Lattice conditional independece モデルにおける分散共分散行列の推定問題について[1]

今野良彦　千葉大学大学院自然科学研究科
email address: konno@math.s.chiba-u.ac.jp
home page: http://www.math.s.chiba-u.ac.jp/~konno

0. はじめに

多変量正規分布のモデリングの多くは平均行列と分散共分散行列を構造化する（制約を加える）ことと考えることができる．たとえば，MANOVA モデル，GMANOVA モデルおよびその拡張モデルは，$X : n \times p \sim N(\Xi, I_p \otimes \Sigma)$ とした[2]ときに，Ξ の構造の制約として表現が可能である．また，変量間の独立性，条件付独立性，対称性は分散共分散行列およびその逆行列の構造化としてとらえることができる．変量間の独立性，条件付独立性を表現するモデルの表現方法とその推測手法 (特に，最尤推定量や尤度比検定量）の導出に関する研究が 90 年代には盛んに行われた．特に，有向グラフや無向グラフなどを使って変量間の独立性，条件付独立性を表現する Graphical model に関わる一連の研究（これらの理論的研究の成果は Lauritzen (1996) を参照されたい）があげられる．

本報告では，Andersson and Perlman (1993) により提案された，有向非巡回グラフモデルの部分集合である Lattice conditional independence モデルにおける分散共分散行列の推定問題を統計的決定理論の立場から議論する．1 節においては，多変量正規分布の分散共分散行列の推定問題を定式化し，既知の結果の簡単なおさらし，2 節では Lattice conditional independence model の定義とその性質を説明し，3 節ではこのモデルのもとでの分散共分散行列の推定問題を議論し，最尤推定量を改良する推定量の導出を行う．最後に，3 変量正規分布の Lattice conditional independence model における推定量を具体的に構成する．

1. 分散共分散行列の推定問題の既知の結果のおさらい

3 節の議論の流れを理解するために，分散共分散行列の推定問題の既知の結果のおさらいする．詳しい議論については Dey and Srinivasan (1985),Eaton (1989), Muirhead (1982),竹村 (1991) を参照されたい．

S を $p \times p$ の正定値確率行列とし，

(1) $$S \sim W_p(n, \Sigma)$$

[1]基盤研究 (B) (1)「統計的領域推定における精確な推定方式の開発と実用化の試み（研究代表者：赤平昌文）」におけるシンポジウム「統計的領域推定とそれに関連する手法の開発とその応用」(2001 年 10 月 29 日（月）〜31 日（水），筑波大学　大学会館）における講演の予稿である．

[2]Ξ は $n \times p$ の行列，Σ は $p \times p$ の正定値行列で，X の確率密度関数は

$$p(x : \Xi, \Sigma) = \frac{1}{(\sqrt{2\pi})^{mp}(\det\Sigma)^{(m/2)}} \mathrm{etr}\left[-\frac{1}{2}(x-\Xi)\Sigma^{-1}(x-\Xi)'\right]$$

で与えられる．

とする．すなわち，S は自由度 n，平均 $n\Sigma$ の Wishart 分布 $W_p(n,\Sigma)$ に従う．Σ の推定問題を損失関数

$$\boldsymbol{L}(\hat{\Sigma},\Sigma)=\mathrm{tr}(\hat{\Sigma}\Sigma^{-1})-\log\det(\hat{\Sigma}\Sigma^{-1})-p \tag{2}$$

のもとで考えることにする．ここで，$\hat{\Sigma}$ は Σ の推定量とする．この損失関数を (1) に関して期待値を取ったものを危険関数とよび，$\boldsymbol{R}(\hat{\Sigma},\Sigma)$ と記すことにする．

注意 1 ：ここでは，推定問題に対してある種の不変性をもつ損失関数のみを考える．

James and Stein (1961) は正の対角成分をもつ下三角行列のなす群のもとで共変な推定量

$$\hat{\Sigma}(ASA')=A\hat{\Sigma}(S)A', \qquad A \text{ は正の対角成分をもつ任意の下三角行列}$$

を考え，この推定量の族の中で最良のものを求めている．実際，この族は

$$\hat{\Sigma}(S)=TDT', \qquad S=TT' \tag{3}$$

と書ける．ただし，T は正の対角成分をもつ下三角行列，D は正の対角成分をもつ定数対角行列である．変換 $(\Sigma,\, A)\longrightarrow A\Sigma A'$ は母数空間に推移的に作用するので，推定量 (3) の危険関数は定数となり，さらに D に関して危険関数を最小化すれば，下三角群に関して共変な推定量の族の中で最良な推定量は (3) において

$$D=\mathrm{diag}\{1/(n+p-2i+1)\}, i=1,\,2,\ldots,n \tag{4}$$

とすればよく，そのリスクは

$$\sum_{i=1}^{p}\{\log(n+p+1-2i)-\boldsymbol{E}[\log\chi^2_{n-i+1}]\} \tag{5}$$

で与えられることを彼らは示している．さらに，Hunt–Stein の定理（概説としては，Bonder and Milnes (1981) がある）を利用することで，これはミニマックスリスクであることが Kiefer (1957) の議論を用いて示されている．

更に，Stein は，推定量の族として，直交群に関して共変なものを考え，それが

$$\hat{\Sigma}=O\mathrm{diag}(\psi_1,\psi_2,\ldots,\psi_p)O' \tag{6}$$

の形で与えられること示している．ここで，

$$S=OLO', \quad L=\mathrm{diag}(l_1,l_2,\ldots,l_p), \quad l_1\ge l_2\ge\cdots\ge l_p$$

を S のスペクトル分解とし，$\psi_i(i=1,2,\ldots,p)$ は $\{(l_1,l_2,\ldots,l_p)\in\Re^p: l_1\ge l_2\ge\cdots\ge l_p>0\}$ 上で偏微分可能な実数値関数関数とする．

推定量 (6) に対する損失関数は

$$\boldsymbol{L}(\hat{\Sigma},\Sigma)=\mathrm{tr}(\hat{\Sigma}\Sigma^{-1})-\sum_{i=1}^{p}\log\psi_i+\log\det(\Sigma)-p$$

なので，

$$\rho(\hat{\Sigma},\Sigma)=\boldsymbol{E}[\mathrm{tr}(\hat{\Sigma}\Sigma^{-1})-\sum_{i=1}^{p}\log\psi_i]$$

として，この ρ を用いて推定量の比較を行えばよいことがわかる．更に，補遺中の補題 2, 3 を使えば，Stein による ρ の不偏推定量

$$(7) \qquad \hat{\rho}(\hat{\Sigma}) = \sum_{i=1}^{p}\left[2\psi_{ii} + (n-p-1)\frac{\psi_i}{l_i} + 2\sum_{i>j}\frac{\psi_i - \psi_j}{l_i - l_j} - \log\psi_i\right]$$

を得る．ただし，$\psi_{ii} = \partial\psi_i/\partial l_i, i = 1, 2, \ldots, p$ である．

注意 2 ：危険関数の不偏推定量の導出についての証明は竹村 (1991) を参照のこと．

ここで，(6) において，

$$(8) \qquad \psi_i = d_i l_i, \quad d_i = \frac{1}{n+p+1-2i} \quad (i = 1, 2, \ldots, p)$$

とおく．推定量 (8) のリスク（正確には ρ ）の不偏推定量 $\hat{\rho}(\hat{\Sigma})$ とミニマックスリスク (5) を比較することにより，推定量 (8) もミニマックスになることを Stein は示している．

2. Lattice conditional independence model の定義とその性質

集合 $I = \{1, 2, \ldots, p\}$ に対して，$\mathcal{K}$ を I の部分集合からなる分配環とし，(集合の）和と積について閉じて空集合を含むものとする．$X \sim N_p(0, \Sigma)$ 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とする．このとき，Lattice conditional independence (LCI) モデル $\boldsymbol{N}(\mathcal{K})$ を，

$$(9) \qquad \begin{array}{c} \text{すべての } K_1, K_2 \in \mathcal{K} \text{ に対して，} X_{K_1 \cap K_2} \text{ を与えたときに} \\ X_{K_1} \text{ と } X_{K_2} \text{ は条件付き独立} \end{array}$$

であるような多変量正規分布 $N(0, \Sigma)$ の集合とする．また，$\boldsymbol{P}_I(\mathcal{K})$ をすべての $L, K \in \mathcal{K}$ に対して，(9) を満足するような多変量正規分布の分散共分散行列の集まりとする．

LCI モデル $\boldsymbol{N}(\mathcal{K})$ からのランダム標本 $X^{(1)}, X^{(2)}, \ldots, X^{(n)} (n \geq p)$ に基づいて分散共分散行列 Σ の推定問題を損失関数 (2) のもとで考える．

<u>記号</u>：$K \in \mathcal{K}$ に対して，

$$\begin{aligned} \langle K \rangle &= \cup(\tilde{K} \in \mathcal{K} | \tilde{K} \subset K \quad \text{かつ} \quad \tilde{K} \neq K) \\ [K] &= K \backslash \langle K \rangle \end{aligned}$$

とする．したがって，$K = \langle K \rangle \cup [K]$ かつ $\langle K \rangle \cap [K] = \emptyset$ である．更に，

$$\mathcal{J}(\mathcal{K}) = \{K \in \mathcal{K} | [K] \neq \emptyset\}$$

とおく． また，$\boldsymbol{M}_I$ と $\boldsymbol{P}_I$ を $p \times p$ の正方行列と正定値行列の集合とする．また， $A = (a_{i_1 i_2} | (i_1, i_2) \in I \times I) \in \boldsymbol{M}_I$ と $K_1, K_2 \in I$ に対して，

$$A_{K_1 \times K_2} = (a_{i_1 i_2} | (i_1, i_2) \in K_1 \times K_2)$$

とし，その集合を $\boldsymbol{M}_{K_1 \times K_2}$ と書く．また，$A_{K \times K}$ と $\boldsymbol{M}_{K \times K}$ を A_K と $\boldsymbol{M}_K$ とそれぞれ書くことにし，$P \in \boldsymbol{P}_I$ に対して P_K と $\boldsymbol{P}_K$ を同様に 定義する．

$A \in \boldsymbol{M}_I$，$K \in \mathcal{K}$ とし，$K = [K] \cup \langle K \rangle$ に対応して A_K の分割を

$$\begin{pmatrix} A_{\langle K \rangle} & A_{\langle K]} \\ A_{[K\rangle} & A_{[K]} \end{pmatrix}$$

とし，$A_{\langle K \rangle} \in \boldsymbol{M}_{\langle K \rangle}$，$A_{[K]} \in \boldsymbol{M}_{[K]}$，$A_{\langle K]} \in \boldsymbol{M}_{\langle K \rangle \times [K]}$，$A_{[K\rangle} \in \boldsymbol{M}_{[K] \times \langle K \rangle}$ とした．また，$\Sigma \in \boldsymbol{P}_I$ と $K \in \mathcal{J}(\mathcal{K})$ に対して，$\Sigma_{[K]\bullet} = \Sigma_{[K]} - \Sigma_{[K\rangle}\Sigma^{-1}_{\langle K \rangle}\Sigma_{\langle K]}$ とする．

行列 $A \in \boldsymbol{M}_I$ が $\mathcal{K}$ - preserving であるとは，すべての $K \in \mathcal{K}$ と $x \in \Re^p$ に対して，$(Ax)_K = A_K x_K$ が成立することである．これは

$$K_1, K_2 \in \mathcal{J}(\mathcal{K}) : K_1 \not\subseteq K_2 \quad \text{ならば} \quad A_{[K_2] \times [K_1]} = 0$$

が成立することと同値である．$\boldsymbol{GL}_I(\mathcal{K})$ と $\boldsymbol{GT}_I(\mathcal{K})$ をそれぞれ $\mathcal{K}$ - preserving で正則な $p \times p$ 行列と正の対角成分をもつ $\mathcal{K}$ - preserving な下三角行列の群とする．

基本的な性質：つぎの性質は推測問題の議論の上で基本的な役割を果たす．

(i) $\boldsymbol{P}_I(\mathcal{K}) \mapsto \mathbf{X}(\boldsymbol{M}_{[K] \times \langle K \rangle} \times \boldsymbol{P}_{[K]} | K \in \mathcal{J}(\mathcal{K}))$ は bijective map である．

(ii) $\Sigma \in \boldsymbol{P}_I(\mathcal{K})$ であるための必要十分条件は，$x \in \Re^p$ に対して，

$$\mathrm{tr}(\Sigma^{-1}xx') = \sum\left(\mathrm{tr}\{\Sigma^{-1}_{[K]\bullet}(x_{[k]} - \Sigma_{[K\rangle}\Sigma^{-1}_{\langle K \rangle}x_{\langle K \rangle})(\cdots)'\}\middle| K \in \mathcal{J}(\mathcal{K})\right)$$

を満足することである．

(iii) $\Sigma \in \boldsymbol{P}_I(\mathcal{K})$ と $L \in \mathcal{K}$ に対して，

$$\det(\Sigma_L) = \Pi(\det \Sigma_{[K]\bullet} | K \in \mathcal{J}(\mathcal{K}), K \subseteq L)$$

が成立する．

(iv) $\boldsymbol{GL}_I(\mathcal{K})$ による $\boldsymbol{P}_I(\mathcal{K})$ への作用 $(A, \Sigma) \mapsto A\Sigma A'$ は推移的である． ただし，$A \in \boldsymbol{GL}_I(\mathcal{K}), \Sigma \in \boldsymbol{P}_I(\mathcal{K})$ である．

$S = \sum_{i=1}^n X_i X_i'$ とおく．$\Sigma \in \boldsymbol{P}_I(\mathcal{K})$ から LCI モデルの性質 (i) を使い，Σ の最尤推定量は

$$n\hat{\Sigma}_{[K]\bullet} = S_{[K]\bullet}, \quad \widehat{\Sigma_{[K\rangle}\Sigma^{-1}_{\langle K \rangle}} = S_{[K\rangle}S^{-1}_{\langle K \rangle}, \quad K \in \mathcal{J}(\mathcal{K}) \tag{10}$$

によって一意的に構成できることを Andersson and Perlman (1993) が示している．(10) によって得られる Σ の最尤推定量を $\hat{\Sigma}^{\mathrm{mle}}_{\mathcal{K}} \in \boldsymbol{P}_I(\mathcal{K})$ と記すことにする．

注意 3 Lattice conditional indepndence モデルと有向非巡回グラフモデルとの関係は Andersson, et al. (1995, 1996) を参照のこと．

3. 改良された推定量の導出

$GT_I(\mathcal{K})$–共変な推定量の中で最良な推定量を求める．そのために，$n\hat{\Sigma}^{\mathrm{mle}}_{\mathcal{K}} = TT'$ と分解する．ただし，$T \in GT_I(\mathcal{K})$ である．更に，$D = \mathrm{diag}(D_{[K]}|K \in \mathcal{J}(\mathcal{K}))$ で $D_{[K]}$ は $|[K]| \times |[K]|$ の対角行列でその対角成分は

$$d^{-1}_{[K],i} = n + |[K]| - |\langle K\rangle| + \sum(|[L]||L \in \mathcal{M}(K)) - 2i + 1,$$

である．ただし，$i = 1, \ldots, |[K]|, K \in \mathcal{J}(\mathcal{K})$ ，$\mathcal{M}(K) = \{L \in \mathcal{J}(\mathcal{K})|K \subseteq \langle L\rangle\}$ とし，$K \in \mathcal{K}$ に対して，$|K|$ を K の中に含まれる要素の個数とする．

定理 1 ：$\hat{\Sigma}^{\mathrm{m}}_{\mathcal{K}} = TDT'$ はミニマックスである．そのミニマックスリスクは

$$\sum_K \left\{ \sum_{i=1}^{|[K]|} \left(-\log d_{[K],i} - \boldsymbol{E} \log \chi^2_{n+|[K]|-|\langle K\rangle|-2i+1} \right) \middle| K \in \mathcal{J}(\mathcal{K}) \right\}$$

で与えられる．

更に，$\max\{|[K]||K \in \mathcal{J}(\mathcal{K})\} \geq 2$ の場合には $\hat{\Sigma}^{\mathrm{m}}_{\mathcal{K}}$ は一様に改良できる．$S_{[K]\bullet}$ をスペクトル分解して，

$$S_{[K]\bullet} : |[K]| \times |[K]| = O_{[K]} \mathrm{diag}(l_{[K],1}, l_{[K],2,}, \ldots, l_{[K],|[K]|}) O'_{[K]}$$

と書く．ただし，$l_{[K],1} \geq l_{[K],2,} \geq \cdots \geq l_{[K],|[K]|} > 0$ である．

ここで，以下の方程式から $\boldsymbol{P}_I(\mathcal{K})$ に含まれることから LCI モデルの性質 (i) を用いて構成される推定量 $\hat{\Sigma}^{\mathrm{im}}_{\mathcal{K}}$ を考える．

$$\begin{aligned} n\hat{\Sigma}^{\mathrm{im}}_{[K]\bullet} &= O_{[K]} \mathrm{diag}(\psi_{[K],1}, \psi_{[K],2}, \ldots, \psi_{[K],|[K]|}) O'_{[K]} \\ \Sigma^{\mathrm{im}}_{[K\rangle} \widehat{(\Sigma^{\mathrm{im}}_{\langle K\rangle})^{-1}} &= S_{[K\rangle} S^{-1}_{\langle K\rangle}, \quad K \in \mathcal{J}(\mathcal{K}). \end{aligned} \tag{11}$$

ただし，$\psi_{[K],i}(i = 1, 2, \ldots, |[K]|, K \in \mathcal{J}(\mathcal{K}))$ は

$$\{(l_{[K],1}, l_{[K],2,}, \ldots, l_{[K],|[K]|}) \in \Re^{|[K]|} | l_{[K],1} \geq l_{[K],2,} \geq \cdots \geq l_{[K],|[K]|} > 0\}$$

上で定義され，その上で偏微分可能な実数値関数とする．

LCI モデルの性質を利用して，(11) で与えれる推定量 $\hat{\Sigma}^{\mathrm{im}}$ の危険関数を分解することにより，節 1 の結果が適用できる．

定理 2 ：$\Sigma \in \boldsymbol{P}_I(\mathcal{K})$ とする．

$$\begin{aligned} \hat{\rho}(\hat{\Sigma}^{\mathrm{im}}_{\mathcal{K}}, \Sigma) = \sum_K \Biggl\{ \sum_{i=1}^{|[K]|} \Biggl(& 2\psi_{[K],ii} + (e_{[K]} - |[K]| - 1) \frac{\psi_{[K],i}}{l_{[K],i}} \\ & + 2 \sum_{j>i} \frac{\psi_{[K],j} - \psi_{[K],i}}{l_{[K],j} - l_{[K],i}} - \log \psi_{[K],i} \Biggr) \Bigg| K \in \mathcal{J}(\mathcal{K}) \Biggr\}. \end{aligned}$$

ただし，$\psi_{[K],ii} = \partial\psi_{[K],i}/\partial l_{[K],i}, i = 1, 2, \ldots, |[K]|$ ，$e_{[K]} = n - |\langle K\rangle| + \{\sum |[L]||L \in \mathcal{M}(K)\}, K \in \mathcal{J}(\mathcal{K})$ である．

定理 3 ：$\max\{|[K]||K\in\mathcal{J}(\mathcal{K})\}\geq 2$ のとき，

$$\psi_{[K],i}=d_{[K],i}l_{[K],i},\quad i=1,2,\ldots,|[K]|,K\in\mathcal{J}(\mathcal{K})$$

と (11) においておけば，$\hat{\Sigma}^{\mathrm{im}}_{\mathcal{K}}$ はミニマックスで $\hat{\Sigma}^{\mathrm{m}}_{\mathcal{K}}$ を一様に改良する．

$I=\{1,2,3\}$ の場合についてふたつの LCI モデルにおけるミニマックス推定量を定理 2 を使い求めてみる．

<u>その 1</u>：$\mathcal{K}_1=\{\phi,1,12,13,I\}$ とする．したがって，X_1 を与えたときに X_2 と X_3 は条件付独立で，$\mathcal{J}(\mathcal{K}_1)=\{1,12,13\}$ となる．$X^{(1)},X^{(2)},\ldots,X^{(n)}$ を $\boldsymbol{N}(\mathcal{K}_1)$ からのランダム標本とし，$S=\sum_{i=1}^n X^{(i)}\{X^{(i)}\}'=(s_{ij})$ とおく．LCI モデルの性質 (i) から Σ の最尤推定量を求めると

$$n\hat{\Sigma}^{\mathrm{mle}}_{\mathcal{K}_1}=\left[\begin{array}{ccc} s_{11} & s_{12} & s_{13}\\ s_{21} & s_{22} & s_{21}s_{11}^{-1}s_{13}\\ s_{31} & s_{31}s_{11}^{-1}s_{12} & s_{33}\end{array}\right]$$

となる．これを分解して，$n\hat{\Sigma}^{\mathrm{mle}}_{\mathcal{K}_1}=T_1T_1'$ と書く．ただし，

$$T_1=\left[\begin{array}{ccc} s_{11}^{1/2} & 0 & 0\\ s_{21}s_{11}^{-1/2} & t_{22} & 0\\ s_{31}s_{11}^{-1/2} & 0 & t_{33}\end{array}\right]\in\boldsymbol{GT}_I(\mathcal{K}_1)$$

である．ここで，$t_{ii}^2=s_{ii}-s_{i1}s_{11}^{-1}s_{1i}$ とおき，$t_{ii}>0(i=2,3)$ である．ミニマックス推定量は $\hat{\Sigma}^{\mathrm{m}}_{\mathcal{K}_1}=T_1D_1T_1'$ で与えられる．ただし，

$$D_1=\ \mathrm{diag}\ \left(\frac{1}{n+2},\frac{1}{n-1},\frac{1}{n-1}\right).$$

である．更に，これを成分で書けば，

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

となることがわかる．

<u>その 2</u>：$\mathcal{K}_2=\{\phi,1,2,12,I\}$ とする．したがって，X_1 と X_2 は独立で，$\mathcal{J}(\mathcal{K}_2)=\{1,2,I\}$ となる．同様に $X_1,X_2,\ldots,X_n$ を $\boldsymbol{N}(\mathcal{K}_2)$ からのランダム標本とする．$\Sigma(\in\boldsymbol{P}_I(\mathcal{K}_2))$ の最尤推定量は $n\hat{\Sigma}^{\mathrm{mle}}_{\mathcal{K}_2}=T_2T_2'$ で与えられる．ここで，

$$T_2=\left[\begin{array}{ccc} s_{11}^{1/2} & 0 & 0\\ 0 & s_{22}^{1/2} & 0\\ t_{31} & t_{32} & t_{33}\end{array}\right]\in\boldsymbol{GT}_I(\mathcal{K}_2)$$

とし，

$$(t_{31},t_{32})=(\frac{s_{31}}{\sqrt{s_{11}}},\frac{s_{32}}{\sqrt{s_{22}}}),\quad t_{33}^2=s_{33}-\left(\frac{s_{31}^2}{s_{11}}+\frac{s_{32}^2}{s_{22}}\right)$$

とおいた．ミニマックス推定量は $\hat{\Sigma}^{\rm m}_{\mathcal{K}_2}=T_2D_2T_2'$ で与えられる．ただし，

$$D_2=\ \mathrm{diag}\ \left(\frac{1}{n+1},\frac{1}{n+1},\frac{1}{n-2}\right)$$

である． さらに，これを具体的に書けば，

$$\hat{\Sigma}^{\rm m}_{\mathcal{K}_2}=\begin{bmatrix}\frac{s_{11}}{n+1} & 0 & \frac{s_{13}}{n+1}\\ 0 & \frac{s_{22}}{n+1} & \frac{s_{32}}{n+1}\\ \frac{s_{31}}{n+1} & \frac{s_{32}}{n+1} & \frac{1}{n-2}\left\{s_{33}-\frac{3}{(n+1)}\left(\frac{s_{31}^2}{s_{11}}+\frac{s_{32}^2}{s_{11}}\right)\right\}\end{bmatrix}$$

となることがわかる．

<u>その 3</u>：$\mathcal{K}_3=\{\phi,1,12,I\}$ とする．したがって，$\mathcal{J}(\mathcal{K}_3)=\{1,12,I\}$ となる．このとき，$\boldsymbol{P}_I(\mathcal{K}_3)$ は 3×3 の正定置行列の空間（ 2 節の場合）となる．さらに，

$$d^{-1}_{[1],1}=n+2,\ d^{-1}_{[12],2}=n,\ d^{-1}_{[I],1}=n-2$$

となり，$p=3$ の場合の (4) と一致する．

<u>その 4</u>：$\mathcal{K}_4=\{\phi,\ 1,\ 2,\ 3,\ 12,\ 13,\ 23,\ I\}$ とする．したがって，$\mathcal{J}(\mathcal{K}_4)=\{1,\ 2,\ 3\}$ となる．このとき，$\boldsymbol{P}_I(\mathcal{K}_3)$ は 3×3 の正の成分をもつ対角行列の空間（すなわち，各変量は独立）となる．さらに，

$$d^{-1}_{[1],1}=n,\ d^{-1}_{[2],2}=n,\ d^{-1}_{[3],1}=n$$

となる．

4. むすびにかえて

節 4 において，最尤推定量を含む $GL_I(\mathcal{K})$ -共変な推定量の族の中で最良な推定量 $\hat{\Sigma}^m(\mathcal{K})$ を構成することで最尤推定量の非許容性が示された．しかし，定理 2 の手法で $\hat{\Sigma}^m(\mathcal{K})$ を改良できない場合（すなわち，$\max\{|[K]|:K\in\mathcal{J}(\mathcal{K})\}=1$ の場合）には，$\hat{\Sigma}^m(\mathcal{K})$ が許容的かどうかは不明である．多変量正規分布の分散共分散行列の推定問題において Bayes 推定の観点からは Haff (1991) による the variational form of Bayes estimators という結果や Leonard and Hsu (1992) や Yang and Berger (1994) の論文があるが，許容性という視点からはほとんど研究がされていないので，この点は課題として残されている．

また，Anderson (1984) に代表される多変量解析にあるモデルを広く拡張した多変量統計モデルを提案とその上で推測理論の展開が Andersson and Perlman (1993) の他に Andersson and Madsen (1998)，Massam and Neher (1998) がある． これらの論文では invariance がモデルの構成に重要な役割を果たしているので，この新しい多変量統計モデルのもとでも Stein の着想はうまく機能することが期待される．

参考文献

Anderson, T.W. (1984): *Introduction to Multivariate Statistical Analysis*, 2nd edtion. New York Wiley.

Andersson,S.A., Madigan, D., Perlman, M.D., and Triggs, C.M. (1995): On the relation between conditional independence models determined by finite distributive lattices and by directed acyclic graphs, *J. Statist Planing Infer.* **48**, 25–46.

Andersson,S.A., Madigan, D., Perlman, M.D., and Triggs, C.M. (1997): A graphical characterization of lattice conditional independence models, *Ann. Math. and Artifical Intellignece* **21**, 27–50.

Andersson, S.A. and Madsen, J. (1998): Symmetry and lattice conditional independence in a multivariate normal distribution, *Ann. Statist.* **26**, 525–572.

Andersson, S.A. and Perlman, M.D. (1993): Lattice models for conditional independence in a multivariate normal distribution, *Ann. Statist.* **21**, 1318–1358.

Bonder, J. and Milnes, P. (1981): Amenability: A survey for statistical applications of Hunt-Stein and related conditions on groups, *Z. Wahrsch. verw. Gebiete* **57**, 103–128.

Dey, D.K. and Srinivasan, C. (1985): Estimation in a covariance matrix under Stein's loss, *Ann. Statist.* **13**, 1581–1591.

Eaton, M.L. (1970): Some problems in covariance estimation. Technical Report No. 49, Department of Statistics, Stanford University.

Eaton, M.L. (1989): Group invariance applications in Statistics, Regional conference series in Probability and Statistics, vol. 1, IMS, California.

Haff, L. R. (1976): Estimation of the inverse covariance matrix: Random mixtures of the inverse Wishart matrix and identity, *Ann. Statist.*1 **6**, 1264–1276.

Haff, L. R. (1991): The variational form of certain Bayes estimators, *Ann. Statist.* **19**, 1163–1190.

James, W. and Stein, C. (1961): Estimation with quadratic loss, *in Proc. Fourth Berkeley Symp. Math. Statist. Probab* **1**, Univ. California Press, 361–380.

Kiefer (1957): Inavriance, minimaax, sequential estimation, and continuous time processes,, *Ann. Math. Statist.* **28**, 573–601.

Konno, Y. (2001): Inaddmissibility of the maximum likelihood estimator of normal covariance matrices with the lattice conditional independence, To appear in it J. Multivariate Anal.

Kubokawa, T. and Srivatava, M.S. (1999): Improved nonnegative estimation of multivariate components of variance, *Ann. Statist.* **27**, 2008–2032.

Lauritzen, S.L. (1996): Graphical models, Oxford statistical science series, Claredon press• Oxford

Leonard, T. and Hsu, J.S.J. (1992): Bayesian inference for a covariance matrix. it Ann. Statist. **20**, 1669–1690.

Massam, H. and Neher, E. (1998): Estimation and testing for lattice conditional independence models on Euclidean Jordan algebras, *Ann. Statist.* **26**, 1051–1082.

宮川雅巳 (1997): グラフィカルモデリング，朝倉書店 Muirhead, R.J. (1982): *Aspects of multivariate statistical analysis*, Wiley.

Stein, C. (1973): Estimation of mean of a multivariate normal distribution, *in Proceedings of Prague Symposium on Asymptotic Statistics*, pp. 345–381.

竹村彰通 (1991): 多変量推測統計の基礎, 共立出版.

Yang, R. and Berger, J.O. (1994): Estimation of a covariance matrix using the refernce prior. *Ann. Statist.* **22**, 1195–1211.

日本品質管理学会　テクノメトリックス研究会編 (1999): グラフィカルモデリングの実際，日科技連

補遺

Wishart identity. 危険関数の不偏推定量を導出するときに重要な道具が部分積分の式である．カイ二乗分布に関する部分積分の式を多次元の分布である Wishart 分布に関する部分積分の式に拡張したものが Wishart identity である．これは Stein (1975) と Haff (1976) に独立に示されている．また，楕円分布のもとでの拡張された公式を Kubokawa and Srivastava(1999) が導出している．

補題 1 (Wishart identity) $n > p+1$ とする．$h : p \times p = (h_{ij})$，$h_{ij}$ は正定値行列上の実数値関数とし，$S : p \times p \sim W_p(n, \Sigma)$ とする．このとき，

$$\boldsymbol{E}[\mathrm{tr}\{\Sigma^{-1}h(S)\}] = 2\boldsymbol{E}[\mathrm{tr}\{D_S h(S)\}] + (n-p-1)\boldsymbol{E}[\mathrm{tr}\{S^{-1}h(S)\}]$$

が成立する．ただし，$D_S = (d_{ij})$ で $d_{ij} = (1/2)(1+\delta_{ij})\partial/\partial S_{ij}$ である．ここで，S_{ij} は行列 S の (i,j) 成分とし，

$$\delta_{ij} = \begin{cases} 1 & \text{もし} \quad i = j \\ 0 & \text{その他} \end{cases}$$

とした.

Calculus on eigenstructure. S の固有根を $\ell_1, \ell_2, \ldots, \ell_p (\ell_1 \geq \ell_2 \geq \cdots \geq \ell_p > 0)$ とかき，$L = \mathrm{diag}(\ell_1, \ell_2, \ldots, \ell_p)$ とおく．O は $p \times p$ の直行行列で $O'SO = L$ を満足するものとする．更に，実数値関数 $\psi_i(\ell_1, \ell_2, \ldots, \ell_p)$ は $\{(\ell_1, \ell_2, \ldots, \ell_p) \in \Re^p : \ell_1 \geq \ell_2 \geq \cdots \geq \ell_p > 0\}$ 上で定義され，その上で偏微分可能とし，$\Psi(L) = \mathrm{diag}(\psi_1(L), \psi_2(L), \ldots, \psi_p(L))$ と $\psi_{ii}(L) = \partial\psi_i(L)/\partial\ell_i, i = 1, 2, \ldots, p$ とする．簡単に，$\psi_i = \psi_i(L)$, $\psi_{ii} = \psi_{ii}(L)$, $i = 1, 2, \ldots, p$ と記すことにする.

補題 2 ：

$$\mathrm{tr}[D_S\{O'\Psi O\}] = \sum_{i=1}^{p}\left\{\psi_{ii} + \sum_{j>i}\frac{\psi_i - \psi_j}{\ell_i - \ell_j}\right\}.$$

例の計算

• その 1 の計算

$$\begin{array}{r|rrrrr} K & \emptyset & 1 & 12 & 13 & I \\ \langle K\rangle & \emptyset & \emptyset & 1 & 1 & \emptyset \\ [K] & \emptyset & 1 & 2 & 3 & \emptyset \end{array}$$

から

$$\mathcal{J}(\mathcal{K}_1) = \{1,\ 12,\ 13\}$$

がわかる.

K	$[K]$	$\mathcal{M}(K)$	$\sum(\lvert[L]\rvert \mid L \in \mathcal{M}(K))$	$d^{-1}_{[K],1}$
1	1	12, 13	$2 = \lvert[12]\rvert + \lvert[13]\rvert$	$n+1-0+2-2+1 = n+2$
12	2	なし	0	$n+1-1+0-2+1 = n-1$
13	3	なし	0	$n+1-1+0-2+1 = n-1$

• その 2 の計算

$$\begin{array}{r|rrrrr} K & \emptyset & 1 & 2 & 12 & I \\ \langle K \rangle & \emptyset & \emptyset & \emptyset & 12 & 12 \\ [K] & \emptyset & 1 & 2 & \emptyset & 3 \end{array}$$

から

$$\mathcal{J}(\mathcal{K}_1) = \{1,\, 2,\, I\}$$

がわかる.

K	$[K]$	$\mathcal{M}(K)$	$\sum(\|[L]\|\|L \in \mathcal{M}(K))$	$d^{-1}_{[K],1}$
1	1	I	$1 = \|[I]\|$	$n+1-0+1-2+1 = n+1$
2	2	I	$1 = \|[I]\|$	$n+1-0+1-2+1 = n+1$
I	3	なし	0	$n+1-2+0-2+1 = n-2$

• その 3 の計算

$$\begin{array}{r|rrrr} K & \emptyset & 1 & 12 & I \\ \langle K \rangle & \emptyset & \emptyset & 1 & 12 \\ [K] & \emptyset & 1 & 2 & 3 \end{array}$$

から

$$\mathcal{J}(\mathcal{K}_3) = \{1,\, 12,\, I\}$$

がわかる.

K	$[K]$	$\mathcal{M}(K)$	$\sum(\|[L]\|\|L \in \mathcal{M}(K))$	$d^{-1}_{[K],1}$
1	1	$12,\, I$	$2 = \|[12]\| + \|[I]\|$	$n+1-0+2-2+1 = n+1$
12	2	I	$1 = \|[I]\|$	$n+1-1+1-2+1 = n$
I	3	なし	0	$n+1-2+0-2+1 = n-2$

• その 4 の計算

$$\begin{array}{r|rrrrrrrr} K & \emptyset & 1 & 2 & 3 & 12 & 13 & 23 & I \\ \langle K \rangle & \emptyset & \emptyset & \emptyset & \emptyset & 12 & 13 & 23 & I \\ [K] & \emptyset & 1 & 2 & 3 & \emptyset & \emptyset & \emptyset & \emptyset \end{array}$$

から

$$\mathcal{J}(\mathcal{K}_3) = \{1,\, 2,\, 3\}$$

がわかる.

K	$[K]$	$\mathcal{M}(K)$	$\sum(\|[L]\|\|L \in \mathcal{M}(K))$	$d^{-1}_{[K],1}$
1	1	なし	0	$n+1-0+0-2+1 = n$
2	2	なし	0	$n+1-0+0-2+1 = n$
3	3	なし	0	$n+1-0+0-2+1 = n$