\hat{\lambda}_i}{\lambda_i} = 1 + o_p(1), \quad \hat{\mathbf{u}}_i^T n^{-1/2} \mathbf{z}_i = 1 + o_p(1) \quad (i = s_1 + 1, \dots, s_2) \quad (\text{A.11})$$

を得る.

また, d の冪が α_{s_3} を有する λ_i ($i = s_2 + 1, \dots, s_3$) については, $U_{11} + U_{12} = \sum_{i=1}^{s_2} \hat{\lambda}_{2i} \hat{\mathbf{u}}_{2i} \hat{\mathbf{u}}_{2i}^T$ なるスペクトル分解を考える. ここで, $\hat{\lambda}_{2i}$ ($i = 1, \dots, s_2$) と $\hat{\mathbf{u}}_{2i}$ ($i = 1, \dots, s_2$) $\in R_n$ は $U_{11} + U_{12}$ の固有値と固有ベクトルである. いま, (A.5)–(A.11) で S_D を $U_{11} + U_{12}$ に書き換えれば, $d \rightarrow \infty, n \rightarrow \infty$ のもとで, $\hat{\lambda}_{2i}/\lambda_i = 1 + o_p(1)$, $i = 1, \dots, s_2$ が示される. よって, (A.8) と同様に, (i)–(ii) のもとで

$$\begin{aligned} \hat{\mathbf{u}}_{2i}^T \frac{S_D}{d^{\alpha_{s_2}}} \hat{\mathbf{u}}_j &= \frac{\hat{\lambda}_j}{d^{\alpha_{s_2}}} \hat{\mathbf{u}}_{2i}^T \hat{\mathbf{u}}_j, \\ \hat{\mathbf{u}}_{2i}^T \frac{S_D}{d^{\alpha_{s_2}}} \hat{\mathbf{u}}_j &= \hat{\mathbf{u}}_{2i}^T \frac{U_{11} + U_{12}}{d^{\alpha_{s_2}}} \hat{\mathbf{u}}_j + O_p(d^{\alpha_{s_3}-\alpha_{s_2}}) = \frac{\hat{\lambda}_{2i}}{d^{\alpha_{s_2}}} \hat{\mathbf{u}}_{2i}^T \hat{\mathbf{u}}_j + O_p(d^{\alpha_{s_3}-\alpha_{s_2}}) \\ (i = 1, \dots, s_2; j = s_2 + 1, \dots, s_3) \end{aligned}$$

を得る. このとき, $j = s_2 + 1, \dots, n$ について

$$\hat{\mathbf{u}}_{2i}^T \hat{\mathbf{u}}_j = \begin{cases} O_p(d^{\alpha_{s_3}-\alpha_{s_1}}) & (i = 1, \dots, s_1), \\ O_p(d^{\alpha_{s_3}-\alpha_{s_2}}) & (i = s_1 + 1, \dots, s_2) \end{cases}$$

となるので, (A.9)–(A.11) と同様にして, $i = s_2 + 1, \dots, s_3$ に対して (A.11) を得る. 以下同様にして, d の冪が α_{s_l} を有する λ_i ($i = s_{l-1} + 1, \dots, s_l$) まで (A.11) が主張でき, それゆえ, 各 i ($= 1, \dots, m$) について条件 (i)–(ii) のもとで

$$\frac{\hat{\lambda}_i}{\lambda_i} = 1 + o_p(1), \quad \hat{\mathbf{u}}_i^T n^{-1/2} \mathbf{z}_i = 1 + o_p(1) \quad (\text{A.12})$$

が主張できる. □

定理 2 の証明 いま, $S_{D(1)}^2 = n^{-2} \mathbf{X}_1^T \mathbf{X}_2 \mathbf{X}_2^T \mathbf{X}_1 = \sum_{i=1}^n \tilde{\lambda}_i^2 \mathbf{u}_{i(1)} \mathbf{u}_{i(1)}^T$, $S_{D(2)}^2 = n^{-2} \mathbf{X}_2^T \mathbf{X}_1 \mathbf{X}_1^T \mathbf{X}_2 = \sum_{i=1}^n \tilde{\lambda}_i^2 \mathbf{u}_{i(2)} \mathbf{u}_{i(2)}^T$ とおく. ただし, $\mathbf{u}_{i(1)}$ ($i = 1, \dots, n$) $\in R_n$ は $S_{D(1)}^2$ の固有ベクトル, $\mathbf{u}_{i(2)}$ ($i = 1, \dots, n$) $\in R_n$ は $S_{D(2)}^2$ の固有ベクトルであ

る. ここで, $S_{D(1)} = n^{-1} \mathbf{X}_1^T \mathbf{X}_2$ とおけば, $S_{D(1)} = \sum_{i=1}^n \tilde{\lambda}_i \mathbf{u}_{i(1)} \mathbf{u}_{i(2)}^T$ と特異値分解される.

まず, 定義から $S_{D(1)} = \mathbf{V}_1 + \mathbf{V}_2$ と表され, $\mathbf{V}_2$ については補題 1 より

(i') $\alpha_i > 1/2$ ならば, $d \rightarrow \infty$, n : 固定,

(ii) $\alpha_i \in (1/4, 1/2]$ ならば, $d \rightarrow \infty$, $d^{2-2\alpha_i}/n \rightarrow 0$

のもとで,

$$\mathbf{e}_{1n}^T \frac{\mathbf{V}_2}{\lambda_i} \mathbf{e}_{2n} = o_p(1) \quad (i = 1, \dots, m) \quad (\text{A.13})$$

となることに注意する. いま, d の冪が α_{s_1} を有する λ_i ($i = 1, \dots, s_1$) について, (A.3)–(A.5) と同様の手順で, (A.13) と (i'), (ii) のもとで

$$\frac{\tilde{\lambda}_i}{\lambda_i} = \tilde{\mathbf{u}}_{i(1)}^T \frac{S_{D(1)}}{\lambda_i} \tilde{\mathbf{u}}_{i(2)} = \tilde{\mathbf{u}}_{i(1)}^T \frac{\mathbf{V}_{11}}{\lambda_i} \tilde{\mathbf{u}}_{i(2)} + o_p(1) \quad (\text{A.14})$$

を得る. ここで, 条件 (i') を

(i) $\alpha_i > 1/2$ ならば, $d \rightarrow \infty$, $n \rightarrow \infty$

に置き換えて補題 3 を使えば, λ_i ($i = 1, \dots, s_1$) について

$$\frac{\tilde{\lambda}_i}{\lambda_i} = \tilde{\mathbf{u}}_{i(1)}^T \frac{S_{D(1)}}{\lambda_i} \tilde{\mathbf{u}}_{i(2)} = 1 + o_p(1), \quad \tilde{\mathbf{u}}_{i(j)}^T n^{-1/2} \mathbf{z}_{ji} = 1 + o_p(1) \quad (j = 1, 2) \quad (\text{A.15})$$

を得る.

次に, d の冪が α_{s_2} を有する λ_i ($i = s_1 + 1, \dots, s_2$) について, (A.7)–(A.8) と同様の手順で, (i), (ii) のもとで

$$\frac{\tilde{\lambda}_j}{d^{\alpha_{s_1}}} \tilde{\mathbf{u}}_{1i(1)}^T \tilde{\mathbf{u}}_{j(1)} = \tilde{\mathbf{u}}_{1i(1)}^T \frac{S_{D(1)}}{d^{\alpha_{s_1}}} \tilde{\mathbf{u}}_{j(2)} = \frac{\tilde{\lambda}_{i1}}{d^{\alpha_{s_1}}} \tilde{\mathbf{u}}_{1i(2)}^T \tilde{\mathbf{u}}_{j(2)} + O_p(d^{\alpha_{s_2} - \alpha_{s_1}}), \quad (\text{A.16})$$

$$\frac{\tilde{\lambda}_j}{d^{\alpha_{s_1}}} \tilde{\mathbf{u}}_{1i(2)}^T \tilde{\mathbf{u}}_{j(2)} = \tilde{\mathbf{u}}_{1i(2)}^T \frac{S_{D(1)}^T}{d^{\alpha_{s_1}}} \tilde{\mathbf{u}}_{j(1)} = \frac{\tilde{\lambda}_{i1}}{d^{\alpha_{s_1}}} \tilde{\mathbf{u}}_{1i(1)}^T \tilde{\mathbf{u}}_{j(1)} + O_p(d^{\alpha_{s_2} - \alpha_{s_1}}) \quad (\text{A.17})$$

$$(i = 1, \dots, s_1; j = s_1 + 1, \dots, s_2)$$

を得る. ここで, (A.14) から $d^{-\alpha_{s_1}} \tilde{\lambda}_i = o_p(1)$, $i = s_1 + 1, \dots, n$ に注意し, (A.16)–(A.17) から, (i), (ii) のもとで $\tilde{\mathbf{u}}_{1i(1)}^T \tilde{\mathbf{u}}_{j(1)} = O_p(d^{\alpha_{s_2} - \alpha_{s_1}})$, $\tilde{\mathbf{u}}_{1i(2)}^T \tilde{\mathbf{u}}_{j(2)} = O_p(d^{\alpha_{s_2} - \alpha_{s_1}})$, $i = 1, \dots, s_1$; $j = s_1 + 1, \dots, s_2$ を得る. それゆえ,

$$\tilde{\mathbf{u}}_{j(1)}^T \frac{S_{D(1)}}{d^{\alpha_{s_2}}} \tilde{\mathbf{u}}_{j(2)} = \tilde{\mathbf{u}}_{j(1)}^T \frac{\mathbf{V}_{12}}{d^{\alpha_{s_2}}} \tilde{\mathbf{u}}_{j(2)} + o_p(1)$$

$$(j = s_1 + 1, \dots, s_2)$$

となり, 補題 3 から $j = s_1 + 1, \dots, s_2$ に対して (A.15) を得る. 以下同様の手順で, d の冪が α_{s_l} を有する λ_i ($i = s_{l-1} + 1, \dots, s_l$) まで (A.15) が主張でき, 各 i ($= 1, \dots, m$) において (A.15) を得る. $\square$

定理3の証明 いま, $\mathbf{X} = \mathbf{H}\mathbf{\Lambda}^{1/2}\mathbf{Z}$ と表され, 各 $i (= 1, \dots, n)$ で $\mathbf{h}_i^T \mathbf{X} = \lambda_i^{1/2} \mathbf{z}_i^T$ である. よって, (5.3) に注意して

$$\mathbf{h}_i^T \hat{\mathbf{h}}_i = (\hat{\lambda}_i n)^{-1/2} \lambda_i^{1/2} \mathbf{z}_i^T \hat{\mathbf{u}}_i = \left(\frac{\lambda_i}{\hat{\lambda}_i} \right)^{1/2} \frac{\mathbf{z}_i^T}{\sqrt{n}} \hat{\mathbf{u}}_i$$

を得る. そのとき, 定理1の証明から, 各 $i (= 1, \dots, m)$ において条件 (i)–(ii) のもとで (A.12) が主張できるので,

$$\mathbf{h}_i^T \hat{\mathbf{h}}_i = 1 + o_p(1) \quad (i = 1, \dots, m)$$

を得る. それゆえ, 各 $i (= 1, \dots, m)$ で, $\text{Angle}(\mathbf{h}_i, \hat{\mathbf{h}}_i) = o_p(1)$ が成立する. $\square$

定理4の証明 定義から, 各 $i (= 1, \dots, n)$ で

$$\begin{aligned} V(\hat{\mathbf{s}}_i) &= \lambda_i n^{-1} \sum_{j=1}^n \left(z_{ij} - \sqrt{n \frac{\hat{\lambda}_i}{\lambda_i}} \hat{u}_{ij} \right)^2 \\ &= \lambda_i \left(n^{-1} \sum_{j=1}^n z_{ij}^2 + \frac{\hat{\lambda}_i}{\lambda_i} \sum_{j=1}^n \hat{u}_{ij}^2 - 2 \sqrt{\frac{\hat{\lambda}_i}{\lambda_i}} \frac{\mathbf{z}_i^T}{\sqrt{n}} \hat{\mathbf{u}}_i \right) \end{aligned}$$

となる. ここで, 各 $i (= 1, \dots, m)$ において $n^{-1} \sum_{j=1}^n z_{ij}^2 = 1 + o_p(1)$ なることと定理1の証明に注意すると, 条件 (i)–(ii) のもとで (A.12) が主張でき, 結果的に

$$\frac{V(\hat{\mathbf{s}}_i)}{\lambda_i} = o_p(1) \quad (i = 1, \dots, m)$$

を得る. $\square$

謝辞 本研究は, 科学研究費補助金 基盤研究 (B) 18300092 研究代表者: 青嶋 誠 「高次元小標本におけるデータ解析の数理統計学的基礎とその応用」から, 研究助成を受けています.

参考文献

- Ahn, J., Marron, J. S., Muller, K. M. and Chi, Y.-Y. (2007). The high-dimension, low-sample-size geometric representation holds under mild conditions. *Biometrika* **94**, 760–766.
- Aoshima, M. and Yata, K. (2008). Asymptotic second-order consistency for two-stage estimation methodologies and its applications. *Ann. Inst. Statist. Math.*, in press.
- Baik, J., Ben Arous, G., and Péché, S. (2005). Phase transition of the largest eigenvalue for non-null complex covariance matrices. *Ann. Probab.* **33**, 1643–1697.

- Baik, J. and Silverstein, J. W. (2006). Eigenvalues of large sample covariance matrices of spiked population models. *J. Multivariate Anal.* **97**, 1382–1408.
- Hall, P., Marron, J. S. and Neeman, A. (2005). Geometric representation of high dimension, low sample size data. *J. R. Statist. Soc.*, **B 67**, 427–444.
- Johnstone, I. M. (2001). On the distribution of the largest eigenvalue in principal components analysis. *Ann. Statist.*, **29**, 295–327.
- Jung, S. and Marron, J. S. (2008). PCA consistency in high dimension, low sample size context. *Ann. Statist.*
- Muller, K. E., Chi, Y.-Y., Ahn, J. and Marron, J. S. (2008). Limitations of high dimension, low sample size principal components for gaussian data. *J. Amer. Statist. Assoc.*
- Paul, D. (2007). Asymptotics of sample eigenstructure for a large dimensional spiked covariance model. *Statistica Sinica* **17**, 1617–1642.
- Yata, K. (2008). Two-stage equivalence tests that control both the size and power. *Seq. Anal.*, **27**, 185–200.
- Yata, K. and Aoshima, M. (2008). Intrinsic dimensionality estimation of high dimension, low sample size data with d -asymptotics, revised.
- Yata, K. and Aoshima, M. (2009). Double shrink methodologies to determine the sample size via covariance structures. *J. Statist. Plan. Infer.*, **139**, 81–99.