科研基盤 (B) 「生物情報を解明するための統計数学的基礎理論とその応用」
(研究代表者：赤平昌文 (筑波大学)) によるシンポジウム

2009 年 10 月 1 日 – 3 日（於：日本女子大学）

Wishart 分布のスケール行列の推定について

日本女子大学理学部　　今野 良彦

September 30, 2009

この講演の構成

- 背景
- 統計的決定問題の用語
- Wishart 分布ののスケール行列の推定問題における古典的な結果を概括
- Andersson and Perlman [2] による多変量正規 Lattice conditional independence model (LCI) モデルの基本的な結果
- 3×3 の分散行分散行列の逆行列の $(3, 2)$ 成分が 0 の条件付独立多変量正規モデルにおけるミニマックス推定量

背景（1）

- 多変量正規分布の Covariance matrix estimation = Wishart 分布のスケール行列の推定

- $X_1, X_2, \ldots, X_n \sim N_r(0, \Sigma)$, $i.i.d.$ のとき，

$$S = \sum_{i=1}^{n} X_i X_i'$$

  の分布を Wishart 分布という．ただし，$X_i\,(i = 1, 2, \ldots, n)$ は縦確率ベクトルで，X_i' は転置ベクトル．

- $X_1, \ldots, X_n$ に基づく Σ の推定問題は，S は Σ の十分統計量なので，S に基づいてのみ考えてばよい．

背景（2）

$\ell_1, \ldots, \ell_r\,(\ell_1 \geq \ell_2 \geq \cdots \geq \ell_r \geq 0)$ を $S = \sum_{i=1}^n X_i X_i'$ の固有根とする．このとき，最尤推定量 $(1/n)S$ は Σ の不偏推定量；

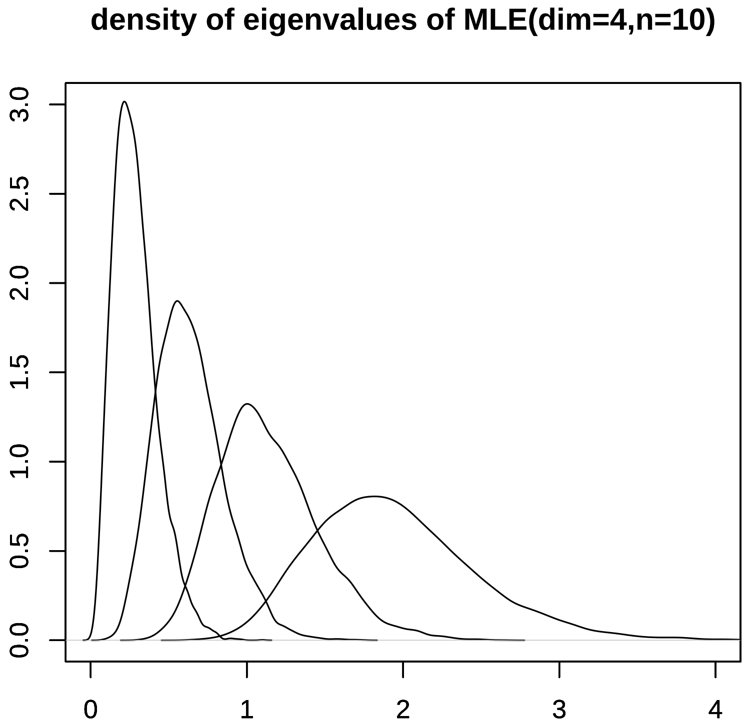
$$\mathbb{E}\left[\frac{1}{n}S\right] = \Sigma. \qquad \Sigma \text{ は } X_i \text{ の分散行分散行列！} X_i \sim N_r(0, \Sigma)$$

しかし，$\sigma_1, \ldots, \sigma_r\,(\sigma_1 \geq \cdots \geq \sigma_r > 0)$ を Σ の固有値としたとき，

$$\mathbb{E}[\ell_1] > \sigma_1,\ \mathbb{E}[\ell_r] < \sigma_r, \qquad \text{provided } n \geq r.$$

背景（３）：最尤推定量の固有根の周辺分布

$$(\Sigma = I_4,\ r = 4,\ n = 10)$$



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

Marčhenko and Pastur law [4] を参照．$r = r(n)$ とし，$r/n \to \tau \le 1 (n \to \infty)$ のとき，

$$\frac{1}{r}\sum_{i=1}^{r} \mathbb{1}_{\lambda_i \le x} \to \int_{-\infty}^{x} \frac{\sqrt{(b-t)(t-a)}}{2\pi t\tau} \mathbb{1}_{a\le t\le b}\, dt$$

(almost surely).

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

Stein [17] の Lecture 4 を参照.

背景（5）： Wishart 行列の固有根の分布

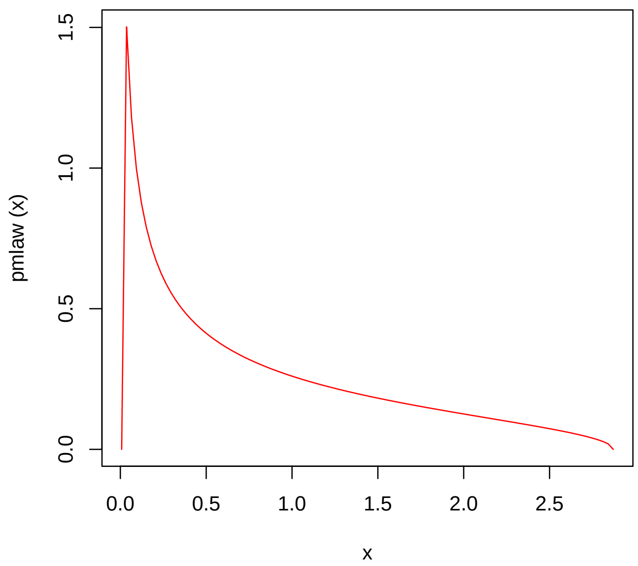


Figure 1: Density of Marčhenko-Pastur's quater-circle law ($\tau = 4/7 \doteqdot 0.57$, $a \doteqdot 0.06$, $b \doteqdot 3.086$).

背景（6）

Covariance matrix estimation

- Classical approach:
  - James and Stein [9], Stein [17], Dey and Srinivasan [5], Haff [6].

- Large covariance matrix estimation:
  - Lediot and Wolf [12]：経験共分散行列と構造化された行列（Sphericity）の convex combination. $\longrightarrow$ Finance におけるポートフォリオ選択
  - sparse genomic data に基づく large-scale covariance matrix の推定：Schäfer and Strimmer [16].

この講演の構成

- 背景
- 統計的決定問題の用語
- Wishart 分布ののスケール行列の推定問題における古典的な結果を概括
- Andersson and Perlman [2] による多変量正規 Lattice conditional independence model (LCI) モデルの基本的な結果
- 3×3 の分散行分散行列の逆行列の $(3, 2)$ 成分が 0 の条件付独立多変量正規モデルにおけるミニマックス推定量

統計的決定問題

(i) 標本空間 $(\mathscr{X}, \mathcal{B}_1)$;

(ii) 母数空間 $(\Theta, \mathcal{B}_2)$;

(iii) 推定量の族 $\mathscr{D}$; ここでは，非確率化推定量のみを対象とする.

(iv) 統計モデル $\{P_\theta : \theta \in \Theta\}$，すなわち，標本空間 $(\mathscr{X}, \mathcal{B}_1)$ 上の確率測度の族;

(v) 損失関数 $\mathbf{L} : \Theta \times \mathscr{D} \to [0, \infty)$，ただし，$\mathbf{L}$ は可測関数である.

危険関数と推定量の最適性（１）

$\boxed{\text{危険関数(リスク)}}$ 与えられた推定量 δ に対して，δ の危険関数 $\mathbf{R}(\theta, \delta)$ を

$$\mathbf{R}(\theta, \delta) = \int_{\mathscr{X}} \mathbf{L}(\theta, a)\, P_\theta(dx).$$

$\boxed{\text{決定関数のよさ}}$ 推定量全体の集合を $\mathscr{D}$ と書くことにする．$\delta_1, \delta_2 \in \mathscr{D}$ に対して，

- $\mathbf{R}(\theta, \delta_1) \leq \mathbf{R}(\theta, \delta_2) \qquad (\forall \theta \in \Theta)$
- $\mathbf{R}(\theta_0, \delta_1) < \mathbf{R}(\theta_0, \delta_2) \qquad (\exists \theta_0 \in \Theta)$

が成り立つとき，δ_1 が δ_2 よりよいという．

危険関数と推定量の最適性（２）

$\boxed{\text{許容性}}$ δ_1 よりよい $\delta \in \mathscr{D}$ が存在しないとき，δ_1 は許容的であるという.

$\boxed{\text{ミニマックス決定関数}}$ δ_1 がミニマックス推定量であるは,

$$\max_{\theta\in\Theta}\mathbf{R}(\theta,\,\delta_1)=\min_{\delta\in\mathscr{D}}\max_{\theta\in\Theta}\mathbf{R}(\theta,\,\delta)$$

が成り立つときにいう.

危険関数と推定量の最適性（３）

$\boxed{\text{不変性}}$ $(\mathscr{X}, G)$, $(\Theta, \bar{G})$, $(\mathscr{D}, \tilde{G})$ は変換群とする．以後，G, $\bar{G}$, $\tilde{G}$ をすべて G とかくことにする．

決定問題が G 不変であるとは，つぎの条件をみたすときをいう．

(i) モデル $\{P_\theta : \theta \in \Theta\}$ は不変：$gP_\theta = P_{g\theta}, \qquad (g \in G, \theta \in \Theta)$.

ただし，任意の $B \in \mathcal{B}_1$ に対して，$gP_\theta(B) = P_\theta(g^{-1}B)$ である．

(ii) 損失関数 $\mathbf{L}$ は不変：$\mathbf{L}(g\theta, g\delta) = \mathbf{L}(\theta, \delta), \qquad (g \in \Theta, \theta \in \Theta, \delta \in \mathscr{D})$.

危険関数と推定量の最適性（４）

決定問題が G 不変，推定量 δ が G 不変：$g\delta(x) = \delta(gx)\ (\forall g \in G)$，$G$ が Θ に推移的に作用していれば，固定した元 $\theta_0 \in G$ に対して，

$$\mathbf{R}(\theta,\, \delta) = \mathbf{R}(\theta_0,\, \delta)$$

となる．

$\boxed{\text{最良 } G \text{ 不変性とミニマックス性(Hunt-Stein の定理の簡易版)}}$

G は三角群とし，決定問題は G 不変とする．このとき，G 不変決定関数 δ_1 に対して，

$$\max_{\theta\in\Theta} \mathbf{R}(\theta,\, \delta_1) = \min_{\delta\in\mathscr{D}^*} \max_{\theta\in\Theta} \mathbf{R}(\theta,\, \delta) \Longrightarrow \delta_1 \text{ はミニマックス}$$

ただし，$\mathscr{D}^*$ は G 不変決定関数の全体である．

Wihart 分布のスケール行列の推定問題：設定

- $X_1, X_2, \ldots, X_n \, (n \geq r) \sim N_r(0, \Sigma)$, i.i.d.　ただし，$\Sigma \in \mathrm{Sym}^+(r, \mathbb{R})^1$.
- $S := \sum_{i=1}^n X_i X_i'$. $\longleftarrow$ 自由度 n，スケール行列 Σ の Wishart 分布 $W_r(\Sigma, n)$.
- Σ の推定問題の損失関数：

$$\mathbf{L}(\Sigma, a) = \mathrm{Tr}\,(a\Sigma^{-1}) - \log \mathrm{Det}\,(a\Sigma^{-1}) - r.$$

ただし，a は Σ の推定量で S に依存.

- リスク関数：$\mathbf{R}(\Sigma, a) = \mathbb{E}[\mathbf{L}(\Sigma, a)]$

[1] $\mathrm{Sym}^+(r, \mathbb{R})$ は r 次正値実対称行列の全体を表わす.

損失関数の性質

(i) $\mathbf{L}(\Sigma, a) \geq 0$ で等号成立は $a = \Sigma$ のときのみ;

(ii) $\mathbf{L}(\Sigma, a)$ は第 2 の変数に関して凸関数;

(iii) 任意の $g \in \mathrm{GL}(r, \mathbb{R})$ [2] に対して,

$$\mathbf{L}(g\Sigma g', gag') = \mathbf{L}(\Sigma, a).$$

ただし，g' は行列 g の転置行列を示す.

[2] $\mathrm{GL}(r, \mathbb{R})$ は r 次正則行列の全体である.

James and Stein (1961) の結果

- $\mathrm{GL}(r,\mathbb{R}) \supset \mathrm{LT}^+(r,\mathbb{R})$[3] も $\mathrm{Sym}^+(r,\mathbb{R})$ に推移的に作用.
- $S = TT'\,(T \in \mathrm{LT}^+(r,\mathbb{R}))$ したとき，$\mathrm{LT}^+(r,\mathbb{R})$ 不変推定量は

$$TAT', \qquad A \in \mathrm{Sym}(r,\mathbb{R}).$$

- $\mathbf{R}(I_r, TAT') \geq \sum_{i=1}^{r} \left\{ \log(n+r-2i+1) - \mathbb{E}[\log \chi^2_{n-i+1}] \right\}$.
- 推定量 $\widehat{\Sigma}_{\min} = T\mathrm{diag}\,(1/(n+r-2i+1)|\, i=1,2,\ldots,r)T'$ は上の下限に到達し，ミニマックス

[3] $\mathrm{LT}^+(r,\mathbb{R})$ は r 次下三角行列で対角成分がすべて正の行列全体を表わす.

直交不変推定量：動機

$\boxed{\text{Marčhenko and Pastur law [4]}}$ を参照. $r = r(n)$ とし，$r/n \to \tau \le 1 (n \to \infty)$ のとき,

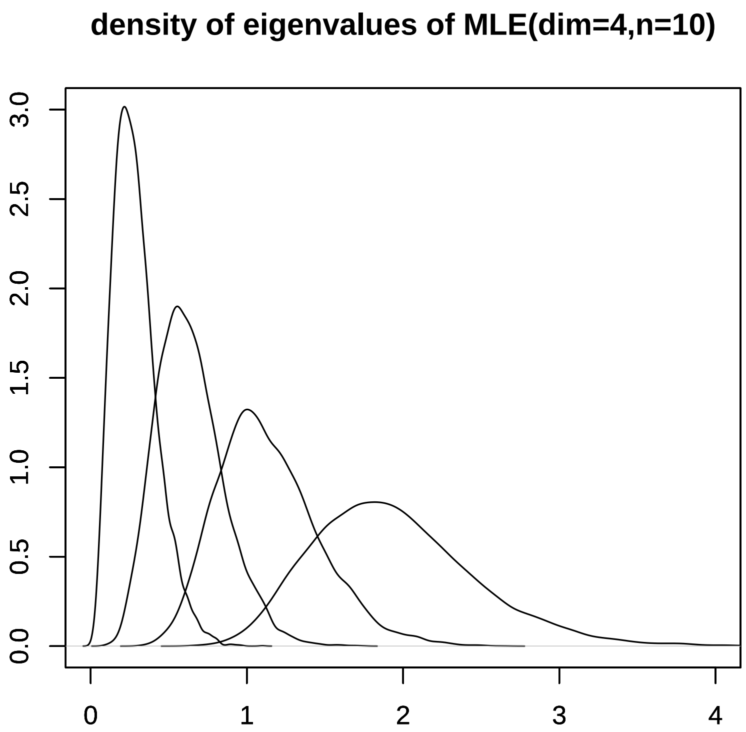
$$\frac{1}{r}\sum_{i=1}^{r} \mathbb{1}_{\lambda_i \le x} \to \int_{-\infty}^{x} \frac{\sqrt{(b-t)(t-a)}}{2\pi t\tau} \mathbb{1}_{a\le t\le b}\, dt$$

(almost surely).

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

Stein [17] の Lecture 4 を参照.

最尤推定量の固有根の周辺分布 $(r = 4,\ n = 10)$



Dey and Srinivasan (1985) の結果

- Σ の直交不変推定量：$\widehat{\Sigma} = H\mathrm{diag}\,(\phi_i |\, i = 1,\, 2,\, \ldots,\, r)H'\,(H \in \mathrm{O}(r,\, \mathbb{R}))$.
  ただし，$\phi_i : (\mathbb{R}_+)^n \mapsto \mathbb{R}_+$，$S = H\mathrm{diag}\,(\ell_1,\, \ldots,\, \ell_r)H'$.

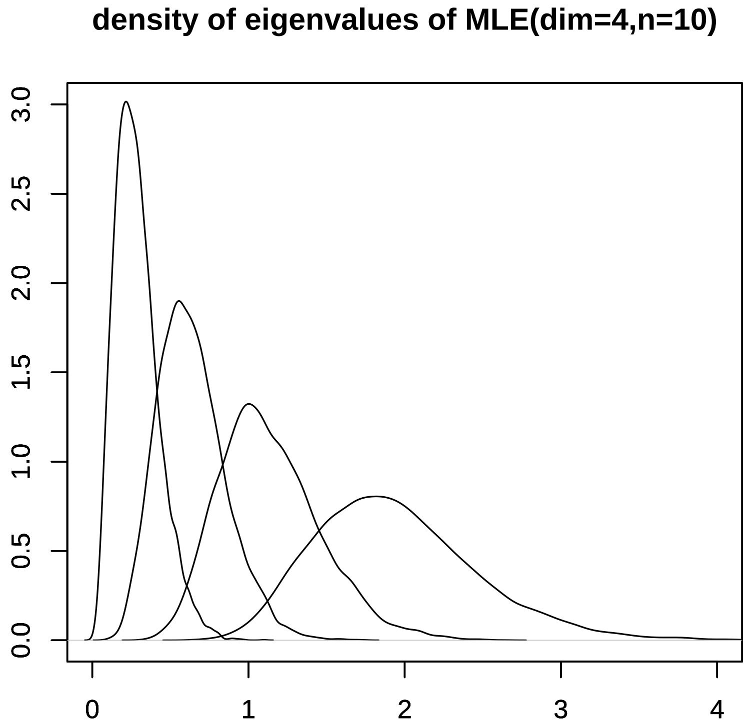
- $\phi_i = (1/n)\ell_i$ とすれば，$\widehat{\Sigma}_{\mathrm{mle}}$ を得る.

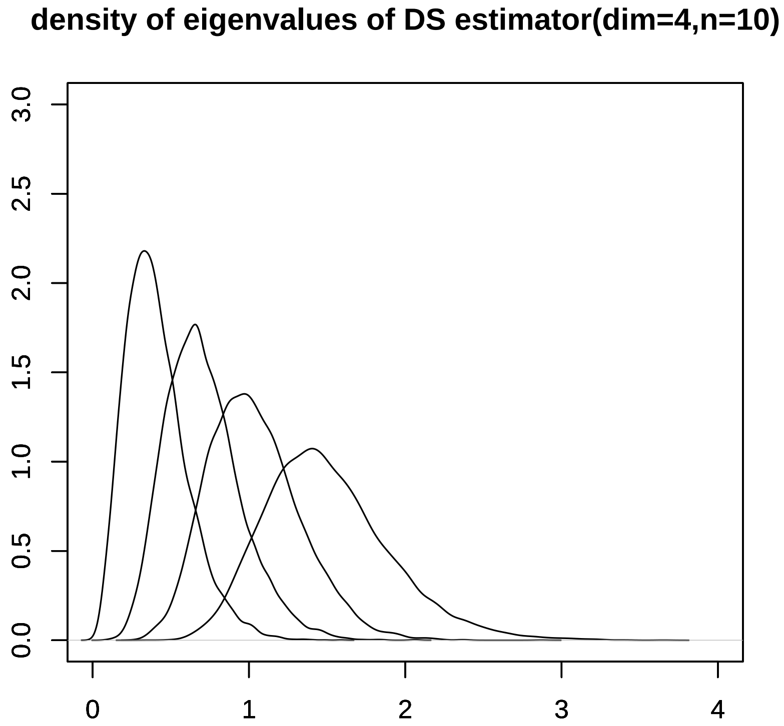
- 直交群 $\mathrm{O}(r,\, \mathbb{R})$ の $\mathrm{Sym}^+(r,\, \mathbb{R})$ への作用は推移的でないので，リスクは一般には定数ではない！

- リスクの直接的な評価 $\longrightarrow$ Stein's Unbiased Risk Estimate 法（SURE 法）

- $\mathbf{R}(\Sigma,\, H\mathrm{diag}\,(d_1\ell_1,\, \ldots,\, d_r\ell)H') \leq -\sum_{i=1}^{r}\left\{\log d_i + \mathbb{E}[\log \chi^2_{n-i+1}]\right\}\ (\forall\Sigma)$.
  ただし，$d = (n + r - 2i + 1)^{-1}\,(i = 1,\, \ldots,\, r)$

最尤推定量と DS 推定量の固有根の周辺分布 $(r = 4,\ n = 10)$





この講演の構成

- 背景
- 統計的決定問題の用語
- Wishart 分布ののスケール行列の推定問題における古典的な結果を概括
- Andersson and Perlman [2] による多変量正規 Lattice conditional independence model (LCI) モデルの基本的な結果
- 3×3 の分散行分散行列の逆行列の $(3, 2)$ 成分が 0 の条件付独立多変量正規モデルにおけるミニマックス推定量

Lattice conditional independence (LCI) モデル（１）

- $\mathcal{K}$：集合 $I = \{1, 2, \ldots, r\}$ の部分集合からなる分配環とし，集合の和（$\cup$）と積（$\cap$）について閉じて空集合を含む.

- $X \sim N_r(0, \Sigma)$ とし、$K \in \mathcal{K}$ に対して，X_K を K の元に対応する座標の成分をもつ部分ベクトル.

- LCI モデル $\boldsymbol{N}(\mathcal{K})$ を

$$X_{L\cap M} \text{ を与えたとき，} X_L \text{ と } X_M \text{ は条件付き独立} \quad (\forall L, M \in \mathcal{K}) \tag{1}$$

であるような多変量正規分布 $N_r(0, \Sigma)$ の集合とする. また，$\boldsymbol{P}(\mathcal{K})$ をすべての L, $M \in \mathcal{K}$ に対して，上の条件を満足するような非退化多変量正規分布の分散共分散行列[4]の集まりとする.

---

[4]したがって，正定値行列.

LCI モデル（２）

$\boxed{\text{記号}}$

- $K \in \mathcal{K}$ に対して，

$$\langle K \rangle = \cup(\tilde{K} \in \mathcal{K} | \tilde{K} \subset K \quad \text{かつ} \quad \tilde{K} \neq K), \qquad [K] = K \backslash \langle K \rangle$$

- $K = \langle K \rangle \cup [K]$ かつ $\langle K \rangle \cap [K]$ は空集合.
- $\langle K \rangle \in \mathcal{K}$ であるが，$[K] \in \mathcal{K}$ は必ずしも成立しない.
- つぎの部分集合族を考える：$\mathcal{J}(\mathcal{K}) = \{K \in \mathcal{K} | [K]$ は空でない$\}$

LCI モデル（３）

- $J(\mathcal{K})$ に含まれる部分集合を $K_1, K_2, \ldots, K_q\,(q \leq r)$ とする．順番は，$i < j$ ならば，$K_j \not\subset K_i$.

- 各 $K_i \in J(\mathcal{K})\,(i = 1, 2, \ldots, q))$ は $[K_i] \neq \emptyset$ で，$i \neq j$ ならば，$[K_i] \cap [K_j] = \emptyset$ であり，任意の $K \in \mathcal{K}$ に対して，

$$K = \cup_{K_i \subset K\ (i=1,\ldots,q)}[K_i]; \qquad I = [K_1] \cup [K_2] \cup \cdots \cup [K_q].$$

以後は，$[K_1], [K_2], \ldots, [K_q]$ の順に行列の成分は並んでいるとする．

LCI モデル（4）

- $L,\, M \subset I$ に対して，行列 A に対して，$(i,\, j)$ 成分（$i \in L,\, j \in M$）を集めた部分行列を $A_{L,\, M}$
- つぎの集合を考える：

$$\begin{aligned}\boldsymbol{LT}(\mathcal{K}) \quad &= \quad \{A \text{ は } I \text{ 次の正則行列};\\ &\qquad\qquad \forall L,\, M \in J(\mathcal{K}) :\, L \neq M, \text{ and } M \not\subset L \Longrightarrow A_{[L],\, [M]} = 0\}.\end{aligned}$$

- $\mathcal{K} = \{\emptyset,\, I\}$ でなければ，$K_1,\, K_2,\, \ldots,\, K_q$ は非減少（包含関係の意味）列であるので，$\boldsymbol{LT}(\mathcal{K})$ はブロック下三角行列になっている．

- $\Sigma \in \boldsymbol{P}(\mathcal{K})$ に対して，

$$\boldsymbol{LT}(\mathcal{K}) \times \boldsymbol{P}(\mathcal{K}) \ni (A,\, S) \mapsto ASA' \in \boldsymbol{P}(\mathcal{K})$$

となり，推移的な作用となる．ただし，A' は行列 A の転置行列とした．

LCI モデル（5）

$\boxed{r = 3}$

- LCI モデル $\boldsymbol{N}(\mathcal{K}_1)$ からのランダム標本 $X_1,\, X_2,\, \ldots,\, X_n (n \geq 3)$ に基づいて分散共分散行列 Σ の推定問題を損失関数

$$\mathbf{L}(\Sigma,\, a) = \mathrm{Tr}\,(a\Sigma^{-1}) - \log \mathrm{Det}\,(a\Sigma^{-1}) - r.$$

のもとで考える.

- $\mathcal{K}_1 = \{\emptyset,\, 1,\, 12,\, 13,\, I\}$ [5]とする. (1) から $\mathcal{K}_1$ は

$$2 \perp\!\!\!\perp 3 \,|\, 1 \Longleftrightarrow (\Sigma^{-1})_{2,\,3} = (\Sigma^{-1})_{3,\,2} = 0 \Longleftrightarrow (\Sigma)_{3,\,2} = \frac{(\Sigma)_{3,\,1}(\Sigma)_{2,\,1}}{(\Sigma)_{1,\,1}}$$

となる制限になることに注意する.

[5] 12 は集合 $\{1,\, 2,\, 3\}$ の部分集合 $\{1,\, 2\}$ の意味である.

LCI モデル（6）

- $\mathcal{J}(\mathcal{K}_1) = \{1, 12, 13\}$ であり，$K_1 = 1, K_2 = 12, K_3 = 13$ とおけば，K_1, K_2, K_3 は非減少列であり，$[K_1] = 1, [K_2] = 2, [K_3] = 3$ となる．
- 変換
$$X_1 \mapsto \begin{pmatrix} a_{11} & 0 & 0 \\ a_{21} & a_{22} & 0 \\ a_{31} & 0 & a_{33} \end{pmatrix} X_1 =: AX_1, \qquad A \in \boldsymbol{LT}(\mathcal{K}_1) \tag{2}$$
のもとで，$\mathcal{K}_1$ により制限されたモデルは不変であることがわかる．
- $\Sigma_A = \mathbb{COV}[AX_1]$ としたとき，$\frac{(\Sigma_A)_{3,1}(\Sigma_A)_{2,1}}{(\Sigma_A)_{1,1}} = (\Sigma_A)_{3,2}$ をみたす．
- $a_{11}, a_{22}, a_{33} > 0$ とし，(2) により定まる変換群を $\boldsymbol{LT}^{+}(\mathcal{K}_1)$ と書くことにする．

LCI モデル（７）

- LCI モデル $\boldsymbol{N}(\mathcal{K}_1)$ からのランダム標本 $X_1, X_2, \ldots, X_n$ に対して,

$$S = \sum_{i=1}^{n} X_i X_i' = (s_{ij})$$

とおく.

- $\mathcal{K}_1$ で制約されたモデルの分散共分散行列 Σ の最尤推定量：

$$n\hat{\Sigma}^{\text{mle}}_{\mathcal{K}_1} = \begin{bmatrix} s_{11} & s_{12} & s_{13} \\ s_{21} & s_{22} & s_{21}s_{11}^{-1}s_{13} \\ s_{31} & s_{31}s_{11}^{-1}s_{12} & s_{33} \end{bmatrix}$$

LCI モデル（8）

- つぎの分解を考える：$n\hat{\Sigma}^{\mathrm{mle}}_{\mathcal{K}_1} = T_1 T_1'$ と書く．ただし，

$$T_1 = \begin{bmatrix} s_{11}^{1/2} & 0 & 0 \\ s_{21}s_{11}^{-1/2} & t_{22} & 0 \\ s_{31}s_{11}^{-1/2} & 0 & t_{33} \end{bmatrix} \in \boldsymbol{LT}^+(\mathcal{K}_1); \quad t_{ii}^2 = s_{ii} - s_{i1}s_{11}^{-1}s_{1i} > 0\,(i = 2,\, 3).$$

- ミニマックス推定量：$\hat{\Sigma}^{\mathrm{m}}_{\mathcal{K}_1} = T_1 D_1 T_1'$; $D_1 =$ diag $\left(\frac{1}{n+2}, \frac{1}{n-1}, \frac{1}{n-1}\right)$.

$$\hat{\Sigma}^{\mathrm{m}}_{\mathcal{K}_1} = \begin{bmatrix} \frac{s_{11}}{n+2} & \frac{s_{12}}{n+2} & \frac{s_{13}}{n+2} \\ \frac{s_{21}}{n+2} & \frac{1}{n-1}\left(s_{22} - \frac{3}{n+2}s_{21}s_{11}^{-1}s_{12}\right) & \frac{s_{21}s_{11}^{-1}s_{13}}{n+2} \\ \frac{s_{31}}{n+2} & \frac{s_{31}s_{11}^{-1}s_{12}}{n+2} & \frac{1}{n-1}\left(s_{33} - \frac{3}{n+2}s_{31}s_{11}^{-1}s_{13}\right) \end{bmatrix}$$

むすび

- 前節の議論において得られた推定量 (3) は，許容的かどうか不明である．すくなくとも 3 節の手法を用いて，推定量 (3) を改良する推定量を求めることはできないように思われる.
- Andersson and Perlman [2] で得られた結果を用いると 前節の結果を一般の Lattice conditional independence model に拡張することが可能である. Konno [10] を参照.

- 古典的な Wishart 分布とその性質を等質錐（すなわち，homogeneous cones, 自己同型群の部分群が推移的に作用する開凸錐）上の分布に一般化を行う研究がいくつがある．Andersson and Wojnar [3], Ishi [7, 8]．さらに，等質性をみたさないものも含んだモデルとしては，Andersson and Klein [1], Letac and Massam [13], Roberat [15] がある．これらの研究で一般化された Wishart 分布が統計推測理論でそのような役割を果たす可能性があるかは興味のあるところである.

References

[1] S.A. Andersson and T. Klein, On Rietz and Wishart distributions associated with decomposable undirected graphs., Preprint Nr. 06/2008. Feb. 2008.

[2] S.A. Andersson, M.D. Perlman, Lattice models for conditional independence in a multivariate normal distribution, *Ann. Statist.* **21** (1993) 1318-1358.

[3] S.A. Andersson, G.G. Wojnar, The Wishart distributions on homogeneous cones, *J. Theoret. Prob.* **16** (2004) 781-818.

[4] Z.D. Bai, Methodologies in spectral analysis of large dimensional random matrices, a reivew. STATISTICA SINICA **9** (1999) 611-677.

[5] D.K. Dey, C. Srinivasan, Estimation of a covariance matrix under Stein's loss, *Ann. Statist.* **13** (1985) 1581-1591.

[6] L.R. Haff, The variational form of certain Bayes estimators, *Ann. Statist.* **19** (1992) 1163-1190.

[7] H. Ishi, Positive Reisz distributions on homogeneous cones, *J. Math. Soc. Japan*, **52** (2000) 161-186.

[8] H. Ishi, 開凸錐に付随する Wishart 分布（P. Graczyk 氏との共同研究），シンポジウム「生物情報を解明するための統計理論とその応用」, 2007, 12-22.

[9] W. James, C. Stein, Estimation with quadratic loss, in Proc. Fourth Berkeley Symp. Math. Statist. Prob. **1** 361-380, Univ. California Press (1961).

[10] Y. Konno, Inadmissibility of the maximum likelihood estimator of normal covariance matrices with the lattice conditional independence, J. Multivariate Anal. **79** (2001) 33-51.

[11] S.L. Lauritzen, Graphical models, Oxford Univ. Press (1996).

[12] O. Ledoit and M. Wolf, A well-conditioned estimator for large-dimensional covariance matrices, J. Multivariate Anal. **88** (2004) 365–411.

[13] G. Letac, H. Massam, Wishart distributions for decomposable graphs, *Ann. Statist.* **35** (2007) 1278–1323.

[14] S.P. Lin, M.D. Perlman, A Monte Carlo comparison of four estimators for a covariance matrix, in Multivariate Analysis VI (P.R. Krishnaiah, ed.) 411-426, North-Holland, Amsterdam (1985).

[15] A. Roverat, Cholesky decomposition of hyper inverse Wishart matrix, Biometrika **87** (2000) 99–112.

[16] J. Schäfer and K. Strimmer, S shrinkage approach to larege-scale covariance matrix estimation and implications for functional genomics, Stat. Appl. Genet. Mol. Biol. **4** (2005), Art. 32, 28 pp. (electronic).

[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).