

久留米大学バイオ統計センター (2010 年 01 月 21 日)

# **Shrinkage estimators for large covariance matrices in multivariate normal distribution**

日本女子大学理学部 今野 良彦

January 19, 2010

## この講演の目的と構成

近年，データ数よりも変量の次元が高いデータ（高次元データ）の解析のための多変量推測理論の構築が注目を集めている．本講演では，高次元データの設定のもとで多変量正規分布の共分散行列 (Large Covariance Matrix) の推定問題を統計的決定理論の枠組みで考察した結果を報告する．

### 本講演の構成

- (1) 問題の背景・記号・問題設定.
- (2) 先行研究について．
- (3) リスクの不偏推定法 (SURE) ．
- (4) 改良型推定量について．
- (5) 数値実験について

## Real Wishart 分布 ( 1 )

$\mathbb{R}^p$ -値確率ベクトル  $X_1, X_2, \dots, X_n$  は独立同一に多変量正規分布  $N_p(0, \Sigma)$  に従う :

$$P(dx) = \frac{1}{(2\pi)^{p/2} \text{Det}(\Sigma)^{1/2}} \exp \left\{ -\frac{1}{2} x' \Sigma^{-1} x \right\} (dx)$$

ただし ,

$$\Sigma \in \text{Sym}_p^+(\mathbb{R}) = \{ \mathbf{S} \in \mathbb{R}^{p \times p} : x' \mathbf{S} x > 0 \text{ for } x \in \mathbb{R}^p, x \neq 0 \},$$

$dx$  は  $\mathbb{R}^p$  上のルベーグ測度 ,  $x'$  は縦ベクトル  $x$  の転置 .

## Real Wishart 分布 ( 2 )

Wishart Matrix

$$\mathbf{W} : p \times p = \sum_{j=1}^n X_j X_j'$$

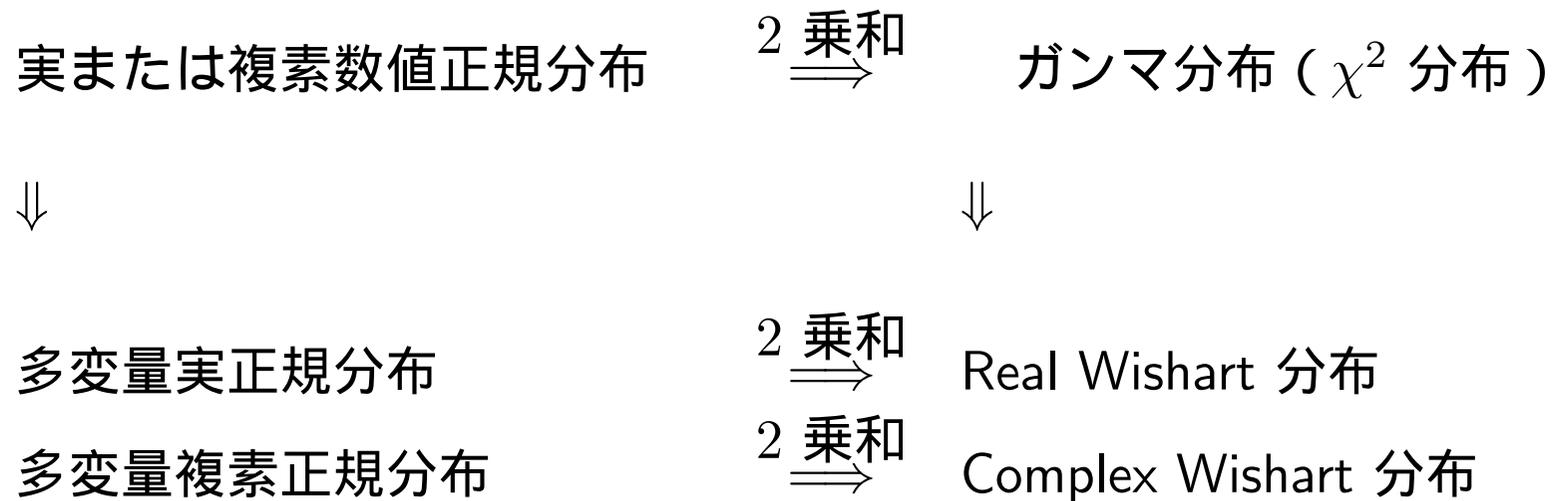
とおく .  $\mathbf{W}$  をウイシャート行列と呼ぶことにする .

定義より ,  $\mathbf{W} \in \overline{\text{Sym}_p^+(\mathbb{R})}$  (  $\text{Sym}_p^+(\mathbb{R})$  の閉包 ) であるが ,

$$\mathbb{P}\{\mathbf{W} \in \text{Sym}_p^+(\mathbb{R})\} = 1 \iff n \geq p \quad \text{かつ} \quad \Sigma \in \text{Sym}_p^+(\mathbb{R})$$

となる . Eaton (1983, page 304) を参照 .

## Real Wishart 分布 ( 3 )



## Real Wishart 分布 ( 4 )

Real Wishart density ( $\Sigma \in \text{Sym}_p^+(\mathbb{R})$ ,  $n \geq p$  を仮定 . )

$$W_p(\Sigma, n)(dx) = \frac{1}{c_{p,n}(\text{Det}\Sigma)^{n/2}} (\text{Det}x)^{n/2} \exp\left(-\frac{1}{2}\text{Tr}(\Sigma^{-1}x)\right) \\ \times \mathbb{1}_{\text{Sym}_p^+(\mathbb{R})}(x) \frac{(dx)}{(\text{Det}x)^{(p+1)/2}}$$

この分布を  $W_p(\Sigma, n)$  とかく . ただし ,

$$c_{p,n} = 2^{np} \pi^{p(p-1)/4} \prod_{i=0}^{p-1} \Gamma((n-i)/2)$$

$$\mathbb{1}_{\text{Sym}_p^+(\mathbb{R})}(x) = \begin{cases} 1 & (x \in \text{Sym}_p^+(\mathbb{R})), \\ 0 & (\text{その他}) \end{cases}$$

## Wishart 行列の極分解

極分解 ( $n \geq p$  の制限は必要なし)

$$W = HLH', \quad H \in O_p(\mathbb{R}), \quad L = \text{diag}(\ell_1, \ell_2, \dots, \ell_p)$$

$$\text{with } \ell_1 \geq \ell_2 \geq \dots \geq \ell_p \geq 0$$

ただし,

$$O_p(\mathbb{R}) = \{H \in \mathbb{R}^{p \times p} : H'H = HH' = I_p\}$$

で  $I_p$  は  $p \times p$  の単位行列とする.

# Real Wishart 行列の固有根の分布 ( 1 )

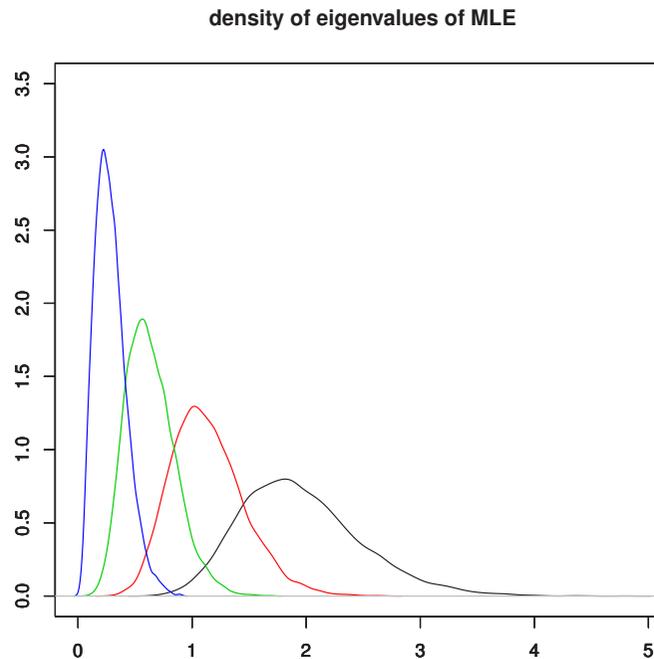


Figure 1: シミュレーションによる  $W_4(\mathbf{I}_4, 10)$  の  $(\ell_1, \dots, \ell_4)/10$  の周辺確率密度関数 ( $p = 4, n = 10$ ) の作図.

## Real Wishart 行列の固有根の分布 ( 2 )

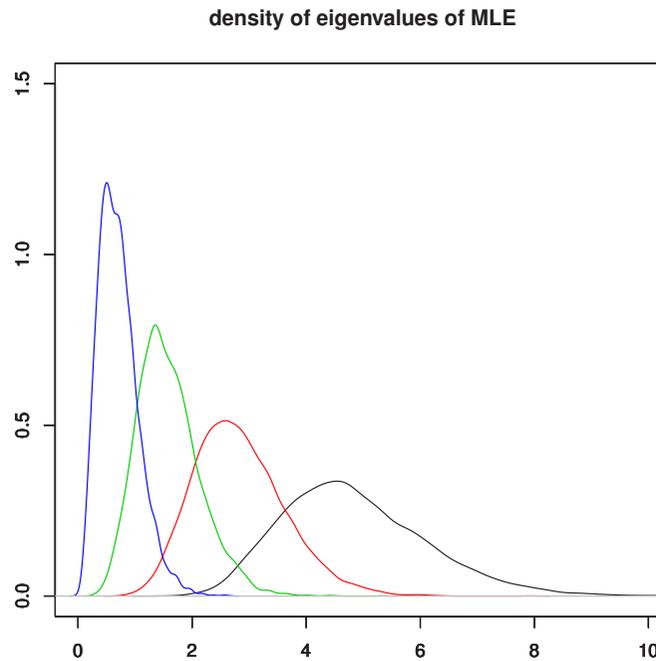


Figure 2: シミュレーションによる  $W_{10}(\mathbf{I}_{10}, 4)$  の  $(\ell_1, \dots, \ell_{10})/4$  の周辺確率密度関数 ( $p = 10, n = 4$ ) の作図.

# Real Wishart 行列の固有根の分布 ( 3 )

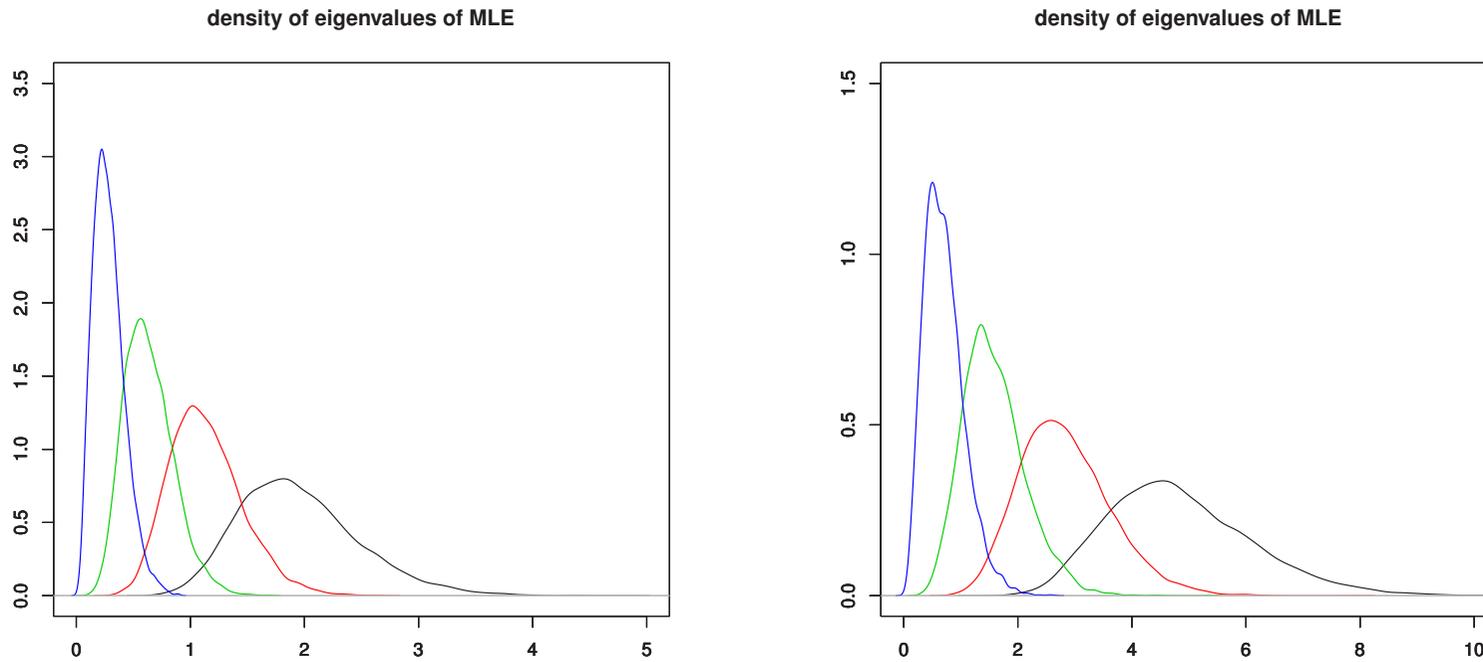


Figure 3: 左は  $W_4(\mathbf{I}_4, 10)$  ; 右は  $W_{10}(\mathbf{I}_{10}, 4)$

## Real Wishart 行列の固有根の分布 ( 4 ) : from RMT view

Marčenko and Pastur law[1]

$p = p(n)$  とし ,  $p/n \rightarrow \tau \leq 1 (n \rightarrow \infty)$  のとき ,

$$\frac{1}{p} \sum_{i=1}^p \mathbb{1}_{\lambda_i \leq x} \rightarrow \int_{-\infty}^x \frac{\sqrt{(b-t)(t-a)}}{2\pi t\tau} \mathbb{1}_{a \leq t \leq b} dt$$

(almost surely).

ただし ,  $a = (1 - \tau^{1/2})^2$  ,  $b = (1 + \tau^{1/2})^2$ .

[17] の Lecture 4 を参照 .

## Real Wishart 行列の固有根の分布 ( 5 )

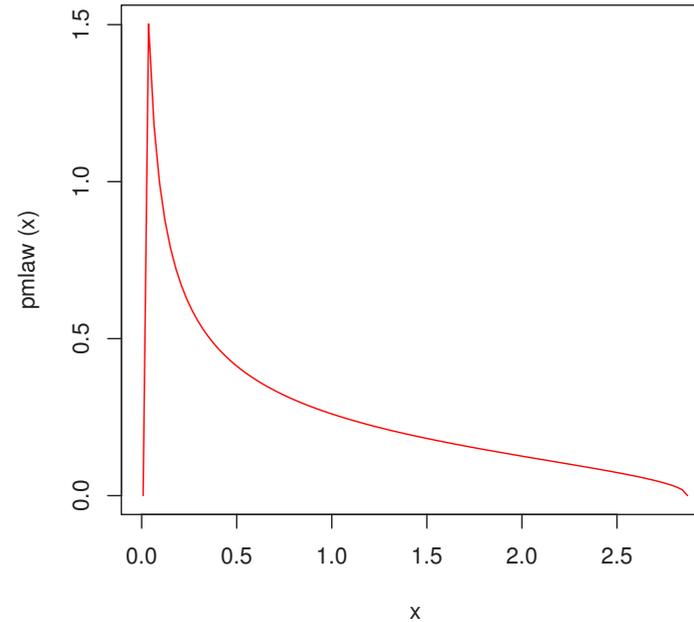


Figure 4: Density of Marčenko-Pastur's quater-circle law ( $\tau = 4/7 \doteq 0.57$ ,  $a \doteq 0.06$ ,  $b \doteq 3.086$ ).

## Real Wishart 行列の固有根の周辺分布 ( 6 )

- $p \times p$  の正値対称行列  $W \sim W_p(\Sigma, n)$  ( $\Sigma > 0$  かつ  $n \geq p$ ) の固有根を  $\ell_1 \geq \ell_2 \geq \dots \geq \ell_p$  とする .  $W/n$  は  $\Sigma$  の不偏推定量 .
- しかし ,  $W/n$  の固有根  $(\ell_1/n, \dots, \ell_p/n)$  は  $\Sigma$  の固有根のバイアス推定量となる !



Shrink eigenvalues toward a center!

- $n < p$  の場合はどうなるのか ?

## 記号と問題設定 ( 1 )

★  $X_1, X_2, \dots, X_n \sim N_p(0, \Sigma)$  . 各  $X_i (i = 1, 2, \dots, n)$  は  $p$  変量確率 (縦) ベクトルで, 独立同一に  $p$  変量正規分布に従う. ただし,  $\Sigma$  は  $p \times p$  の正値対称行列で未知.

★  $n$  は (標本数 - 1) で,  $p$  は変量の次元.

★ Wishart 確率行列 ( $p \times p$  の行列)  $W := \sum_{k=1}^n X_k X_k'$  . ただし, “'” はベクトルや行列の転置を示す.

★ 共分散行列  $\Sigma$  の推定問題を損失関数

$$L(\hat{\Sigma}, \Sigma) = \text{Tr} (\hat{\Sigma} \Sigma^{-1} - I_p)^2$$

のもとで考える. ここで,  $\hat{\Sigma}$  は  $\Sigma$  の推定量,  $I_p$  は  $p \times p$  の単位行列,  $\text{Tr}$  は行列のトレースを表す.

★  $W$  の分布に関する損失関数  $L$  の期待値  $R(\hat{\Sigma}, \Sigma) := \mathbb{E}[L(\hat{\Sigma}, \Sigma)]$  をリスクとよぶ.  $\Sigma$  に関して一様に推定量のリスクを比較したい.

## 記号と問題設定 ( 2 )

- ★ 平均を 0 としたことは本質的ではない .
- ★ Wishart 確率行列  $W$  は正定値  $\iff n \geq p$  .
- ★ Wishart 確率行列  $W$  の分布は常に存在するが ,  $p \times p$  の対称行列の空間上の Lebesgue 測度に関する確率密度関数は  $n \geq p$  のとき存在 .
- ★ 変換  $\hat{\Sigma} \mapsto A\hat{\Sigma}A'$ ;  $\Sigma \mapsto A\Sigma A'$  ( $A$  は  $p \times p$  の正則行列) に関して不変な損失関数 :

$$L(\hat{\Sigma}, \Sigma) = \text{Tr}(\hat{\Sigma}\Sigma^{-1} - I_p)^2; L_S(\hat{\Sigma}, \Sigma) = \text{Tr}(\hat{\Sigma}\Sigma^{-1}) - \log \text{Det}(\hat{\Sigma}\Sigma^{-1}) - p.$$

ただし ,  $\text{Det}$  は行列式 . しかし ,  $n < p$  のとき ,  $L_S$  ( $L_S$  の期待値) は  $n^{-1}W$  を評価できない .

## 先行研究について ( 1 )

推定量  $n^{-1}W$  の問題点

- ★  $E[n^{-1}W] = \Sigma$  だが,  $n^{-1}W$  の固有根は,  $\Sigma$  の固有根よりも広がっている.
- ★  $n < p$  のとき,  $\Sigma$  は正定値であるにもかかわらず,  $n^{-1}W$  は正定値ではない.

 $n \geq p$  の場合の先行研究

- ★ 損失関数  $L_S$  のもとでは,  $n^{-1}W$  の固有根を Shrinkage-expansion method を用いた改良型推定量. Stein (1977), Dey and Srivastava (1985), Haff (1991) 等を参照.
- ★ リスクを評価するために, SURE 法が有効 — 部分積分の公式と eigenvalue-calculus →  $n < p$  の場合は?
- ★ 損失関数  $L$  のもとでは, Haff (1980) の結果 →  $n < p$  の場合は?

## 先行研究について ( 2 )

### $n < p$ の場合の先行研究

- ★ Ledoit and Wolf (2004): 損失関数  $\text{Tr}(\hat{\Sigma} - \Sigma)^2$  のもとで,  $n^{-1}W$  と  $I_p$  の線形結合のなかで漸近的 ( $n/p$  は有界) に最適なもの. 積率の条件のみで分布に依存しない結果.
- ★ Wu and Pourahmadi (2003), Bickel and Levina (2008): banding approach. 漸近的に評価.
- ★ Furrer and Bengtsson (2007): “tapering”.

## 問題設定の復習

★  $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n \sim N_p(\mathbf{0}, \Sigma)$  . 各  $\mathbf{X}_i (i = 1, 2, \dots, n)$  は  $p$  変量確率 (縦) ベクトルで, 独立同一に  $p$  変量正規分布に従う . ただし,  $\Sigma$  は  $p \times p$  の正値対称行列で未知 .

★  $n$  は (標本数 - 1) で,  $p$  は変量の次元;

★ Wishart 確率行列 ( $p \times p$  の行列)  $W := \sum_{k=1}^n \mathbf{X}_k \mathbf{X}_k'$  に基づき, 共分散行列  $\Sigma$  の推定問題を損失関数

$$L(\hat{\Sigma}, \Sigma) = \text{Tr}(\hat{\Sigma}\Sigma^{-1} - I_p)^2$$

のもとで考える . ここで,  $\hat{\Sigma}$  は  $\Sigma$  の推定量;

★  $W$  の分布に関する損失関数  $L$  の期待値  $R(\hat{\Sigma}, \Sigma) := \mathbb{E}[L(\hat{\Sigma}, \Sigma)]$  をリスクとよぶ .  $\Sigma$  に関して一様に推定量のリスクを比較したい .

## 推定量のクラス

★  $W = \sum_{i=1}^n X_i X_i'$  を分解する :  $\ell_1 \geq \dots \geq \ell_n$  は  $W$  の固有値で ,

$$W = O_1 L O_1', \quad L = \text{Diag}(\ell_1, \dots, \ell_n);$$

$O_1$  は  $p \times n$  の半直交行列 s.t.  $O_1' O_1 = I_n$ .

推定量のクラス

$$\hat{\Sigma} = O_1 \Psi(L) O_1', \quad (1)$$

ただし ,  $\Psi := \Psi(L) = \text{Diag}(\psi_1, \psi_2, \dots, \psi_n)$  で  $\psi_k := \psi_k(L) (k = 1, 2, \dots, n)$  は  $\mathbb{R}_{\geq}^n$  から  $\mathbb{R}$  への可微分関数.

目標

$\Sigma$  に依存するリスク  $\mathbb{E}[\text{Tr}(\hat{\Sigma} \Sigma^{-1} - I_p)^2]$  を評価したい!

## 推定量のクラスとリスクの評価の方針 ( SURE 法)

★ リスク  $\mathbb{E}[\text{Tr}(\hat{\Sigma}\Sigma^{-1} - I_p)^2]$  の不偏推定量  $\hat{R}(\hat{\Sigma})$  ( $\psi_1, \dots, \psi_n$  と  $\ell_1, \dots, \ell_n$  を通してのみ  $W$  に依存) を導出:

$$\mathbb{E}[\text{Tr}(\hat{\Sigma}\Sigma^{-1} - I_p)^2] = \mathbb{E}[\hat{R}(\hat{\Sigma})]$$

★  $\mathbb{E}[\text{Tr}(n^{-1}W\Sigma^{-1} - I_p)^2]$  は定数リスクなので,

$$\hat{R}(\hat{\Sigma}) \leq \mathbb{E}[\text{Tr}(n^{-1}W\Sigma^{-1} - I_p)^2]$$

ならば,

$$\mathbb{E}[\text{Tr}(\hat{\Sigma}\Sigma^{-1} - I_p)^2] \leq \mathbb{E}[\text{Tr}(n^{-1}W\Sigma^{-1} - I_p)^2]$$

がわかる.

SURE の導出 推定量の族 (1) に対して, リスクの不偏推定量  $\hat{R}(\hat{\Sigma})$  を導出する.

## 部分積分の公式と SURE 法 ( 1 )

★  $(x_{ij})_{i=1, \dots, n; j=1, \dots, p} := [\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n]' \sim N_{n \times p}(\mathbf{0}, \mathbf{I}_n \otimes \boldsymbol{\Sigma});$

★  $n \times p$  の行列作用素  $\nabla_{\mathbf{X}}$  を次で定める :

$$\nabla_{\mathbf{X}} = \left( \frac{\partial}{\partial x_{ij}} \right)_{\substack{i=1, 2, \dots, n \\ j=1, 2, \dots, p}} ;$$

★ 行列  $\nabla_{\mathbf{X}} \mathbf{A}$  の  $(i, j)$  成分を

$$(\nabla_{\mathbf{X}} \mathbf{A})_{ij} = \sum_{k=1}^p \frac{\partial a_{kj}}{\partial x_{ik}} \quad \text{for } i = 1, 2, \dots, n; j = 1, 2, \dots, p.$$

## 部分積分の公式と SURE 法 ( 2 )

**補題 1**  $[\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n]' \sim N_{n \times p}(\mathbf{0}, \mathbf{I}_n \otimes \Sigma)$  とし,  $\mathbf{W} = \sum_{i=1}^n \mathbf{X}_i \mathbf{X}_i'$  とおいたとき,  $p \times p$  関数  $G = G(\mathbf{W})$  に対して,

$$\mathbb{E}[\Sigma^{-1} \mathbf{W} G] = \mathbb{E}[nG + (\mathbf{X}' \nabla_{\mathbf{X}})' G].$$

特に,

$$\mathbb{E}[\text{Tr}(\Sigma^{-1} \mathbf{W} G)] = \mathbb{E}[n \text{Tr}(G) + \text{Tr}(\mathbf{X}' \nabla_{\mathbf{X}} G')].$$

ただし, ' は転置.

## 部分積分の公式と SURE 法 ( 3 )

★ 補題 1 において,  $G = O_1 \text{Diag}(\ell_1^{-1} \psi_1, \dots, \ell_n^{-1} \psi_n) O_1'$  とおく :

**補題 2** それぞれの期待値が存在するとき,

$$\mathbb{E} [\Sigma^{-1} O_1 \Psi O_1'] = \mathbb{E} \left[ O_1 \Psi^{(1r)} O_1' + \text{Tr} (L^{-1} \Psi) (I_p - O_1 O_1') \right].$$

ただし,  $\Psi^{(1r)} = \text{Diag}(\psi_1^{(1r)}, \psi_2^{(1r)}, \dots, \psi_n^{(1r)})$  で

$$\psi_k^{(1r)} = \sum_{b \neq k}^n \frac{\psi_k - \psi_b}{\ell_k - \ell_b} + 2 \frac{\partial \psi_k}{\partial \ell_k} - \frac{\psi_k}{\ell_k}, \quad k = 1, 2, \dots, n.$$

特に, Kubokawa and Srivastava (2008)'s identity として,

$$\mathbb{E} [\text{Tr} \{ \Sigma^{-1} O_1 \Psi O_1' \}] = \mathbb{E} \left[ \sum_{k=1}^n \left\{ (p - n - 1) \frac{\psi_k}{\ell_k} + 2 \frac{\partial \psi_k}{\partial \ell_k} + \sum_{b \neq k}^n \frac{\psi_k - \psi_b}{\ell_k - \ell_b} \right\} \right].$$

## 部分積分の公式と SURE 法 ( 4 )

補題 3

$\hat{\Sigma} = \mathbf{O}_1 \Psi(L) \mathbf{O}'_1$  に対して,

$$\mathbb{E}[\text{Tr} \{ \Sigma^{-1} \mathbf{O}_1 \Psi \mathbf{O}'_1 \Sigma^{-1} \mathbf{O}_1 \Psi \mathbf{O}'_1 \}] = \mathbb{E}[\text{Tr} \{ \Sigma^{-1} \mathbf{O}_1 \tilde{\Psi}^{(1)} \mathbf{O}'_1 \}].$$

ただし,  $\tilde{\Psi}^{(1)} = \text{Diag}(\tilde{\psi}_1^{(1)}, \tilde{\psi}_2^{(1)}, \dots, \tilde{\psi}_n^{(1)})$  で

$$\tilde{\psi}_k^{(1)} = (p - n - 1) \frac{\psi_k^2}{\ell_k} + 4\psi_k \cdot \frac{\partial \psi_k}{\partial \ell_k} + 2\psi_k \cdot \sum_{b \neq k}^n \frac{\psi_k - \psi_b}{\ell_k - \ell_b}, \quad k = 1, 2, \dots, n.$$

## 部分積分の公式と SURE 法 ( 5 )

定理 4

$\hat{\Sigma} = O_1 \Psi(L) O_1'$  に対して,

$$R(\hat{\Sigma}, \Sigma) = \mathbb{E} \left[ \sum_{k=1}^n \left\{ (p - n - 1) \left( \frac{\tilde{\psi}_k^{(1)}}{\ell_k} - 2 \frac{\psi_k}{\ell_k} \right) + 2 \left( \frac{\partial \tilde{\psi}_k^{(1)}}{\partial \ell_k} - 2 \frac{\partial \psi_k}{\partial \ell_k} \right) + \sum_{b \neq k}^n \frac{(\tilde{\psi}_k^{(1)} - 2\psi_k) - (\tilde{\psi}_b^{(1)} - 2\psi_b)}{\ell_k - \ell_b} \right\} + p \right].$$

ただし,  $(k = 1, 2, \dots, n)$  で

$$\tilde{\psi}_k^{(1)} = (p - n - 1) \psi_k^2 / \ell_k + 4\psi_k (\partial \psi_k / \partial \ell_k) + 2\psi_k \sum_{b \neq k}^n (\psi_k - \psi_b) / (\ell_k - \ell_b)$$

## 改良型推定量 ( 1 )

推定量の族

 $n < p$  とする . つぎの推定量の族を考える :

$$\hat{\Sigma}_t = \frac{1}{p+n+1} \left( \mathbf{W} + \frac{t}{\text{Tr } \mathbf{W}^+} \mathbf{O}_1 \mathbf{O}'_1 \right).$$

ただし,  $\mathbf{O}_1$  は  $p \times n$  の半直交行列で,  $\mathbf{W}$  の正の固有値に対応する固有ベクトルを並べたのもの,  $\mathbf{W}^+$  は  $\mathbf{W}$  の Moore-Penrose の逆行列,  $t$  は正の定数である .

結果

$\hat{\Sigma}_t$  のリスクの不偏推定量 (SURE) を導出し, リスクを評価することにより次の結果を得る .

$$0 < t < 2(n-1)(p-n-1) / \{(p-n+1)(p-n+2)\}$$

のとき, すべての  $\Sigma$  に対して,  $R(\hat{\Sigma}_t, \Sigma) \leq R(n^{-1}\mathbf{W}, \Sigma)$  が成立する .

## 改良型推定量 ( 2 )

★  $\hat{\Sigma}_t$  は正定値ではない .

★  $\frac{1}{p+n+1}(\mathbf{W} + \frac{t}{\text{Tr } \mathbf{W}^+} \mathbf{O}_1 \mathbf{O}'_1)$  を修正したもの:

$$\tilde{\Sigma}_{\text{HF}} = \frac{1}{p+n+1} \left\{ \mathbf{W} + \frac{t_0}{\text{Tr } \mathbf{W}} \mathbf{I}_p \right\}, \quad t_0 = \frac{2(n-1)(p+n+1)}{(p-n+1)(p-n+3)}.$$

★ 残念なことに , 推定量  $\tilde{\Sigma}_t$  のリスクを SURE を用いて評価できない !

## 数値実験の結果 ( 1 )

the percentage relative improvement in average loss of  $\hat{\Sigma}_{\text{HF}}$  over  $\mathbf{W}/(n+p+1)$

$$\text{PRIAL}(\hat{\Sigma}_{\text{HF}}) = \frac{\text{average loss of } \mathbf{W}/(n+p+1) - \text{average loss of } \hat{\Sigma}_{\text{HF}}}{\text{average loss of } \mathbf{W}/(n+p+1)}$$

parameter

$$\alpha = \{p - (\text{Tr } \Sigma^{-1})^2 / \text{Tr } \Sigma^{-2}\} / p.$$

注意

$$\Sigma = \mathbf{I}_p \implies \alpha = 0.$$

## 数値実験の結果 ( 2 )

Table 1: Result of 1000 Monte-Carlo simulations for  $p/n = 1/2$  and  $\Sigma = \mathbf{I}_p$ .

| $p$ | $n$ | $\mathbf{W}/(n + p + 1)$ | $\hat{\Sigma}_{\text{HF}}$ | PRIAL |
|-----|-----|--------------------------|----------------------------|-------|
| 10  | 5   | 6.863 (0.013)            | 6.3465 (0.017)             | 7.5%  |
| 20  | 10  | 13.549 (0.014)           | 12.654 (0.017)             | 6.6%  |
| 40  | 20  | 26.891 (0.014)           | 25.647 (0.016)             | 4.6%  |
| 60  | 30  | 40.216 (0.014)           | 38.821 (0.016)             | 3.5%  |
| 80  | 40  | 53.570 (0.014)           | 52.095 (0.015)             | 2.8%  |
| 100 | 50  | 66.894 (0.014)           | 65.362 (0.015)             | 2.3%  |

The values in parentheses refer to the standard error on average loss.

## 数値実験の結果 ( 3 )

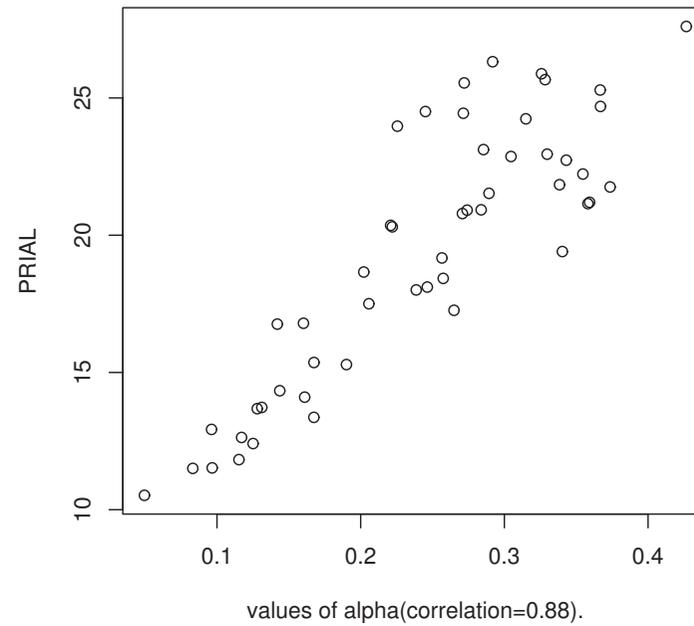


Figure 5: Effect of  $\alpha$  for  $p = 20$  and  $n = 4$ .

## 数値実験の結果 ( 4 )

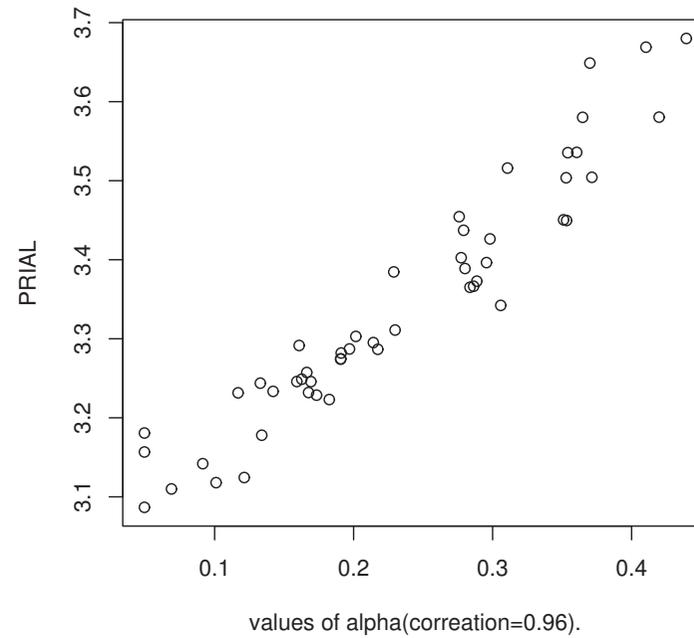


Figure 6: Effect of  $\alpha$  for  $p = 20$  and  $n = 16$ .

## 数値実験の結果 ( 5 )

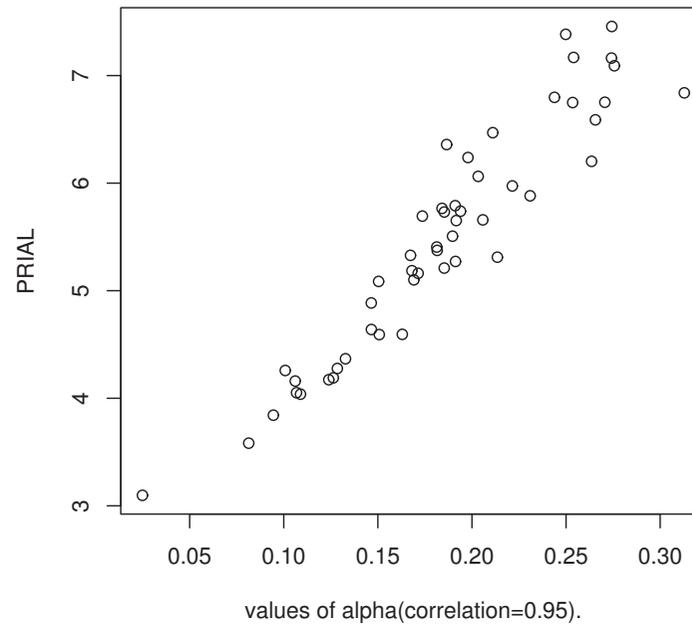


Figure 7: Effect of  $\alpha$  for  $p = 100$  and  $n = 20$ .

## 数値実験の結果 ( 6 )

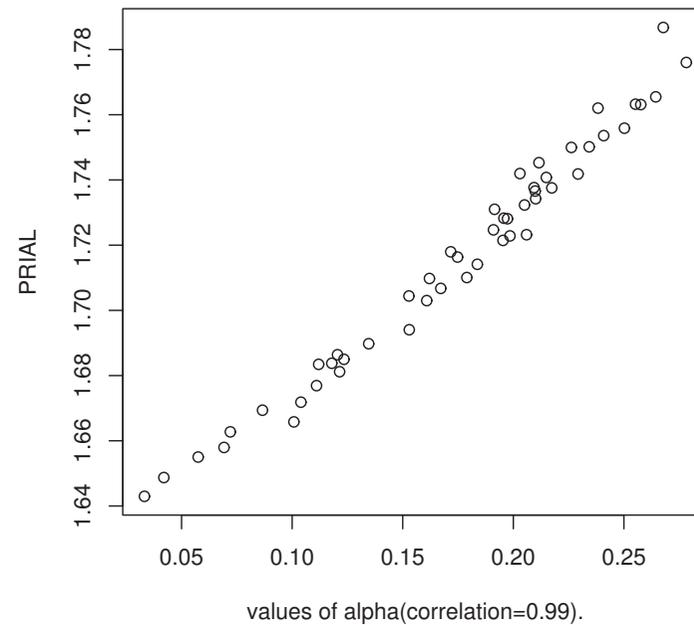


Figure 8: Effect of  $\alpha$  for  $p = 100$  and  $n = 80$ .

## References

- [1] Z.D. Bai, Methodologies in spectral analysis of large dimensional random matrices, a reievw. *Statistica Sinica* **9** (1999) 611-677.
- [2] P.J. Bickel and E. Levina, Covariance regularization by thresholding. *Ann. Statist.* **36** (2008) 2577–2604.
- [3] D.K. Dey and C. Srinivasan, Estimation of a covariance matrix under Stein's loss, *Ann. Statist.* **13** (1985) 1581-1591.
- [4] M.L. Eaton, Group Invariance Application in Statistics, REGIONAL CONFERENCE SERIES IN PROBABILITY AND STATISTICS Vol. 1, Institute of Mathematical Statistics (1989).
- [5] R. Furrer and T. Bengtsson, Estimation of high-dimensional prior and posterior covariance matrices in Kalman filter variants *J. Multivariate Analysis* **98** (2007) 227-255.
- [6] L.R. Haff, Empirical Bayes estimation of the multivariate normal covariance matrix. *Ann. Statist.* **8** (1980) 586–597.

- [7] L.R. Haff, The variational form of certain Bayes estimators, *Ann. Statist.* **19** (1992) 1163-1190.
- [8] W. James and C. Stein, Estimation with quadratic loss, in PROC. FOURTH BERKELEY SYMP. MATH. STATIST. PROB. **1** 361-380, Univ. California Press (1961).
- [9] Y. Konno, Estimation of a normal covariance matrix parametrized by irreducible symmetric cones under Stein's loss, *J. Multivariate Anal.* **98** (2007) 295-316.
- [10] Y. Konno, Shrinkage estimators for large covariance matrices in multivariate real and complex normal distributions under an invariant quadratic loss. *J. Multivariate Anal.* **100** (2009), 2237-2253.
- [11] T. Kubokawa and M. Srivastava, Estimation of the precision matrix of a singular Wishart distribution and its application in high dimensional data. *J. Multivariate Anal.* **99** (2008) 1906-1928.
- [12] O. Ledoit and M. Wolf, A well-conditioned estimator for large-dimensional covariance matrices. (English summary) *J. Multivariate Analysis* **88** (2004) 365-411.

- [13] H. Massam and E. Neher, Estimation and testing for lattice conditional independence models of Euclidean Jordan algebras, *Ann. Statist.* **26** (1998) 1051-1082.
- [14] R.J. Muirhead, ASPECTS OF MULTIVARIATE STATISTICAL ANALYSIS, John Wiley & Sons, Inc (1982).
- [15] J. Schäfer and K. Strimmer, A shrinkage approach to large-scale covariance matrix estimation and implications for functional genomics, *Stat. Appl. Genet. Mol. Biol.* **4** (2005), Art. 32, 28 pp. (electronic).
- [16] M. Srivastava, Singular Wishart and multivariate beta distribution. *Ann. Statist.* **31** (2003) 1537-1560.
- [17] C. Stein, Lectures on the theory of estimation of many parameters, in STUDIES IN THE STATISTICAL THEORY OF ESTIMATION I(I. A. Ibragimov and M. S. Nikulin, eds.) (1977).
- [18] W.B. Wu and M. Pourahmadi, Nonparametric estimation of large covariance matrices of longitudinal data *Biometrika* **90** (2003) 831-844