強スパイク固有値モデルにおける高次元相関行列の検定

筑波大学・数理物質科学研究群　　巖名　佑務
Yumu Iwana
Graduate School of Pure and Applied Sciences,
Tsukuba University
筑波大学・数理物質系　　矢田　和善
Kazuyoshi Yata
Institute of Mathematics,
University of Tsukuba
東京理科大学・創域理工学部　　石井　晶
Aki Ishii
Department of Information Sciences,
Tokyo University of Science
筑波大学・数理物質系　　青嶋　誠
Makoto Aoshima
Institute of Mathematics,
University of Tsukuba

1　はじめに

本論文では，強スパイク固有値モデルのもとで高次元相関行列の検定を考える．高次元データにおける相関行列の検定は，Aoshima and Yata [1]，Yata and Aoshima [5, 7] によって，弱スパイク固有値 (Non-Strongly Spiked Eigenvalue) モデルのもとで拡張クロスデータ行列法 (ECDM) を用いた検定統計量が与えられた．ここで，弱スパイク固有値モデルは Aoshima and Yata [2] で考案され，以下で与えられる．

$$\frac{\lambda_1}{\sqrt{\mathrm{tr}(\boldsymbol{\Sigma}^2)}} \to 0, \quad p \to \infty.$$

ただし，$\boldsymbol{\Sigma}$ は p 次の共分散行列であり，λ_1 はその最大固有値である．一方で，Aoshima and Yata [2] は，強スパイク固有値 (Strongly Spiked Eigenvalue) モデルも考案した．強

スパイク固有値モデルは，次の条件を満たす固有値モデルである．

$$\liminf_{p\to\infty}\frac{\lambda_1}{\sqrt{\mathrm{tr}(\boldsymbol{\Sigma}^2)}}>0.$$

詳細は，Aoshima and Yata [2] を参照のこと．本論文では，強スパイク固有値モデルのもとで ECDM を用いた新たな検定統計量を与え，その漸近分布を用いた高次元相関行列の検定手法を提案する．最後に，数値シミュレーションによって，今回提案する新たな検定手法の精度を確認する．

2　相関行列の検定

母集団分布に p 次元の分布を考え，n 個のデータ $\boldsymbol{x}_1,...,\boldsymbol{x}_n$ を無作為に抽出する．ただし，

$$\boldsymbol{x}_j=(\boldsymbol{x}_{1j}^\top,\boldsymbol{x}_{2j}^\top)^\top,\quad j=1,...,n$$

とし，$\boldsymbol{x}_{ij}\in\mathbb{R}^{p_i}$，$p_1\in\{1,...,p-1\}$，$p_2=p-p_1$ とする．ここで，$\boldsymbol{x}_j$ は未知の平均ベクトル $\boldsymbol{\mu}=(\boldsymbol{\mu}_1^\top,\boldsymbol{\mu}_2^\top)^\top$，共分散行列

$$\boldsymbol{\Sigma}=\begin{pmatrix}\boldsymbol{\Sigma}_1 & \boldsymbol{\Sigma}_*\\ \boldsymbol{\Sigma}_*^\top & \boldsymbol{\Sigma}_2\end{pmatrix}(>\boldsymbol{O})$$

をもつとする．ただし，各 i,j で

$$\mathrm{E}(\boldsymbol{x}_{ij})=\boldsymbol{\mu}_i,\quad \mathrm{Var}(\boldsymbol{x}_{ij})=\boldsymbol{\Sigma}_i,\quad \mathrm{Cov}(\boldsymbol{x}_{1j},\boldsymbol{x}_{2j})=\mathrm{E}(\boldsymbol{x}_{1j}\boldsymbol{x}_{2j}^\top)-\boldsymbol{\mu}_1\boldsymbol{\mu}_2^\top=\boldsymbol{\Sigma}_*$$

とする．相関行列を

$$\boldsymbol{P}=\mathrm{Corr}(\boldsymbol{x}_{1j},\boldsymbol{x}_{2j})=\mathrm{diag}(\sigma_{11},\sigma_{12},\ldots,\sigma_{1p_1})^{-1/2}\boldsymbol{\Sigma}_*\mathrm{diag}(\sigma_{21},\sigma_{22},\ldots,\sigma_{2p_2})^{-1/2}$$

とおく．ただし，$\sigma_{i1},\sigma_{12},\ldots,\sigma_{ip_i}$ は $\boldsymbol{\Sigma}_i$ の対角成分である．このとき，次の検定を考える．

$$H_0:\boldsymbol{P}=\boldsymbol{O}\ \ vs.\ \ H_1:\boldsymbol{P}\neq\boldsymbol{O}. \tag{1}$$

Yata and Aoshima [7] は，拡張クロスデータ行列法 (ECDM) による推定量・検定統計量を一般化し，弱スパイク固有値モデルのもと，相関行列の検定 (1) の高次元検定手法を構築した．Aoshima and Yata [2] は，強スパイク固有値モデルのもとで高次元二標本検定を考案した．また，Ishii, Yata and Aoshima [3] では，強スパイク固有値モデルのもとで高次元共分散行列の構造に関する検定を考案した．本論文は，検定 (1) において，強スパイク固有値モデルのもと ECDM を用いた検定統計量の漸近分布を導出し，新たな検定手法を提案する．

3 ECDM を用いた弱スパイク固有値モデルにおける相関行列の検定

本節では，Yata and Aoshima [7] が与えた ECDM を用いた弱スパイク固有値モデルのもとでの (1) の検定方式を紹介する．次のモデルを考える．各 j に対し，

$$\boldsymbol{x}_j = \boldsymbol{\Gamma}\boldsymbol{w}_j + \boldsymbol{\mu}, \quad \boldsymbol{w}_j = (w_{1j}, \ldots, w_{qj})^\top.$$

ただし，$\boldsymbol{\Gamma}$ は $\boldsymbol{\Gamma}\boldsymbol{\Gamma}^\top = \boldsymbol{\Sigma}$ を満たすある $p \times q$ 行列である．そのとき，$\mathrm{E}(\boldsymbol{w}_j) = \boldsymbol{0}, \mathrm{Var}(\boldsymbol{w}_j) = \boldsymbol{I}_q$ となることに注意する．ただし，$\boldsymbol{I}_q$ は q 次元の単位行列を表す．もし $\boldsymbol{x}_j$ が正規分布に従うならば，$\boldsymbol{w}_j$ は $N_q(\boldsymbol{0}, \boldsymbol{I}_q)$ に従う．$r = 1, \ldots, q$ に対し，$\mathrm{Var}(w_{rj}^2) = M_r$ とおき，$\limsup_{p\to\infty} M_r < \infty$ と仮定する．$\boldsymbol{\Sigma}_1$ と $\boldsymbol{\Sigma}_2$ に対し，次を仮定する．

$$\textbf{(A-i)} \quad \min\left\{ \frac{\lambda_{11}}{\sqrt{\mathrm{tr}(\boldsymbol{\Sigma}_1^2)}}, \frac{\lambda_{21}}{\sqrt{\mathrm{tr}(\boldsymbol{\Sigma}_2^2)}} \right\} \to 0, \quad p \to \infty.$$

ただし，λ_{i1} は $\boldsymbol{\Sigma}_i$ の最大固有値である．固有値モデル (A-i) は，弱スパイク固有値モデルの 1 つである．$m = \min\{p, n\}$，$\Delta = \mathrm{tr}(\boldsymbol{\Sigma}_*\boldsymbol{\Sigma}_*^\top)$ $(= \|\boldsymbol{\Sigma}_*\|_F^2)$ とおく．ただし，$\|\cdot\|_F$ はフロベニウスノルムを表す．$\Delta = 0$ と $\boldsymbol{P} = \boldsymbol{O}$ は同値であることに注意する．さらに，必要に応じて次を仮定する．

(A-ii) 全ての $v \in \{2, ..., 8\}$，$r_1 \neq r_2 \neq \cdots \neq r_v \in [1, q]$ ，$\alpha_1, ..., \alpha_v \in [1, 4]$，$\alpha_1 + \cdots + \alpha_v \leq 8$ に対し，

$$\mathrm{E}(w_{r_1 j}^{\alpha_1} \cdots w_{r_v j}^{\alpha_v}) = E(w_{r_1 j}^{\alpha_1}) \cdots E(w_{r_v j}^{\alpha_v}).$$

$\boldsymbol{x}_j$ が正規分布に従うならば，(A-ii) を満たすことに注意する．

ここで, ECDM を用いた Δ の不偏推定量について考える．$n_{(1)} = \lceil n/2 \rceil$, $n_{(2)} = n - n_{(1)}$ とおく．ここで，$\lceil x \rceil$ は x 以上の最小の整数を表す．2 つの集合 $\boldsymbol{V}_{n(1)(k)}$，$\boldsymbol{V}_{n(2)(k)}$，$k = 3, \ldots, 2n-1$ を次のように定義する．

$$\boldsymbol{V}_{n(1)(k)} = \begin{cases} \{\lfloor k/2 \rfloor - n_{(1)} + 1, ..., \lfloor k/2 \rfloor\}, & \lfloor k/2 \rfloor \geq n_{(1)} \text{のとき}, \\ \{1, \ldots, \lfloor k/2 \rfloor\} \cup \{\lfloor k/2 \rfloor + n_{(2)} + 1, ..., n\}, & \text{それ以外}. \end{cases}$$

$$\boldsymbol{V}_{n(2)(k)} = \begin{cases} \{\lfloor k/2 \rfloor + 1, ..., \lfloor k/2 \rfloor + n_{(2)}\}, & \lfloor k/2 \rfloor \leq n_{(1)} \text{のとき}, \\ \{1, ..., \lfloor k/2 \rfloor - n_{(1)}\} \cup \{\lfloor k/2 \rfloor + 1, ..., n\}, & \text{それ以外}. \end{cases}$$

ここで，$\lfloor x \rfloor$ は x 以下の最大の整数を表す．そのとき，$k = 3, ..., 2n-1$ について，

$$\#\boldsymbol{V}_{n(\ell)(k)} = n_{(\ell)},\ \ell = 1, 2, \quad \boldsymbol{V}_{n(1)(k)} \cap \boldsymbol{V}_{n(2)(k)} = \emptyset, \quad \boldsymbol{V}_{n(1)(k)} \cup \boldsymbol{V}_{n(2)(k)} = \{1, ..., n\}$$

となること，及び，$i < j\ (\leq n)$ について，

$$i \in \boldsymbol{V}_{n(1)(i+j)}, \quad j \in \boldsymbol{V}_{n(2)(i+j)}$$

となることに注意する．ここで，$\#\boldsymbol{S}$ は集合 $\boldsymbol{S}$ の要素の個数を表す．$\ell\ (= 1, 2)$，$k\ (= 3, ..., 2n-1)$ について，2 分割した集合の平均を

$$\overline{\boldsymbol{x}}_{\ell(1)(k)} = \frac{1}{n_{(1)}} \sum_{j \in \boldsymbol{V}_{n(1)(k)}} \boldsymbol{x}_{\ell j}, \quad \overline{\boldsymbol{x}}_{\ell(2)(k)} = \frac{1}{n_{(2)}} \sum_{j \in \boldsymbol{V}_{n(2)(k)}} \boldsymbol{x}_{\ell j}$$

とし，ある $i, j\ (1 \leq i < j \leq n)$ について，Δ の 1 つの不偏推定量として

$$\widehat{\Delta}_{ij} = u_n(\boldsymbol{x}_{1i} - \overline{\boldsymbol{x}}_{1(1)(i+j)})^\top(\boldsymbol{x}_{1j} - \overline{\boldsymbol{x}}_{1(2)(i+j)})(\boldsymbol{x}_{2i} - \overline{\boldsymbol{x}}_{2(1)(i+j)})^\top(\boldsymbol{x}_{2j} - \overline{\boldsymbol{x}}_{2(2)(i+j)})$$

を計算する．ただし，$u_n = n_{(1)}n_{(2)}/\{(n_{(1)} - 1)(n_{(2)} - 1)\}$ である．全ての組み合わせで平均を取り，

$$\widehat{T}_n = \frac{2u_n}{n(n-1)} \sum_{i<j}^{n} \widehat{\Delta}_{ij}$$

を定義する．このとき，$\mathrm{E}(\widehat{T}_n) = \Delta$ となることに注意する．

$$\delta = \sqrt{2\mathrm{tr}(\boldsymbol{\Sigma}_1^2)\mathrm{tr}(\boldsymbol{\Sigma}_2^2)}/n$$

とおく．Yata and Aoshima [7] は，次の定理を与えた．

定理 1 ([7])**.** *(A-i)* と *(A-ii)* を仮定する．H_0 のもと，$m \to \infty$ のとき次が成り立つ．

$$\mathrm{Var}(\widehat{T}_n) = \delta^2\{1 + o(1)\}.$$

定理 2 ([7])**.** *(A-i)* と *(A-ii)* を仮定する．H_0 のもと，$m \to \infty$ のとき次が成り立つ．

$$\frac{\widehat{T}_n}{\delta} \Rightarrow N(0, 1).$$

ただし，$\Rightarrow$ は分布収束を表す．

Yata and Aoshima [7] は $i=1,2$ に対し，$\mathrm{tr}(\boldsymbol{\Sigma}_i^2)$ の推定量を次で与えた．

$$W_{in}=\frac{2u_n}{n(n-1)}\sum_{r<s}^{n}\{(\boldsymbol{x}_{ir}-\overline{\boldsymbol{x}}_{i(1)(r+s)})^{\top}(\boldsymbol{x}_{is}-\overline{\boldsymbol{x}}_{i(2)(r+s)})\}^2.$$

$\mathrm{E}(W_{in})=\mathrm{tr}(\boldsymbol{\Sigma}_i^2)$ となる．定理 2 より，Yata and Aoshima [7] は，(1) の検定方式を以下で与えた．

$$H_0\text{を棄却} \quad \Leftrightarrow \quad T/\widehat{\delta}>z_{\alpha}. \tag{2}$$

ただし，$\widehat{\delta}=n^{-1}(2W_{1n}W_{2n})^{1/2}$ であり，z_{α} は $N(0,1)$ の上側 $100\alpha\%$ 点である．このとき，検定方式 (2) における第一種の誤り確率 (Size) は

$$\mathrm{Size}=\alpha+o(1)$$

となる．

4 ECDM を用いた強スパイク固有値モデルにおける相関行列の検定

本節では，強スパイク固有値モデルのもとでの (1) の検定方式について考える．次を仮定する．

(C-i) $\quad p_1=1; \quad \dfrac{\lambda_{21}}{\sqrt{\mathrm{tr}(\boldsymbol{\Sigma}_2^2)}}\to 1, \quad p_2\to\infty.$

固有値モデル (C-i) は，強スパイク固有値モデルの 1 つである．各 j に対し，

$$\begin{aligned}x_{1j}&=\sigma_{11}^{1/2}z_{1j}+\mu_1;\\ \boldsymbol{x}_{2j}&=\boldsymbol{H}_2\boldsymbol{\Lambda}_2^{1/2}\boldsymbol{z}_{2j}+\boldsymbol{\mu}_2\end{aligned}$$

とする．ただし，$\sigma_{11}>0$，$\boldsymbol{\Lambda}_2=\mathrm{diag}(\lambda_{21},\ldots,\lambda_{2p_2})$ は $\boldsymbol{\Sigma}_2$ の固有値 $\lambda_{21}\geq\cdots\geq\lambda_{2p_2}>0$ を対角成分にもつ対角行列，$\boldsymbol{H}_2$ は対応する固有ベクトルを並べた直交行列である．そのとき，$\mathrm{E}(z_{1j})=0,\mathrm{Var}(z_{1j})=1$，$\mathrm{E}(\boldsymbol{z}_{2j})=\boldsymbol{0}$，$\mathrm{Var}(\boldsymbol{z}_{2j})=\boldsymbol{I}_{p_2}$ となることに注意する．さらに，H_0 のもと $E(z_{1j}\boldsymbol{z}_{2j})=\boldsymbol{0}$ となる．$\boldsymbol{z}_{2j}=(z_{2j},z_{3j},...,z_{pj})^{\top}$ とおく．ここで，すべての r で $\limsup_{p\to\infty}\mathrm{E}(z_{rj}^4)<\infty$ と仮定する．さらに，次を仮定する．

(C-ii) $\mathrm{E}(z_{sj}^2z_{tj}^2)=1, \quad \mathrm{E}(z_{sj}z_{tj}z_{s'j}z_{t'j})=0, \quad s\neq t,s',t'.$

$\boldsymbol{x}_j$ が正規分布に従うならば，H_0 のもと (C-ii) を満たすことに注意する．このとき，次が成り立つ．

補題 1. *(C-i)* と *(C-ii)* を仮定する．H_0 のもと，$m \to \infty$ のとき次が成り立つ．

$$\mathrm{Var}(\widehat{T}_n) = 2\frac{\sigma_{11}^2\lambda_{21}^2}{n^2}\{1 + o(1)\}.$$

定理 3. *(C-i)* と *(C-ii)* を仮定する．H_0 のもと，$m \to \infty$ のとき次が成り立つ．

$$\frac{n\widehat{T}_n}{\sigma_{11}\lambda_{21}} + 1 \Rightarrow \chi_1^2.$$

ただし，χ_1^2 は自由度 *1* のカイ二乗分布を表す．

ここで，λ_{21} の推定については，Yata and Aoshima [4] で与えられたノイズ掃き出し法を用いる．いま，$\boldsymbol{x}_{2j}$ の標本共分散行列を

$$\boldsymbol{S}_2 = (n-1)^{-1}\sum_{j=1}^{n}(\boldsymbol{x}_{2j} - \overline{\boldsymbol{x}}_2)(\boldsymbol{x}_{2j} - \overline{\boldsymbol{x}}_2)^\top$$

とおく．ただし，

$$\overline{\boldsymbol{x}}_2 = n^{-1}\sum_{j=1}^{n}\boldsymbol{x}_{2j}$$

である．ノイズ掃き出し法を用いた λ_{21} の推定量 $\hat{\lambda}_{2(\mathrm{NR})}$ は次のように計算する．

$$\hat{\lambda}_{2(\mathrm{NR})} = \lambda_{\mathrm{Max}}(\boldsymbol{S}_2) - \frac{\mathrm{tr}(\boldsymbol{S}_2) - \lambda_{\mathrm{Max}}(\boldsymbol{S}_2)}{n-2}.$$

ただし，$\lambda_{\mathrm{Max}}(\boldsymbol{S}_2)$ は $\boldsymbol{S}_2$ の最大固有値である．そのとき，Yata and Aoshima [4, 6] より，次の一致性が成り立つ．

補題 2 ([4, 5])**.** *(C-i)* と *(C-ii)* を仮定する．$m \to \infty$ のとき次が成り立つ．

$$\frac{\hat{\lambda}_{2(\mathrm{NR})}}{\lambda_{21}} = 1 + o_p(1).$$

いま，

$$U = \frac{n\widehat{T}_n}{\widehat{\sigma}_{11}\hat{\lambda}_{2(\mathrm{NR})}} + 1$$

とおく．ただし，

$$\widehat{\sigma}_{11} = \frac{1}{n-1}\sum_{j=1}^{n}(x_{1j} - \overline{x}_1)^2; \quad \overline{x}_1 = \frac{1}{n}\sum_{j=1}^{n} x_{1j}$$

である．

定理 3 と補題 2 より，次が成り立つ．

系 1. *(C-i)* と *(C-ii)* を仮定する．H_0 のもと，$m \to \infty$ のとき次が成り立つ．

$$U \Rightarrow \chi_1^2.$$

系 1 より，(1) の検定方式を以下で与える．

$$H_0\text{を棄却} \quad \Leftrightarrow \quad U > \chi_1^2(\alpha). \tag{3}$$

ただし，$\chi_1^2(\alpha)$ は χ_1^2 の上側 $100\alpha\%$ 点である．それゆえ，検定 (3) において，(C-i) と (C-ii) のもと，第一種の過誤の確率は α に収束する．

5　数値シミュレーション

検定方式 (3) の性能を数値シミュレーションで確認する．$\alpha = 0.05$, $p_1 = 1$, $p_2 = 2^s (s = 6, \ldots, 11)$，$\sigma_{11} = 1$，$\boldsymbol{\mu} = \mathbf{0}$ とし，

$$\boldsymbol{\Sigma}_2 = \begin{pmatrix} \boldsymbol{\Sigma}_{(1)} & \boldsymbol{O}_{2,p-2} \\ \boldsymbol{O}_{p-2,2} & \boldsymbol{\Sigma}_{(2)} \end{pmatrix}, \quad \boldsymbol{\Sigma}_{(1)} = \mathrm{diag}(p_2^{4/5}, p_2^{1/2}), \quad \boldsymbol{\Sigma}_{(2)} = (0.3^{|i-j|^{1/2}})$$

とする．ただし，$\boldsymbol{O}_{\ell,\ell'}$ は $\ell \times \ell'$ のゼロ行列とする．H_0 の場合について考える．

$\boldsymbol{x}_j$ を次の分布から発生させる．

(I) $\boldsymbol{x}_j \sim N_p(\mathbf{0}, \boldsymbol{\Sigma})$;

(II) $z_{sj} = 2^{-1/2}(v_{sj} - 1)$，$s = 1, ..., p$，$v_{sj}$s are i.i.d. as χ_1^2.

$n = 4\lceil p_2^{2/3} \rceil$ とする．図 1 は，(I)，(II) について，それぞれ 2000 回の結果の平均をプロットしたものである．$\Pr\ (r = 1, ..., 2000)$ を，H_0 を誤って棄却すれば 1，そうでなければ 0 と定義する．第一種の過誤の確率（Size）を

$$\overline{\alpha} = \frac{1}{2000}\sum_{r=1}^{2000} \Pr$$

で推定した．

理論通り，強スパイク固有値モデル (C-i) のもとで，検定方式 (3) の Size が α に収束していくことが確認できた．

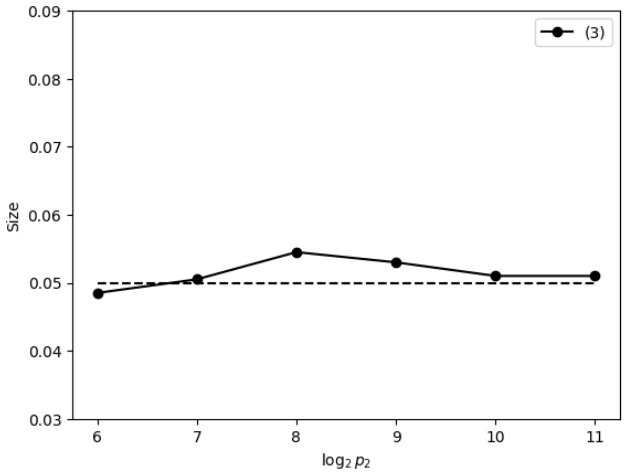

図 1　(I) $\boldsymbol{x}_j \sim N_p(\mathbf{0}, \boldsymbol{\Sigma})$

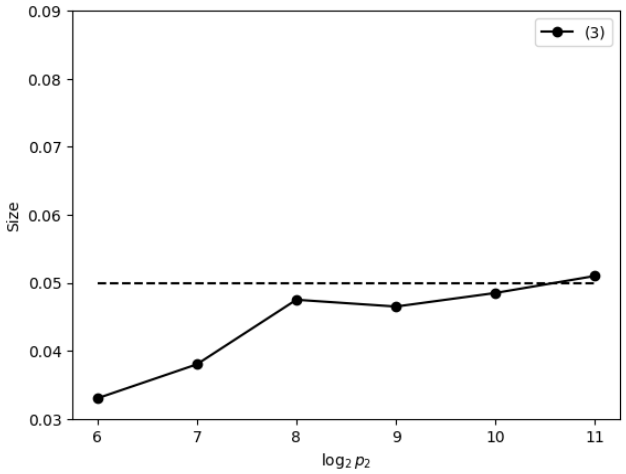

図 2　(II) $z_{sj} = 2^{-1/2}(v_{sj} - 1)$,　$s = 1, ..., p$,　v_{sj}s are i.i.d. as　χ^2_1

6　付録

補題 *1* の証明．Yata and Aoshima [7] の補題 3.1 と補題 3.2 の証明より補題を得る．　□

定理 *3* の証明．H_0，(C-i)，(C-ii) を仮定する．補題 1 と $\widehat{T}_n$ の主要項を考えると，$m \to \infty$

のもと，

$$\begin{aligned}\frac{n\widehat{T}_n}{\lambda_{21}} &= \sigma_{11}\frac{2}{n}\sum_{j<j'}^{n} z_{1j}z_{1j'}z_{2j}z_{2j'} + o_p(1) \\ &= \sigma_{11}\left\{\left(\frac{1}{\sqrt{n}}\sum_{j=1}^{n} z_{1j}z_{2j}\right)^2 - \frac{1}{n}\sum_{j=1}^{n} z_{1j}^2 z_{2j}^2\right\} + o_p(1)\end{aligned}$$

となる．ここで，中心極限定理より，

$$\left(\frac{1}{\sqrt{n}}\sum_{j=1}^{n} z_{1j}z_{2j}\right)^2 \Rightarrow \chi_1^2$$

となり，また，大数の法則より，$n \to \infty$ のとき，

$$\frac{1}{n}\sum_{j=1}^{n} z_{1j}^2 z_{2j}^2 = 1 + o_p(1)$$

となるので，結果を得る． □

系 *1* の証明. 定理 3 と補題 3 より，系 1 が成り立つ. □

謝辞

科学研究費補助金 基盤研究 (A) 20H00576 研究代表者: 青嶋 誠「大規模複雑データの理論と方法論の革新的展開」，学術研究助成基金助成金 挑戦的研究 (萌芽) 22K19769 研究代表者: 青嶋 誠「テンソル構造をもつ巨大データの統計的圧縮技術の開発」，学術研究助成基金助成金 基盤研究 (C) 22K03412 研究代表者: 矢田 和善「非線形特徴量に基づく新たな高次元統計理論の開発とその応用」，および，学術研究助成基金助成金 若手研究 24K20749 研究代表者: 石井 晶「高次元従属標本の強スパイク性による統計的推測の構築」から研究助成を受けています．また，京都大学数理解析研究所の国際共同利用・共同研究拠点事業により研究助成を受けています.

参考文献

[1] M. Aoshima, K. Yata, Two-stage procedures for high-dimensional data. Sequential Anal. (*Editor's special invited paper*) 30 (2011) 356-399.

[2] M. Aoshima, K. Yata, Two-sample tests for high-dimension, strongly spiked eigenvalue models. Statist. Sinica 28 (2018) 43-62.

[3] A. Ishii, K. Yata, M. Aoshima, Hypothesis tests for high-dimensional covariance structures. Ann. Inst. Statist. Math. 73 (2021) 599-622.

[4] K. Yata, M. Aoshima, Effective PCA for high-dimension, low-sample-size data with noise reduction via geometric representations. J. Multivariate Anal. 105 (2012) 193-215.

[5] K. Yata, M. Aoshima, Correlation tests for high-dimensional data using extended cross-data-matrix methodology. J. Multivariate Anal. 117 (2013) 313-331.

[6] K. Yata, M. Aoshima, PCA consistency for the power spiked model in high-dimensional settings. J. Multivariate Anal. 122 (2013) 334-354.

[7] K. Yata, M. Aoshima, High-dimensional infrence on covariance structures via the extended cross-data-matrix methodology. J. Multivariate Anal. 151 (2016) 151-166.