

高次元相互共分散ベクトルの自動スパース推定

筑波大学・数理物質科学研究群 海野 哲也

Tetsuya Umino

Graduate School of Science and Technology

University of Tsukuba

筑波大学・数理物質系 矢田 和善

Kazuyoshi Yata

Institute of Mathematics

University of Tsukuba

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

本稿では、高次元特徴量の1つである高次元共分散ベクトルに対するスパース推定を考える．共分散ベクトル σ に対し、以下の検定を考える．

$$H_0 : \sigma = \mathbf{0} \quad \text{vs.} \quad H_1 : \sigma \neq \mathbf{0}. \quad (1)$$

Aoshima and Yata [1] はクロスデータ行列法により検定 (1) に対する検定統計量を与え、Yata and Aoshima [5] は拡張クロスデータ行列法 (ECDM 法) により、検定統計量の性能を向上させた．本稿では検定 (1) において H_0 が棄却された後に、 σ の推定を行うことを考える．

共分散行列全体に対するスパース推定は、これまでに広く研究されている．Bien and Tibshirani [3] は L1-罰則項に基づくスパース推定を提案し、Bickel and Levina [2] は閾値法によるスパース推定を提案した．一方で、高次元共分散ベクトルのスパース推定に焦点を当てた研究は、これまで限定的なものであった．

従来のスパース推定では、得られる結果が一致性を有するには適切な閾値や正則化パラメータを選択する必要があるものの、高次元においては計算コストや結果の安定性に関する問題が生じる．最近、Yata and Aoshima [7] は閾値や正則化パラメータを必要としない、高次元主成分ベクトルの新たなスパース推定の方法論を提案し、その漸近的性質を求めた．さらに、Umino et al. [4] は同手法を高次元相互共分散行列のスパース推定に応用した．本稿では、Yata and Aoshima [7] の方法論を高次元共分散ベクトルの推定に応用することで、閾値や正則化パラメータを用いずに、

$$\frac{\|\sigma - \hat{\sigma}\|^2}{\Delta} = o_P(1)$$

の意味での一致性を有する σ のスパース推定法を提案する．ただし、 $\|\cdot\|$ をユークリッドノルム

とし、 $\Delta = \|\sigma\|^2$ とする.

2 標本共分散ベクトルによる σ の推定

2.1 高次元の分布に関する仮定

平均に $p+1$ 次のベクトル μ , 共分散行列に $p+1$ 次の非負定値行列 Σ を持つ母集団から $n(\geq 4)$ 個の $p+1$ 次元データベクトル $\mathbf{x}_1, \dots, \mathbf{x}_n$ を無作為に抽出し, これらを $\mathbf{x}_j = (\mathbf{x}_{1j}^T, x_{j(*)})^T$ と分割する. ただし, $\mathbf{x}_{1j} = (x_{1j(1)}, \dots, x_{1j(p)})^T \in \mathbf{R}^p$, $x_{j(*)} \in \mathbf{R}$ である. $\mathbf{x}_{1j}$ は平均に $\mu_1 = (\mu_{1(1)}, \dots, \mu_{1(p)})^T$, 共分散行列に Σ_1 を持ち, $x_{j(*)}$ は平均に μ_* , 分散に σ_* を持つとする. さらに, $\mathbf{x}_{1j}$ と $x_{j(*)}$ の共分散ベクトルを $\text{Cov}(\mathbf{x}_{1j}, x_{j(*)}) = \sigma = (\sigma_1, \dots, \sigma_p)^T$ とおく. すなわち, $\mu = (\mu_1^T, \mu_*)^T$ かつ

$$\Sigma = \begin{pmatrix} \Sigma_1 & \sigma \\ \sigma^T & \sigma_* \end{pmatrix}$$

であるとする.

Σ の固有値分解を $\Sigma = \mathbf{H}\mathbf{\Lambda}\mathbf{H}^T$ とする. ここで, $\mathbf{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_{p+1})$ は Σ の固有値 $\lambda_1 \geq \dots \geq \lambda_{p+1} \geq 0$ からなる対角行列であり, $\mathbf{H}$ は対応する固有ベクトルからなる直交行列である. このとき, $\mathbf{E}(\mathbf{z}_j) = \mathbf{0}$ かつ $\text{Cov}(\mathbf{z}_j) = \mathbf{I}_{p+1}$ となる確率変数ベクトル $\mathbf{z}_j$ を用いて, $\mathbf{x}_j = \mathbf{H}\mathbf{\Lambda}^{1/2}\mathbf{z}_j + \mu$ と書くことができる. ただし, $t > 0$ に対し $\mathbf{I}_t$ は t 次の単位行列を表す. また, Σ_1 の第 j 対角成分を $\sigma_{1(j)}$ とおき, 各 j に対し $p \rightarrow \infty$ で $\sigma_{1(j)} \in (0, \infty)$ かつ $\sigma_* \in (0, \infty)$ なる正則条件を仮定する. ここで, 関数 $f(\cdot)$ に対し “ $p \rightarrow \infty$ で $f(p) \in (0, \infty)$ ” とは, $\liminf_{p \rightarrow \infty} f(p) > 0$ かつ $\limsup_{p \rightarrow \infty} f(p) < \infty$ であることを意味する. したがって, $p \rightarrow \infty$ で $\text{tr}(\Sigma_1)/p \in (0, \infty)$ である.

Yata and Aoshima [5, 6] に基づき, 本稿では各 $\mathbf{x}_j$ に以下のモデルを仮定する.

$$\mathbf{x}_j = \mathbf{\Gamma}\mathbf{w}_j + \mu.$$

ただし $q > 0$ をある自然数として, $\mathbf{\Gamma}$ は $\mathbf{\Gamma}\mathbf{\Gamma}^T = \Sigma$ を満たす $(p+1) \times q$ 行列であり, $\mathbf{w}_j = (w_{1j}, \dots, w_{qj})^T \in \mathbf{R}^q$, $j = 1, \dots, n$ は独立に $\mathbf{E}(\mathbf{w}_j) = \mathbf{0}$, $\text{Cov}(\mathbf{w}_j) = \mathbf{I}_q$ を持つ分布に従う確率変数ベクトルとする. また, $s = 1, \dots, q$ に対し $\text{Var}(w_{sj}^2) = M_s$ とし, $\limsup_{p \rightarrow \infty} M_s < \infty$ であるとする. さらに, $\mathbf{\Gamma} = (\mathbf{\Gamma}_1^T, \gamma_*^T)^T$, $\mathbf{\Gamma}_1 = (\gamma_{11}, \dots, \gamma_{1q})$, $\gamma_{1s} \in \mathbf{R}^p$ かつ $\gamma_* = (\gamma_{1*}, \dots, \gamma_{q*})^T \in \mathbf{R}^q$ とする. このとき,

$$\mathbf{x}_{1j} = \mathbf{\Gamma}_1\mathbf{w}_j + \mu_1, \quad x_{j(*)} = \gamma_*^T\mathbf{w}_j + \mu_* \quad \text{かつ} \quad \sigma = \mathbf{\Gamma}_1\gamma_* = \sum_{s=1}^q \gamma_{1s}\gamma_{s*}$$

とできることに注意する. 本稿を通じて, 以下を仮定する.

(A-i) 任意の $r \neq s, t, u$ を満たす $r, s, t, u = 1, \dots, q$ に対し,

$$\mathbf{E}(w_{rj}^2 w_{sj}^2) = \mathbf{E}(w_{rj}^2) \mathbf{E}(w_{sj}^2) = 1 \quad \text{かつ} \quad \mathbf{E}(w_{rj} w_{sj} w_{tj} w_{uj}) = 0.$$

(A-ii) 各 $r = 1, \dots, p$ に対し, $\limsup_{p \rightarrow \infty} E [| (x_{1j(r)} - \mu_{1(r)})(x_{j(*)} - \mu_*) - \sigma_r |^8] < \infty$.

x_j が正規分布に従う場合, (A-i) 及び (A-ii) は自然に成立することに注意する.

2.2 高次元における標本共分散ベクトルの漸近的性質

簡単のため, 以下では平均ベクトル $\boldsymbol{\mu}$ が既知であるとする. 標本共分散ベクトルを $\mathbf{s} = \sum_{j=1}^n (\mathbf{x}_{1j} - \boldsymbol{\mu}_1)(x_{j(*)} - \mu_*)/n$ とする. このとき, $E(\mathbf{s}) = \boldsymbol{\sigma}$ である. 必要に応じて, 以下を仮定する.

(C-i) $p, n \rightarrow \infty$ で $\frac{\text{tr}(\boldsymbol{\Sigma}_1^2)}{n^2 \Delta^2} = o(1)$.

$\text{tr}(\boldsymbol{\Sigma}_1^2) \leq \text{tr}(\boldsymbol{\Sigma}_1)^2$ が成立することより, (C-i) は $\text{tr}(\boldsymbol{\Sigma}_1)/(n\Delta) = o(1)$ よりも緩い条件であることに注意する. $\mathbf{s}$ について, Umino et al. [4] の定理 1 と命題 1 より以下の結果を得る.

命題 1. (C-i) と適当な条件のもと, $p, n \rightarrow \infty$ で以下が成立する.

$$\begin{aligned} \frac{\|\mathbf{s}\|^2}{\Delta} &= \left\{ \frac{n+1}{n} + \frac{\text{tr}(\boldsymbol{\Sigma}_1)\sigma_*}{n\Delta} + \frac{1}{n\Delta} \sum_{s=1}^q (M_s - 2) \|\gamma_{1s}\|^2 \gamma_{s*}^2 \right\} \{1 + o_P(1)\} \quad \text{かつ} \\ \frac{\boldsymbol{\sigma}^T \mathbf{s}}{\Delta} &= 1 + o_P(1). \end{aligned}$$

x_j が正規分布に従う場合, 各 s に対し $M_s = 2$ となることに注意する. この場合, 上の命題より

$$\frac{\|\boldsymbol{\sigma} - \mathbf{s}\|^2}{\Delta} = \left\{ \frac{\text{tr}(\boldsymbol{\Sigma}_1)\sigma_*}{n\Delta} \right\} \{1 + o_P(1)\} + o_P(1)$$

となるため, 条件 $\text{tr}(\boldsymbol{\Sigma}_1)/(n\Delta) = o(1)$ は $\mathbf{s}$ が一貫性を有するための必要条件となる. しかし, この条件は $p/(n\Delta) = o(1)$ であることを意味しており, $p/n \rightarrow \infty$ となる高次元小標本設定において非常に厳しいものとなる.

3 高次元共分散ベクトルのスパース推定

前節より, 標本共分散ベクトル $\mathbf{s}$ は高次元において巨大なノイズを内包することがわかる. 本節では Yata and Aoshima [7] で提案されたスパース推定のアイデアを応用することで, このノイズを除去することを考える.

$\hat{\Delta}$ を Yata and Aoshima [5] で提案された Δ の不偏推定量とし, 集合 $\mathcal{D} \subset \{1, \dots, p\}$ を

$$\mathcal{D} = \{r \mid \sigma_r \neq 0, r = 1, \dots, p\}$$

とおく. $\boldsymbol{\sigma}$ について, 以下のモデルを仮定する.

(C-ii) 各 $r \in \mathcal{D}$ に対し, $\liminf_{p \rightarrow \infty} \sigma_r^2 > 0$.

また、標本共分散ベクトルを $\mathbf{s} = (s_1, \dots, s_p)^T$ とおき、 $s_1, \dots, s_p$ を絶対値の大きい順に並べ替えたものを $s_{(1)}, \dots, s_{(p)}$ とおく。すなわち、 $|s_{(1)}| \geq \dots \geq |s_{(p)}|$ とする。命題 1 と Yata and Aoshima [5] の系 2.1 より、(A-i),(C-i) および $\liminf_{p \rightarrow \infty} \text{tr}(\Sigma_1)/(n\Delta) > 0$ のもと、漸近的に確率 1 で $\|\mathbf{s}\|^2 \geq \hat{\Delta}$ が成立する。このとき、

$$\sum_{r=1}^{k-1} \frac{s_{(r)}^2}{\hat{\Delta}} < 1 \quad \text{かつ} \quad \sum_{r=1}^k \frac{s_{(r)}^2}{\hat{\Delta}} \geq 1$$

となる自然数 $k(\leq p)$ が一意に存在する。そのとき、 $\mathbf{s}$ を次のようにスパース化する。

$$\tilde{\sigma} = (\tilde{s}_1, \dots, \tilde{s}_p)^T.$$

ただし、

$$\tilde{s}_r = \begin{cases} s_r & \text{if } |s_r| \geq |s_{(k)}| \\ 0 & \text{otherwise} \end{cases}, \quad r = 1, \dots, p \quad (2)$$

とする。このとき、次の定理を得る。

定理 1. (A-i),(A-ii),(C-i),(C-ii) と、以下の (C-iii) を仮定する。

(C-iii) $p, n \rightarrow \infty$ で $\frac{p}{n^4} = o(1)$.

このとき、 $p, n \rightarrow \infty$ で以下が成立する。

$$\frac{\|\sigma - \tilde{\sigma}\|^2}{\Delta} = o_P(1).$$

注意. $\|\mathbf{s}\|^2 < \hat{\Delta}$ である場合、(2) において $k = p$ とする。

注意. 通常、平均ベクトル μ は未知であり、 σ の推定には不偏推定量 $\mathbf{s}_* = \sum_{j=1}^n (\mathbf{x}_{1j} - \bar{\mathbf{x}}_1)(x_{j(*)} - \bar{x}_*)/(n-1)$ が用いられる。ただし、 $\bar{\mathbf{x}}_1 = \sum_{j=1}^n \mathbf{x}_{1j}/n$ かつ $\bar{x}_* = \sum_{j=1}^n x_{j(*)}/n$ である。本稿では割愛するが、 $\mathbf{s}_*$ を上記と同様の手順でスパース化することで得られる $\tilde{\sigma}_*$ についても、定理 1 と同様の結果を得ることが可能である。

4 シミュレーション

本節では高次元小標本のもとで、 $\mathbf{s}$ と $\tilde{\sigma}$ の精度を数値的に検証する。 $p = q$, $\mu = \mathbf{0}$, $\Gamma = \mathbf{H}\Lambda^{1/2}$ とおき、 Σ_1 の固有値分解を $\Sigma_1 = \mathbf{H}_1\Lambda_1\mathbf{H}_1^T$ とする。ここで、 $\Lambda_1 = \text{diag}(\lambda_{11}, \dots, \lambda_{1p})$ は Σ_1 の固有値 $\lambda_{11} \geq \dots \geq \lambda_{1p} \geq 0$ からなる対角行列であり、 $\mathbf{H}_1$ は対応する固有ベクトルからなる直交行列である。 Σ_1 について、次の 2 つの設定を考える。

(a) $p = 2^s$, $s = 6, \dots, 11$, $n = \lceil p^{2/3} \rceil$ とおく。

$$\Sigma_1 = \begin{pmatrix} \Gamma_5 & \mathbf{O} \\ \mathbf{O} & I_{p-5} \end{pmatrix} \quad (= \Sigma_{1a})$$

とおく. ただし, $\lceil x \rceil$ は x 以上の最小の整数を表し, $t > 0$ に対し $\mathbf{1}_t = (1, \dots, 1)^T \in \mathbf{R}^t$ とし $\mathbf{\Gamma}_t = (\mathbf{I}_t + \mathbf{1}_t \mathbf{1}_t^T)/2$ とする.

(b) $p = 500$, $n = 2^s$, $s = 3 \dots, 8$ とおく.

$$\mathbf{\Sigma}_1 = \begin{pmatrix} \mathbf{\Gamma}_{4\lfloor p^{1/3} \rfloor} & \mathbf{O} \\ \mathbf{O} & \mathbf{I}_{p-4\lfloor p^{1/3} \rfloor} \end{pmatrix} \quad (= \mathbf{\Sigma}_{1b})$$

とおく. ただし, $\lfloor x \rfloor$ は x 以下の最大の整数を表す.

上記 2 つの設定のもと, $\mathbf{x}_{1j}$ と $x_{j(*)}$ を

$$\mathbf{x}_{1j} = \mathbf{H}_1 \mathbf{\Lambda}_1^{1/2} (w_{1j}, \dots, w_{pj})^T, \quad x_{j(*)} = w_{1j} \quad (3)$$

で発生させ, $\mathbf{x}_j$ には以下の 3 つの分布を考える.

(I) $\mathcal{N}_{p+1}(\mathbf{0}, \mathbf{\Sigma})$.

(II) $w_{rj} = (\nu_{rj} - 5)/\sqrt{10}$. ただし, ν_{rj} , $r = 1, \dots, p$ は独立に自由度 5 のカイ二乗分布に従う確率変数.

(III) $w_{rj} = t_{rj}/\sqrt{10/8}$. ただし, t_{rj} , $r = 1, \dots, p$ は独立に自由度 10 の t 分布に従う確率変数.

(A-i) は全ての設定において満たされることに注意する. 一方で, (A-ii) は (I) と (II) では満たされ, (III) では満たされないことに注意する. また (3) のもと, $\boldsymbol{\sigma}$ は設定 (a) では最初の 5 個の成分が非ゼロであるベクトル $\boldsymbol{\sigma} = (\sqrt{\lambda_{11}/5}, \dots, \sqrt{\lambda_{11}/5}, 0, \dots, 0)^T$ となり, 設定 (b) では最初の $4\lfloor p^{1/3} \rfloor$ 個の成分が非ゼロであるベクトル $\boldsymbol{\sigma} = (\sqrt{\lambda_{11}/(4\lfloor p^{1/3} \rfloor)}, \dots, \sqrt{\lambda_{11}/(4\lfloor p^{1/3} \rfloor)}, 0, \dots, 0)^T$ となる.

p 次元ベクトル $\mathbf{a}$ に対し, $e(\mathbf{a}) = \|\mathbf{a} - \boldsymbol{\sigma}\|^2/\Delta$ とする. 上記の設定のもと, $e(\mathbf{s})$ と $e(\tilde{\boldsymbol{\sigma}})$ をそれぞれ 1000 回発生させ, その平均 (左図) と分散 (右図) を図 1 から図 3 にプロットした. これらの図からもわかるように, スパースな推定量 $\tilde{\boldsymbol{\sigma}}$ が従来の推定量 $\mathbf{s}$ に比べ, 高次元小標本のもと非常に良い推定量となっている. ここでは割愛するが, 設定を変えて実験をしたときにも, $\tilde{\boldsymbol{\sigma}}$ が $\mathbf{s}$ に比べ, 高次元小標本のもと優れた推定量となっていることが確認されている.

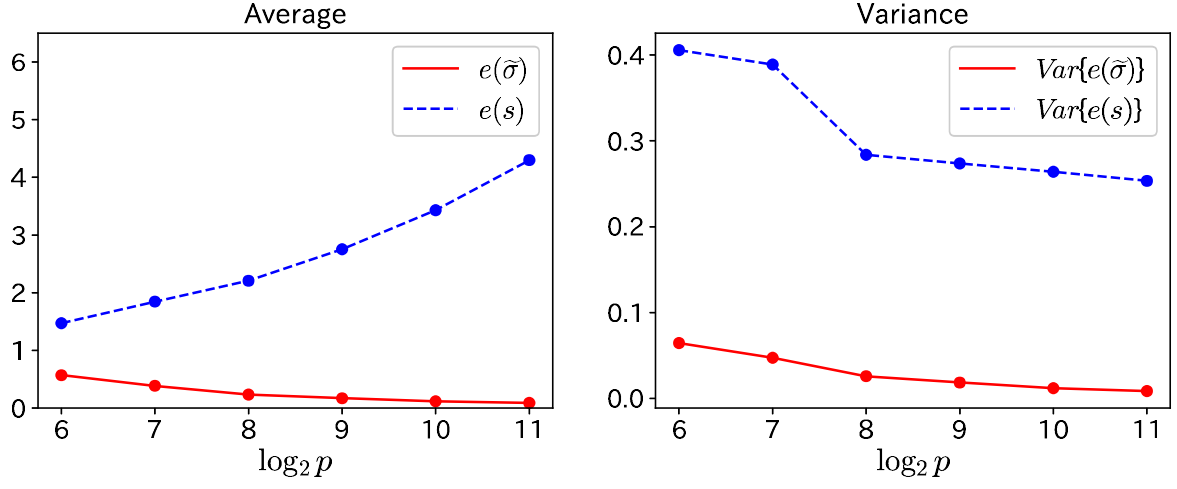
5 定理 1 の証明

簡単のために $|\sigma_1| \geq \dots \geq |\sigma_p|$ であるとする. また, $d = |\mathcal{D}|$ とする. ただし, 有限集合 A に対しその要素数を $|A|$ で表す. マルコフの不等式より, 任意の $\tau > 0$ に対し (A-ii) と (C-iii) の仮定のもと

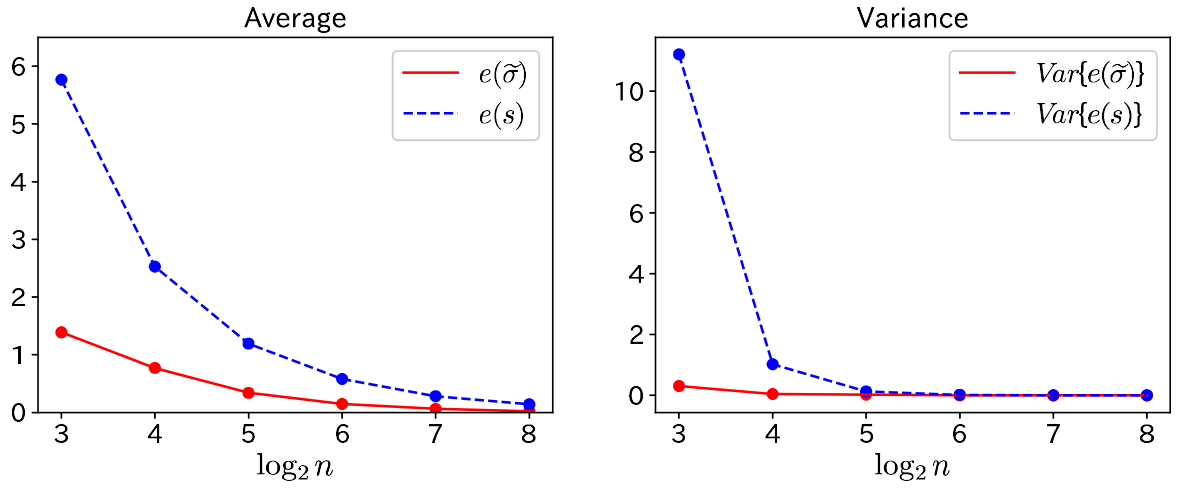
$$\sum_{r=1}^p \mathbb{P}(|s_r - \sigma_r| \geq \tau) \leq \sum_{r=1}^p \frac{\mathbb{E}(|s_r - \sigma_r|^8)}{\tau^8} = O\left(\frac{p}{n^4}\right) = o(1)$$

となるため,

$$\max_{r=1, \dots, p} |s_r - \sigma_r| = o_P(1)$$



(a) $p = 2^s$, $s = 6, \dots, 11$, $n = \lceil p^{2/3} \rceil$, $\Sigma_1 = \Sigma_{1a}$.



(b) $p = 500$, $n = 2^s$, $s = 3, \dots, 8$, $\Sigma_1 = \Sigma_{1b}$.

図 1: $\mathbf{x}_j \sim \mathcal{N}_{p+1}(\mathbf{0}, \Sigma)$.

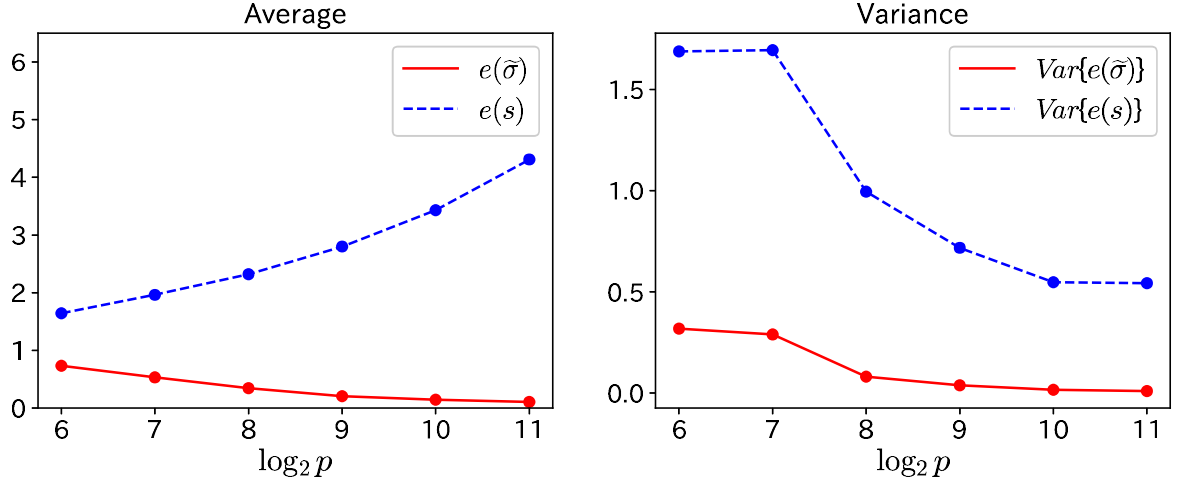
が成立する。よって (C-ii) の仮定より

$$s_r^2 = \begin{cases} \sigma_r^2 \{1 + o_P(1)\} & r \leq d, \\ o_P(\sigma_d^2) & r \geq d+1 \end{cases} \quad (4)$$

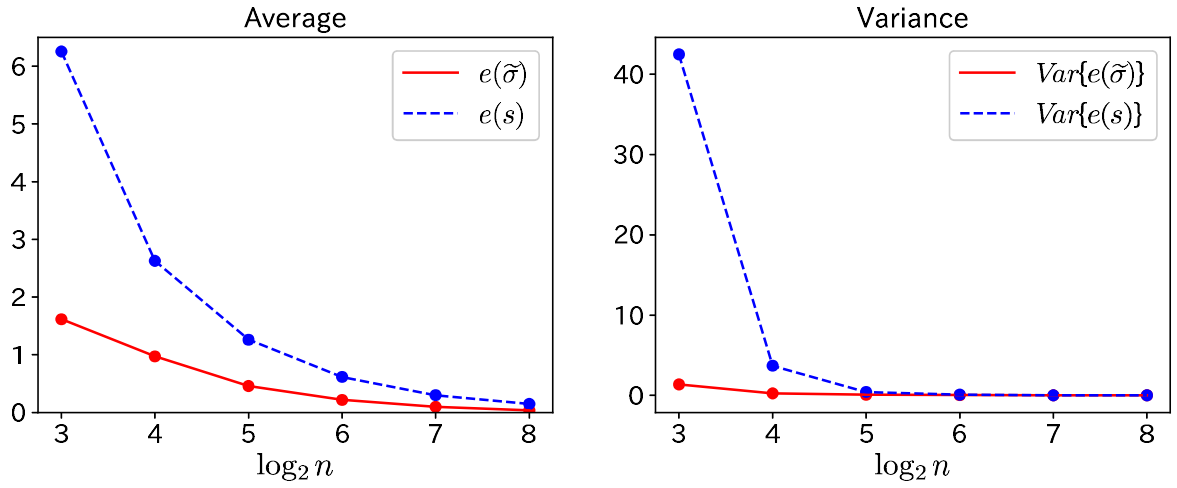
となるため, $p, n \rightarrow \infty$ で

$$\sum_{r \in \mathcal{D}} \frac{s_r^2}{\Delta} = 1 + o_P(1) \quad \text{かつ} \quad \sum_{r \in \mathcal{D}} \frac{s_r \sigma_r}{\Delta} = 1 + o_P(1) \quad (5)$$

が成立する。



(a) $p = 2^s$, $s = 6, \dots, 11$, $n = \lceil p^{2/3} \rceil$, $\Sigma_1 = \Sigma_{1a}$.

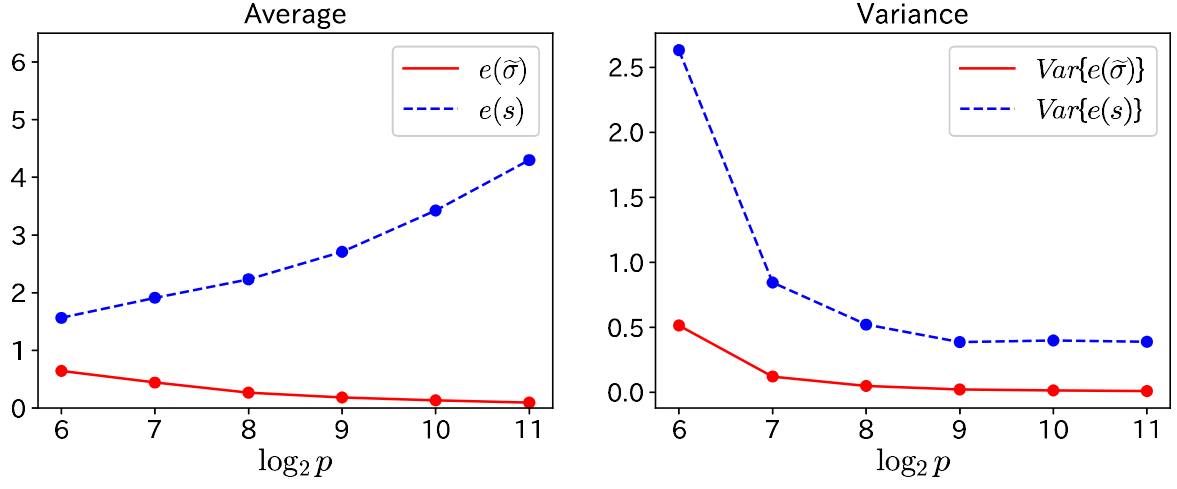


(b) $p = 500$, $n = 2^s$, $s = 3, \dots, 8$, $\Sigma_1 = \Sigma_{1b}$.

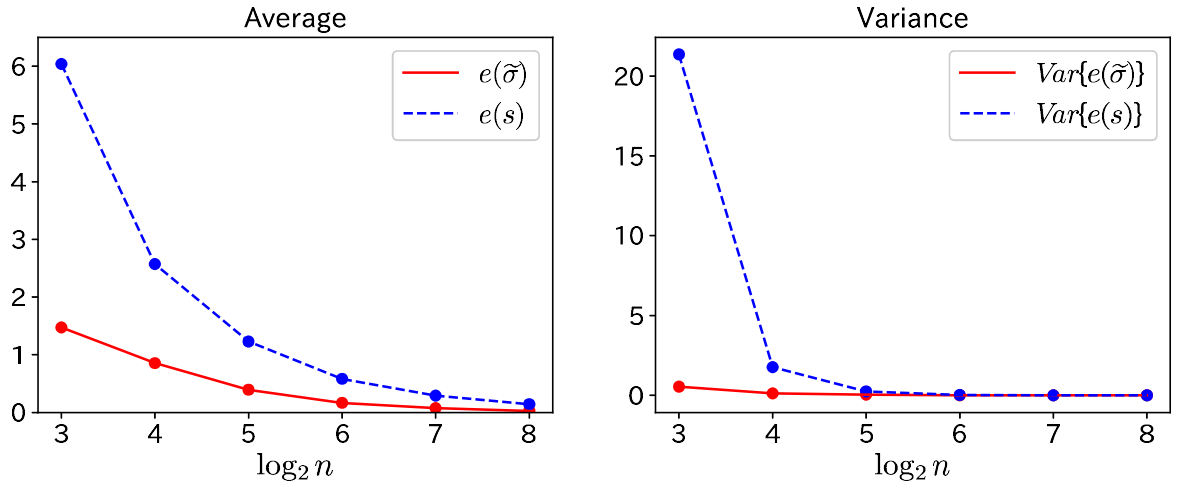
図 2: $w_{rj} = (\nu_{rj} - 5)/\sqrt{10}$. ただし, ν_{rj} , $r = 1, \dots, p$ は独立に自由度 5 のカイ二乗分布に従う確率変数.

Yata and Aoshima [5] の系 2.1 より, (A-i) と (C-i) の仮定のもと, $p, n \rightarrow \infty$ で $\hat{\Delta}/\Delta = 1 + o_P(1)$ が成立する. さらに, 集合 $\mathcal{D}_1 \subset \{1, \dots, p\}$ を $\mathcal{D}_1 = \{r \mid \tilde{s}_r \neq 0, r = 1, \dots, p\}$ とする. このとき $|\mathcal{D}_1| = k$ となり, (4) より漸近的に確率 1 で

$$\min_{1 \leq r \leq d} s_r^2 > \max_{d+1 \leq r \leq p} s_r^2$$



(a) $p = 2^s$, $s = 6, \dots, 11$, $n = \lceil p^{2/3} \rceil$, $\Sigma_1 = \Sigma_{1a}$.



(b) $p = 500$, $n = 2^s$, $s = 3, \dots, 8$, $\Sigma_1 = \Sigma_{1b}$.

図 3: $w_{rj} = t_{rj}/\sqrt{10/8}$. ただし, t_{rj} , $r = 1, \dots, p$ は独立に自由度 10 の t 分布に従う確率変数.

となるため, 集合 $\mathcal{D}, \mathcal{D}_1$ について, 漸近的に確率 1 で次が成立する.

$$\mathcal{D} \subset \mathcal{D}_1 \quad \text{and} \quad s_{(k)} \in \{s_{d+1}, \dots, s_p\} \quad \text{if} \quad k > d, \quad (6)$$

$$\mathcal{D}_1 \subset \mathcal{D} \quad \text{if} \quad k \leq d. \quad (7)$$

$k > d$ であるとき, (4) から (6) より

$$\sum_{r \in \mathcal{D}_1} \frac{s_r^2}{\Delta} \leq \frac{\hat{\Delta}}{\Delta} + \frac{s_{(k)}^2}{\Delta} = 1 + o_P(1) \quad \text{かつ} \quad \sum_{r \in \mathcal{D}_1} \frac{s_r \sigma_r}{\Delta} = \sum_{r \in \mathcal{D}} \frac{s_r \sigma_r}{\Delta} = 1 + o_P(1)$$

となる。よって

$$\frac{\|\sigma - \tilde{\sigma}\|^2}{\Delta} = \sum_{r \in \mathcal{D}_1} \frac{s_r^2}{\Delta} + 1 - 2 \sum_{r \in \mathcal{D}_1} \frac{s_r \sigma_r}{\Delta} = o_P(1) \quad (8)$$

が成立する。

$k \leq d$ であるとき, (5),(7) とシュワルツの不等式より

$$\sum_{r \in \mathcal{D} \setminus \mathcal{D}_1} \frac{s_r^2}{\Delta} \leq \sum_{r \in \mathcal{D}} \frac{s_r^2}{\Delta} - \frac{\hat{\Delta}}{\Delta} = o_P(1) \quad \text{かつ} \quad \sum_{r \in \mathcal{D} \setminus \mathcal{D}_1} \frac{s_r \sigma_r}{\Delta} \leq \sqrt{\sum_{r \in \mathcal{D} \setminus \mathcal{D}_1} \frac{s_r^2}{\Delta}} \sqrt{\sum_{r \in \mathcal{D} \setminus \mathcal{D}_1} \frac{\sigma_r^2}{\Delta}} = o_P(1)$$

が成立する。したがって (5) と組み合わせることで

$$\begin{aligned} \sum_{r \in \mathcal{D}_1} \frac{s_r^2}{\Delta} &= \sum_{r \in \mathcal{D}} \frac{s_r^2}{\Delta} - \sum_{r \in \mathcal{D} \setminus \mathcal{D}_1} \frac{s_r^2}{\Delta} = 1 + o_P(1) \quad \text{かつ} \\ \sum_{r \in \mathcal{D}_1} \frac{s_r \sigma_r}{\Delta} &= \sum_{r \in \mathcal{D}} \frac{s_r \sigma_r}{\Delta} - \sum_{r \in \mathcal{D} \setminus \mathcal{D}_1} \frac{s_r \sigma_r}{\Delta} = 1 + o_P(1) \end{aligned}$$

となることから (8) が成立し, 題意を得る。

謝辞

本研究は, 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 挑戦的萌芽研究 (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」, および, 科学研究費補助金 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」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・共同研究拠点事業より研究助成を受けています。

参考文献

- [1] Aoshima, M. and Yata, K. (2011). Two-stage procedures for high-dimensional data. *Sequential Analysis (Editor's special invited paper)*, 30(4):356–399.
- [2] Bickel, P. J. and Levina, E. (2008). Covariance regularization by thresholding. *The Annals of Statistics*, 36(6):2577 – 2604.
- [3] Bien, J. and Tibshirani, R. J. (2011). Sparse estimation of a covariance matrix. *Biometrika*, 98(4):807–820.
- [4] Umino, T., Yata, K., and Aoshima, M. (2025). Automatic sparse estimation of the high-dimensional cross-covariance matrix. *submitted*.
- [5] Yata, K. and Aoshima, M. (2013). Correlation tests for high-dimensional data using extended cross-data-matrix methodology. *Journal of Multivariate Analysis*, 117:313–331.

- [6] Yata, K. and Aoshima, M. (2016). High-dimensional inference on covariance structures via the extended cross-data-matrix methodology. *Journal of Multivariate Analysis*, 151:151–166.
- [7] Yata, K. and Aoshima, M. (2025). Automatic sparse pca for high-dimensional data. *Statistica Sinica*, 35:1069–1090.