Stein's approach for constructing improved estimators for matrix estimation problems.

今野良彦[1]　千葉大学理学部数学・情報数理学科

0. はじめに

多変量正規分布の平均ベクトルの最尤推定量がいわゆる縮小型の推定量で平均２乗誤差の和の下で一様に改良できること（いわゆる Stein 現象または Stein 効果）が示されて以来、様々な同時推定問題で同様の現象が研究されている．この報告では Stein 現象の中でも matrix estimation problems と呼ばれる推定問題に焦点を絞ってその議論をおこなう．Stein 現象のいろいろな方向への最近の発展に関しては、久保川 (1996)、Lehmann and Casella (1998)、Robert (1994), 篠崎 (1991) 等を参照のこと．ここで例として取り上げる推定問題は簡単な推定量を除いて危険関数を直接評価するのが非常に困難である．しかし、推定量の危険関数の直接評価をするのではなく、ある族に含まれる推定量の危険関数にたいする不偏推定量を導出し、個々の推定量の危険関数の不偏推定量を比較することにより、推定量の危険関数を評価する方針が Stein (1973) に提示されてから、正規分布の分散共分散行列の推定問題を含むいくつかの推定問題においてこの方針に沿い、改良型の推定量が提案されている．Stein の方針に沿い、いくつかの推定問題において縮小型の推定量の構成を議論する．1 節においては、基本的な例である多変量正規分布の平均行列と分散共分散行列の推定問題を定式化し、2 節では Stein 方針を解説し、3 節ではこの方針を実行するための技術的な道具をあげ、4 節では基本的な推定問題をふくむ 4 つの推定問題において縮小型の推定量の構成を議論する．

1. 推定問題

$X: m\times p$ と $S: p\times p$ は独立とし、

$$X \sim N_{m\times p}(\mu, T\otimes\Sigma), \qquad S \sim W_p(n,\Sigma) \tag{1}$$

[1] email address: konno@math.s.chiba-u.ac.jp

とする．ただし、μ は $m\times p$ の未知の平均行列とし、T は $m\times m$ の既知の正定値行列、Σ は $p\times p$ の未知の正定値行列とし、X の確率密度関数は

$$\frac{1}{(\sqrt{2\pi})^{mp}(\det T)^{(p/2)}(\det\Sigma)^{(m/2)}}\mathrm{etr}\left[-\frac{1}{2}T^{-1}(x-\mu)\Sigma^{-1}(x-\mu)'\right]$$

で与えられる．$W_p(n,\Sigma)$ は自由度 n、平均 $n\Sigma$ の Wishart 分布とする．ここで、$\exp\mathrm{tr}(\cdot)$ を $\mathrm{etr}(\cdot)$、行列 A の転置行列を A' とそれぞれ記した．

未知の母数 μ と Σ の推定問題を個々に考えることにする．

例 1（平均行列の推定）：μ の推定問題を扱うために損失関数として

$$L_1(\hat{\mu},\mu;\Sigma)=\mathrm{tr}[Q^{-1}(\hat{\mu}-\mu)\Sigma^{-1}(\hat{\mu}-\mu)']$$

を採用する．ただし、Q は既知の $m\times m$ の正定値行列で、$\hat{\mu}$ は (X,S) に基づく μ の推定量である．

例 2（分散共分散行列の推定）：Σ の推定問題に対して

$$L_2(\hat{\Sigma},\Sigma)=\mathrm{tr}(\hat{\Sigma}\Sigma^{-1})-\log\det(\hat{\Sigma}\Sigma^{-1})-p \tag{2}$$

を用いる．ここで、$\hat{\Sigma}$ は Σ の推定量とする．

それぞれの損失関数を (1) に関して期待値を取ったものを危険関数とよび、$R_1(\hat{\mu},\mu)$ と $R_2(\hat{\Sigma},\Sigma)$ 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． 他の損失関数の下での議論として、例 1 では、Honda (1991)、例 2 では、Haff (1991) 等がある．

2. 基本的なアプローチ

改良型の推定量を導出する基本的な方針は以下の段階を踏むことになる．

1. 考察の対象の統計モデル（尤度関数）を分解し、条件付き独立性等を用いてモデルの分解に対応して推定量や危険関数の分解を行い、より簡単な統計モデルにおける推定問題に帰着させる．

2. モデルの構造を反映した共変性（推移的に作用していないもの！）を用いて、推定の族を制限する．

3. 共変推定量の族を陽に求める．

4. Normal identity や Wishart identity 等の部分積分の式と calculus on eigenstructure を用いて、求めた共変推定量に対する危険関数の不偏推定量を計算する．

5. 危険関数の不偏推定量から導出された偏微分不等式を利用して、改良型推定量を導く．

注意 2 ：個々の問題でこの方針に沿って改良型推定量を導出するのはモデルの構造に対する深い洞察もしくは直観的な解釈が必要となる．通常の推定量（たとえば、最尤推定量）は、母数空間に推移的に作用する群のもとでも共変なので、この共変構造に対して不変な損失関数（この報告で採用している損失関数）を採用しているがぎりは定数の危険関数をもつ．しかし、段階 2 において得られる推定量の族は、母数空間に推移的に作用しない群のもとの共変推定量なので、推定量の危険関数を母数に依存するものとなる．また、推定量を制限をするときに、推定量の族を広くとれば、それだけ段階 4 で得られる危険関数の不偏推定量が複雑になり、逆にその扱いが困難になる．

注意 3 ：前章の基本的な推定問題では、段階 1 は不要になる．

注意 4 ：正規分布の平均ベクトルの推定では、危険関数の不偏推定量の評価を経由せずに危険関数をより直接的に評価することは可能であるが、この報告で扱う正規分布の分散共分散行列の推定のように行列母数の同時推定問題では、危険関数の不偏推定量を用いての評価を経由しない直接的な評価は困難である．しかし、例 1 の推定問題において、Zidek (1978) はある共変推定量の族に対する危険関数を zonal 多項式を用いた表現を与えているが、改良型の推定量を導出するには見通しの良い表現ではなかった．

3. 技術的な補題

Normal and Wishart identities．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．この式の意味を理解するために 1 次元の分布にたいする式をあげる．

補題１：確率変数 Y_1 は $N(\xi_1, 1)$ に従うとし、$g_1 : \Re \mapsto \Re$ は任意の有界閉区間で絶対連続で、$E_{\xi_1}[|\dot{g}_1(Y_1)|] < \infty$ なる関数とする．このとき、

$$E_{\xi_1}[g_1(Y_1)(Y_1 - \xi_1)] = E_{\xi_1}[\dot{g}_1(Y_1)]$$

が成立する．ここで、$\dot{g}_1$ は g_1 の導関数とした．

補題２：確率変数 V が $\sigma^2\chi_n$ に従い、$h_1 : \Re^+ \mapsto \Re$ は任意の有界閉区間で絶対連続で、$E[V^{-1}h_1(V)]$ と $E[|\dot{h}_1(V)|]$ は有限とする．このとき、

$$E[\frac{h_1(V)}{\sigma^2}] = 2E[\dot{h}_1(V)] + (n-2)E[\frac{h_1(V)}{V}]$$

が成立する．ここで、χ_n は自由度 n のカイ二乗分布とし、$\dot{h}_1$ は h_1 の導関数とした．

１次元の正規分布とカイ二乗分布に関する部分積分の式をそれぞれ多次元の分布である多変量変量の正規分布 $N(\xi, \Sigma)$ と Wishart 分布に関する部分積分の式に拡張したものが Normal identity と Wishart identity である．前者は Stein (1973) により導入され、後者が Stein (1975) と Haff (1977) に独立に示されている．

Normal identity：$Y = (Y_1, Y_2, \ldots, Y_p)' \sim N_p(\xi, \Sigma)$ とし、$g : p \times 1 = (g_1, \ldots, g_p)'$、$g_i : \Re^p \mapsto \Re (i = 1, 2, \ldots, p)$ とする．このとき、

$$E[\Sigma^{-1}(Y - \xi)g'(Y)] = E[\nabla_Y g'(Y)]$$

が成立する．ただし、$\nabla_Y = (\partial/\partial Y_1, \partial/\partial Y_2, \ldots, \partial/\partial Y_p)'$ である．

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、

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

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_1 と S_2 を $p\times p$ の正定値行列とし、$S_1S_2^{-1}$ の固有根を $f_1,f_2,\ldots,f_p(f_1\geq f_2\geq\cdots\geq f_p>0)$ とかき、$F=\mathrm{diag}(f_1,f_2,\ldots,f_p)$ とおく．A は $p\times p$ の正則行列で $A'S_2A=I_p$ と $A'S_1A=F$ を満足するものとする．D_{S_1} と D_{S_2} を S_1 と S_2 の成分をそれぞれ用いて D_S と同様に定義したものとする．更に、実数値関数 $\varphi_i(f_1,f_2,\ldots,f_p)$ は $\{(f_1,f_2,\ldots,f_p)\in\Re^p: f_1\geq f_2\geq\cdots\geq f_p>0\}$ 上で定義され、その上で偏微分可能とし、$\Phi=\mathrm{diag}(\varphi_1,\varphi_2,\ldots,\varphi_p)$ と $\varphi_{ii}=\partial\varphi_i/\partial f_i, i=1,2,\ldots,p$ と書くことにする．

補題 3 ：

$$
\begin{aligned}
\mathrm{tr}[D_{S_1}\{(A')^{-1}\Phi A^{-1}\}] &= \sum_{i=1}^{p}\left\{\varphi_{ii}+\sum_{j>i}\frac{\varphi_i-\varphi_j}{f_i-f_j}\right\}\\
\mathrm{tr}[D_{S_2}\{(A')^{-1}\Phi A^{-1}\}] &= \sum_{i=1}^{p}\left\{p\varphi_i-f_i\varphi_{ii}-\sum_{j>i}\frac{f_i\varphi_i-f_j\varphi_j}{f_i-f_j}\right\}\\
D_{S_1}f_i &= a_ia_i' \quad i=1,2,\ldots,p
\end{aligned}
$$

ただし、a_i は A の i 列ベクトルとする．

4. リスクの不偏推定量と推定量の構成

ここでは、2 節にある方針に沿い、1 節にあげた例を含む推定問題について改良型の推定量を構成する．

例 1 （続き）：Bilodeau and Kariya (1989) では、μ の推定量として mp 個の (X,S) 上の実数値関数によって構成される推定量を考えているが、ここでは対象の族を狭めて、

$$\hat{\mu}(X,S)=X+\frac{1}{2}T\{\nabla_X h(F)\}S \tag{3}$$

なる形の推定量を考える．ただし、$X=(x_{ij})$ に対して、$\nabla_X: m\times p=(\partial/\partial x_{ij})$、$F=\mathrm{diag}(f_1,f_2,\ldots,f_{m\wedge p})$ とし、$f_1,f_2,\ldots,f_{m\wedge p}(f_1\geq f_2\geq\cdots\geq f_{m\wedge p}>0)$ は $X'(TQT)^{-1}XS^{-1}$ の固有根で、h は $\{(f_1,f_2,\ldots,f_{m\wedge p})\in\Re^{m\wedge p}: f_1\geq f_2\geq\cdots\geq f_{m\wedge p}>0\}$ 上で定義されている実数値関数で、2 回偏微分可能とする．ただし、$m\wedge p=\min(m,p)$ である．

推定量 (3) の危険関数の不偏推定量を表現するために以下を定義する．

記号：自然数 a,b,c と推定量 (3) に対して、

$$J(h;a,b,c) = \sum_{i=1}^{c}\left\{2(b-c+1)h_i + 4f_ih_{ii} + 4\sum_{j>i}\frac{f_ih_i - f_jh_j}{f_i - f_j}\right.$$
$$\left. + (a+c-3)f_ih_i^2 - 4f_i^2h_ih_{ii} - 2\sum_{j>i}\frac{(f_ih_i)^2-(f_jh_j)^2}{f_i-f_j}\right\}$$

とおく．ただし、$h_i = \partial h/\partial f_i, h_{ii} = \partial^2 h/\partial f_i^2 (i=1,2,\ldots,c)$ である．

定理１：$|p-m|>1$ かつ $n>p+1$ とする．このとき、推定量 (3) の危険関数の不偏推定量は

$$\hat{R}(\hat{\mu},\mu) = p\mathrm{tr}(QT) + J(h; n+\{(m-p)\wedge 0\}, m\vee p, m\wedge p) \quad (4)$$

で与えられる．すなわち、

$$R(\hat{\mu},\mu) = E[\hat{R}(\hat{\mu},\mu)]$$

である．ここで、$m\vee p = \max(m,p)$ とした．

上の定理より、推定量 (3) が X を一様に改良するための条件は $J\leq 0$ を満たすような h であればよいことがわかる．

たとえば、h として、

$$h = \{(m\vee p)-(m\wedge p)-1\}\sum_{i=1}^{m\wedge p}\log f_i$$

と取れば、簡単な計算により (4) の右辺の２項目が負になることがわかる．これを書き換えるといわゆる Efron - Morris 型の推定量

$$\hat{\mu}(X,S) = \begin{cases} X + (m-p-1)Q^{-1}T^{-1}X\{X'(TQT)^{-1}X\}^{-1}S \\ \qquad (m>p+1 \text{ の場合}) \\ X + (p-m-1)T\{XS^{-1}X'\}^{-1}X \\ \qquad (p>m+1 \text{ の場合}) \end{cases}$$

となることがわかる．

注意 5 ：定理 1 におけるパラメータの対応は $X'XS^{-1}$ の固有根の同時確率密度関数のパラメータの対応と同じである．

$$\begin{array}{ccc} m>p+1 \text{ の場合} & & p>m+1 \text{ の場合} \\ p & \mapsto & m \\ m & \mapsto & p \\ n & \mapsto & n+m-p. \end{array}$$

注意 6 ：定理 1 の結果は分散共分散が既知の場合の多変量正規分布の平均行列の推定問題における Stein (1973) の結果の拡張とみなすことができる．また、Kuriki (1993) は分散共分散行列が既知の場合の多変量正規分布の $p\times p$ の平均行列が歪対称な構造をもつときの平均行列の推定問題において、Stein (1973) の結果に対応するものを得ている．

注意 7 ：Bilodeau and Kariya (1989) により得られた危険関数の不偏推定量と比較すると推定量の族を制限したことにより、改良型の推定量を構成するにはより見通しの良い式になっている．ただし、彼らが危険行列のもとで X を一様に改良する推定量を提案しているが、そのなかのいくつかの推定量はわれわれの族に含まれないので、この結果は適用できない．

例 2（続き）: James and Stein (1961) は上三角行列の群のもとで共変な推定量の族（母数空間に推移的に作用するので、その危険関数は定数）の中で最良のものを求めている．Hunt–Stein の定理（Bonder and Milnes (1981) を参照）を利用することで、これはミニマックス推定量であり、ミニマックスリスクが

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

で与えられることを彼らは示している．

更に、Stein (1975) では、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

$$\hat{\Sigma}=O\mathrm{diag}(\psi_1,\psi_2,\ldots,\psi_p)O' \tag{5}$$

の形で与えられること示している．ここで、

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

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

推定量 (5) に対する損失関数は

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

なので、

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

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、Wishart identity と補題 3 を使えば、Stein (1975) による ρ の不偏推定量

$$\hat{\rho}(\hat{\Sigma},\Sigma) = \sum_{i=1}^{p} \left[2\psi_{ii} + (n-p-1)\frac{\psi_i}{l_i} + 2\sum_{i>j} \frac{\psi_i - \psi_j}{l_i - l_j} - \log \psi_i \right] \quad (6)$$

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

注意 8 ：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．また、Sheena (1995) は、S の固有根の同時確率密度関数から (6) をより直接的な方法で導出している．

ここで、(5) において、

$$\psi_i = d_i l_i, \quad d_i = \frac{1}{n+p+1-2i} \quad (i=1,2,\ldots,p) \qquad (7)$$

とおけば、ミニマックス推定量を得る．

しかし、これは $\psi_1 \geq \psi_2 \geq \cdots \geq \psi_p$ を常に満足するとはかぎらないので、非許容的になることを Sheena and Takemura (1992) は示し、(7) で与えられるミニマックス推定量を一様に改良するものを得られている．一方、Perron (1992) は上三角行列の群の下で共変なミニマックス推定量を直交行列の群上の一様分布で平均を取り修正をおこなった推定量の近似解を求めることにより、別のタイプのミニマックス推定量を求めている．

もっとも簡単な $p=2$ の場合を考えると、Perron の推定量は (5) において

$$\begin{aligned}\psi_1 &= \left(\frac{l_1}{l_1+l_2}d_1+\frac{l_2}{l_1+l_2}d_2\right)l_1\\ \psi_2 &= \left(\frac{l_2}{l_1+l_2}d_1+\frac{l_1}{l_1+l_2}d_2\right)l_2\end{aligned}$$

とおけばよいことがわかる．$p=2$ の場合において、

$$\begin{aligned}\psi_1 &= \left(\frac{\sqrt{l_1}}{\sqrt{l_1}+\sqrt{l_2}}d_1+\frac{\sqrt{l_2}}{\sqrt{l_1}+\sqrt{l_2}}d_2\right)l_1\\ \psi_2 &= \left(\frac{\sqrt{l_2}}{\sqrt{l_1}+\sqrt{l_2}}d_1+\frac{\sqrt{l_1}}{\sqrt{l_1}+\sqrt{l_2}}d_2\right)l_2\end{aligned}$$

とすれば、目標の推定量を明示的に与えることを Takemura (1984) が示している．

例 3（MANOVA – GMANOVA 混合モデルの回帰係数行列の推定）: ここではモデル

$$Z: n\times p = A_1\beta_1C_1+A_2\beta_2+\epsilon,\quad \epsilon\sim N_{n\times p}(0,I_n\otimes\tilde{\Sigma})$$

を考える．ただし、$A_i: n\times a_i (i=1,2)$ と C_1 は既知の行列でそれぞれのランクを a_i と c とする．$\beta_1: a_1\times c_1$ と $\beta_2: a_2\times p$ は未知の回帰係数行列とする．更に仮定として、$A: n\times(a_1+a_2)=[A_1,A_2]$ のランクはa_1+a_2 とし、$\tilde{\Sigma}: p\times p$ は未知の正定値行列とする．

(β_1,β_2) の同時推定問題を考えるために損失関数として、

$$L((\hat{\beta}_1,\hat{\beta}_2),(\beta_1,\beta_2)) = \mathrm{tr}[\{A_1(\hat{\beta}_1-\beta_1)C_1+A_2(\hat{\beta}_2-\beta_2)\}\tilde{\Sigma}^{-1}\{\cdots\}']$$

を採用する．

一方、(β_1,β_2) の最尤推定量は

$$\begin{aligned}\hat{\beta}_1^{\mathrm{mle}} &= (A_1'PA_1)^{-1}A_1'PZ\tilde{S}^{-1}C_1'(C_1\tilde{S}^{-1}C_1')^{-1}\\ \hat{\beta}_2^{\mathrm{mle}} &= (A_2'A_2)^{-1}A_2'[Z-A_1\hat{\beta}_1^{\mathrm{mle}}C_1]\end{aligned}$$

で与えられる．ただし、

$$\tilde{S}: p\times p = Z'[I_n-A(A'A)^{-1}A']Z,\quad P: n\times n = I_n-A_2(A_2'A_2)^{-1}A_2'$$

とした．

推定問題を分解するためにつぎの記号を定義する．

$$\begin{aligned} V : a_1 \times p &= (A_1' P A_1)^{-1} A_1' P Z \\ U : a_2 \times p &= (A_2' A_2)^{-1} A_2' Z \end{aligned}$$

としたとき、$V, U, \tilde{S}$ は互いに独立で

$$\begin{aligned} V &\sim N_{a_1 \times p}(\beta_1 C_1, (A_1' P A_1)^{-1} \otimes \tilde{\Sigma}) && (8) \\ U &\sim N_{a_2 \times p}(\eta, (A_2' A_2)^{-1} \otimes \tilde{\Sigma}) && (9) \\ & \quad \text{ただし、} \eta = \beta_2 + (A_2' A_2)^{-1} A_2' A_1 \beta_1 C_1 \\ \tilde{S} &\sim W_p(n - (a_1 + a_2), \tilde{\Sigma}) \end{aligned}$$

に従う．更に、$G : c_1 \times p = (C_1 C_1')^{-1} C_1$ とおき、$G_o : (p - c_1) \times p$ は $GG_o' = 0$ かつ $G_o G_o' = I_{p-c_1}$ を満足するものとし、

$$V_1 : a_1 \times c_1 = VG', \quad V_2 : a_2 \times (p - c_1) = VG_o'$$

とおき、

$$S = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} = \begin{bmatrix} G\tilde{S}G' & G\tilde{S}G_o' \\ G_o\tilde{S}G' & G_o\tilde{S}G_o' \end{bmatrix}$$

とし、

$$S_{11\cdot 2} = S_{11} - S_{12} S_{22}^{-1} S_{21}$$

と定める．また、Σ と $\Sigma_{11\cdot 2}$ も G と G_o を用いて $\tilde{\Sigma}$ から同様に定める．

$(V_1, V_2, S_{11\cdot 2}, S_{22}^{-1} S_{21}, S_{22})$ と (U, S) はそれぞれ (β_1, Σ) と (β_2, Σ) の十分統計量なので、推定量の族として

$$\begin{aligned} \hat{\beta}_1(B_1) &= \hat{\beta}_1^{\text{mle}} + B_1(V_1, V_2, S_{11\cdot 2}, S_{22}^{-1} S_{21}, S_{22}) \\ \hat{\beta}_2(B_2) &= \hat{\beta}_2^{\text{mle}} + B_2(U, S) \end{aligned}$$

を考える．ただし、$B_1 : a_1 \times c_1$ と $B_2 : a_2 : \times p$ とする．すると $\hat{\beta}_1(B_1)$ と $\hat{\beta}_2(B_2)$ の危険関数は

$$R((\hat{\beta}_1(B_1), \hat{\beta}_2(B_2)), (\beta_1, \beta_2)) = R_1 + R_2$$

となる．ここで、

$$
\begin{aligned}
R_1 &= E[\mathrm{tr}\{(A_1'PA_1)(\hat{\beta}_1^{\mathrm{mle}} + B_1 - \beta_1)\Sigma_{11\cdot 2}^{-1}(\cdots)'\}] \\
R_2 &= E[\mathrm{tr}\{(A_2'A_2)(U + B_2 - \eta)\Sigma^{-1}(U + B_2 - \eta)'\}]
\end{aligned}
$$

となる．これより、(β_1, β_2) の同時推定問題はふたつの独立した推定問題に分解ができる．

$$
[V_1, V_2] \sim N_{a_1 \times p}([\beta_1, 0], (A_1'PA_1)^{-1} \otimes \Sigma) \tag{10}
$$

と (9) に注意すれば、ひとつは GMANOVA モデル (10) における回帰係数行列 β_1 の推定を危険関数 R_1 のもとで考えた問題となり、他方は MANOVA モデル (9) における回帰係数行列 η の推定を危険関数 R_2 のもとの問題となる．後者については例 1 の結果を直接適応可能であり、前者については GMANOVA モデルを条件付き独立性で分解し、それに対応する推定量の族を考えることで β_1 の推定問題をさらに分解し、$\beta_1 + V_2\Sigma_{22}^{-1}\Sigma_{21}$ と $\Sigma_{22}^{-1}\Sigma_{21}$ に関する (V_2, S_{22}) を条件付きした推定問題として別個に扱うことができることを Kariya, et al (1996) は示している．
例 4（LCI 構造をもつ分散共分散行列の推定）: 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 $N(\mathcal{K})$ を、

$$
\begin{aligned}
&\text{すべての } L, M \in \mathcal{K} \text{ に対して、} X_{L\cap M} \text{ を与えたときに} \\
&X_L \text{ と } X_M \text{ は条件付き独立}
\end{aligned} \tag{11}
$$

であるような多変量正規分布 $N(0, \Sigma)$ の集合とする．また、$P_I(\mathcal{K})$ をすべての $L, M \in \mathcal{K}$ に対して、(11) 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 $N(\mathcal{K})$ 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、$\langle K \rangle = \cup(\tilde{K} \in \mathcal{K} | \tilde{K} \subset K$　かつ　$\tilde{K} \neq K)$ 、$[K] = K \backslash \langle K \rangle$ とする．したがって、$K = \langle K \rangle \cup [K]$ かつ $\langle K \rangle \cap [K]$

は空集合である．更に、

$$\mathcal{J}(\mathcal{K}) = \{K \in \mathcal{K} | [K] \text{ は空でない}\}$$

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 $K, L \in I$ に対して、

$$A_{K \times L} = (a_{i_1 i_2} | (i_1, i_2) \in K \times L)$$

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

$A \in \boldsymbol{M}_I$、$K \in \mathcal{K}$ 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、$\Sigma_{[K]\bullet} = \Sigma_{[K]} - \Sigma_{[K \rangle} \Sigma_{\langle K \rangle}^{-1} \Sigma_{\langle K]}$ 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

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

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 $p \times p$ 行列と正の値の対角成分をもつ $\mathcal{K}$ - preserving な下三角行列の群とする．

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) $\boldsymbol{P}_I(\mathcal{K}) \mapsto \mathrm{X}(\boldsymbol{M}_{[K] \times \langle K \rangle} \times \boldsymbol{P}_{[K]} | K \in \mathcal{J}(\mathcal{K}))$ 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、

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

を満足することである．

(iii) $\Sigma \in P_I(\mathcal{K})$ 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) $GL_I(\mathcal{K})$ による $P_I(\mathcal{K})$ への作用 $(A, \Sigma) \mapsto A\Sigma A'$ は推移的である．ただし、$A \in GL_I(\mathcal{K}), \Sigma \in P_I(\mathcal{K})$ である．

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

$$n\hat{\Sigma}_{[K]\bullet} = S_{[K]\bullet}, \quad \widehat{\Sigma_{[K\rangle}\Sigma_{\langle K\rangle}^{-1}} = S_{[K\rangle}S_{\langle K\rangle}^{-1}, \quad K \in \mathcal{J}(\mathcal{K}) \qquad (12)$$

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．(12) によって得られる Σ の最尤推定量を $\hat{\Sigma}_{\mathcal{K}}^{\mathrm{mle}} \in P_I(\mathcal{K})$ と記すことにする．例 2 の場合と同様、$GT_I(\mathcal{K})$-共変な推定量の中で最良な推定量がミニマックス推定量になることがわかる．そこで、$n\hat{\Sigma}_{\mathcal{K}}^{\mathrm{mle}} = TT'$ 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

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

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

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

で与えられる．

更に、$\max\{|[K]| | K \in \mathcal{J}(\mathcal{K})\} \geq 2$ の場合には $\hat{\Sigma}_{\mathcal{K}}^{\mathrm{m}}$ 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 $P_I(\mathcal{K})$ 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 $\hat{\Sigma}_{\mathcal{K}}$ を考える．

$$\begin{aligned} n\hat{\Sigma}_{[K]\bullet} &= O_{[K]}\mathrm{diag}(\psi_{[K],1}, \psi_{[K],2}, \ldots, \psi_{[K],\|[K]\|})O'_{[K]} \\ \Sigma_{[K\rangle}\widehat{\Sigma^{-1}_{\langle K\rangle}} &= S_{[K\rangle}S^{-1}_{\langle K\rangle}, \quad K \in \mathcal{J}(\mathcal{K}). \end{aligned} \tag{13}$$

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、(13) で与えれる推定量 $\hat{\Sigma}$ の危険関数を分解することにより、例 2 の結果が適用できる．

定理 3 ：$\Sigma \in P_I(\mathcal{K})$ とする．

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

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 4 ：$\max\{\|[K]\| | K \in \mathcal{J}(\mathcal{K})\} \geq 2$ 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 (13) においておけば、$\hat{\Sigma}_{\mathcal{K}}$ 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．

その 1：$\mathcal{K}_1 = \{\phi, 1, 12, 13, I\}$ とする．したがって、$\mathcal{J}(\mathcal{K}_1) = \{1, 12, 13\}$ となる． $X_1, X_2, \ldots, X_n$ を $N(\mathcal{K}_1)$ 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

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

となる．これを分解して、$n\hat{\Sigma}^{\rm mle}_{\mathcal{K}_1} = T_1 T_1'$ と書く．ただし、

$$T_1 = \begin{bmatrix} s_{11}^{1/2} & 0 & 0 \\ s_{21}s_{11}^{-1/2} & t_{22} & 0 \\ s_{31}s_{11}^{-1/2} & 0 & t_{33} \end{bmatrix} \in GT_I(\mathcal{K}_1)$$

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 $\hat{\Sigma}^{\rm m}_{\mathcal{K}_1} = T_1 D_1 T_1'$ 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、

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

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、$\mathcal{J}(\mathcal{K}_2) = \{1, 2, I\}$ となる．同様に $X_1, X_2, \ldots, X_n$ を $N(\mathcal{K}_2)$ からのランダム標本とする．$\Sigma(\in P_I(\mathcal{K}_2))$ の最尤推定量は $n\hat{\Sigma}^{\rm mle}_{\mathcal{K}_2} = T_2 T_2'$ で与えられる．ここで、

$$T_2 = \begin{bmatrix} s_{11}^{1/2} & 0 & 0 \\ 0 & s_{22}^{1/2} & 0 \\ t_{31} & t_{32} & t_{33} \end{bmatrix} \in GT_I(\mathcal{K}_2)$$

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 $\hat{\Sigma}^{\rm m}_{\mathcal{K}_2} = T_2 D_2 T_2'$ 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．

5. むすびにかえて

4 節で取り上げた例以外にも matrix estimation problems に関する研究は多々ある．その中でも特に Stein の着想に沿って改良型の推定量を導出した研究として以下のものがある．多変量正規分布の２標本問題における分散共分散行列の同時推定問題（Loh (1988, 1991a, 1991c) ）、同じ問題を別の方向からの定式化したもの（Muirhead and Verathaworn (1985), Bilodeau and Srivastava (1989)）、多変量正規分布の２標本問題における共通平均の推定問題（Loh (1988, 1991b))、付加情報がある場合の多変量正規分布の共分散行列の推定問題（Konno (1995), Krishinamoorthy (1992))、4 節の例 1 と 2 の結果を多変量 t 分布や楕円型分布のもとに拡張した研究（Joarder and Ali (1997), Kubokawa (1997)）、非心 Wishart 分布と非心多変量 F 分布の非心母数の推定問題（Leung (1994a, 1994b) ）等がある．

最近、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．

ミニマックス推定という立場から matrix estimation problems の取り組む研究は多いが、Stein の方針という手段以外に有効に機能する道具がこの問題ではないことから別の方向からの研究はまだまだ数が少ない．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．

参考文献

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

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.

Bilodeau, M., and Kariya, T. (1989): Minimax estimators in the normal MANOVA model, *J. Multivariate Anal.* 28, 260–270.

Bilodeau, M. and Srivastava, M.S. (1992): Estimation of the eigenvalues of $\Sigma_1\Sigma_2^{-1}$, *J. Multivariate Anal.* 41, 1-13.

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.

Efron, B. and Morris, C. (1976): Families of minimax estimators of the mean of a multivariate normal distribution, *Ann. Statist.* 4, 11–21.

Gleser, L. (1986): Minimax estimators of a normal mean vector of a normal distribution with unknown covariance matrix, *Ann. Statist.* 14, 1625–1633.

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

Haff, L. R. (1979): Identity for the Wishart distribution with applications, *J. Multivariate Anal.* 9, 531–544.

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

Honda, T. (1991): Minimax estimators in the MANOVA model for arbitrary quadratic loss and unknown covariance matrix, *J. Multivariate Anal.* 36, 113–120.

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

Joarder, A.H. and Ali, M.M. (1997): Estimation of the scale matrix of a multivariate T-model under entropy loss. *Metrika* 46, 21–32.

Kariya, T., Konno, Y. and Strawderman, W.E. (1996): Double Shrinkage estimation in the GMANOVA model, *J. Multivariate Anal.* 56, 245–258.

Kariya, Y., Konno, Y. and Strawderman, W.E. (1999): Construction of shrinkage estimators for the regression coefficient matrix in the GMANOVA models, to appear in *Communication in Statistics, Theory and Method.*

Konno, Y. (1991): On estimation of a matrix of normal means with unknown covariance matrix, *J. Multivariate Anal.*, 36, 44–55.

Konno, Y. (1995): Estimation of a normal covariance matrix with incomplete data under Stein's loss, *J. Multivariate Anal.* 52, 308–324.

Konno, Y. (1998): Inaddmissibility of the maximum likelihood estimator of normal covariance matrices with the lattice conditional independence, Technical Report No. 340, Department of Statistics, University of Washington.

Konno, Y., Kubokawa, T. and Saleh, A.K.Md.E. (1997): Shrinkage estimators in a mixed MANOVA and GMANOVA mdoel, *Statistics & Decisions* 15, 37–49.

Krishnamoorthy, K. (1992): Estimation of normal covariance matrix and precision matrices with incomplete data, *Comm. Statist. Theory Methods* 20, 757–770.

久保川達也 (1996): 縮小推定の理論と応用 (1), 経済学論集 61, no. 4, 2–31.

Kubokawa, T., Saleh, A.K.Md.E., and Morita, K. (1992): Improving on MLE of coefficient matrix in a growth curve model, *J. Statist. Plan. and Infer.* 31, 169–177.

Kubokawa, T. and Srivatava, M.S. (1997): Robust improvements in estimation of mean and covariance matrices in ellipitically contoured distribution, 「推定論とその応用の研究」シンポジウム予稿集, 1–10.

Kuriki, S. (1993): Orthogonally invariant estimation of the skew-symmetric normal mean matrix. *Ann. Inst. Statist. Math.* 45, 731–739.

Lehmann, E. L. and Casella, G. (1998): *Theory of Point Estimation*, 2nd edtion. Springer.

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

Leung, P.-L. (1994a): An identity for the noncentral Wishart distribution with application. *J. Multivariate Anal.* 48, 107–114.

Leung, P.-L. (1994b): An identity for the noncentral multivariate F distribution with application. Unpublished manuscript.

Loh, W.-L. (1988): Estimating covariance matrices. Ph.D. thesis. Stanford University.

Loh, W.-L. (1991a): Estimating covariance matrices, *Ann. Statist.* 19, 293–296.

Loh, W.-L. (1991b): Estimating the common mean of two multivariate normal distributions, *Ann. Statist.* 19, 297–323.

Loh, W.-L. (1991c): Estimating covariance matrices II, *J. Multivariate Anal.* 36, 163–174.

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

Muirhead, R.J. (1982): *Aspects of multivariate statistical analysis*, Wiley.

Muirhead, R.J. and Verathaworn, T. (1985): Estimating the latent roots of $\Sigma_1\Sigma_2^{-1}$, In *Multivariate Analysis VI*, (ed. Krishnaiah, P.R.), 431–438.

Perron, F. (1992): Minimax estimators of a covariance matrix, *J. Multivariate Anal.* 43, 6–28.

Robert, C.P. (1994): *The Bayesian Choice.* Springer-Verlag.

Sheena, Y. (1995): Unbiased estimator of risk for an orthogonally invariant estimator of a covariance matrix. *J. Japan Statist. Soc.* 25, 35–48.

Sheena, Y and Takemura, A. (1992): Inadmissibility of non-order-preserving orthogonally invariant estimators of the covariance matrix in the case of Stein's loss, *J. Multivariate Anal.* 41, 117–131.

篠崎信雄 (1991): Stein タイプの縮小推定量とその応用, 応用統計学 20, 59–76.

Stein, C. (1956): Inadmissibility of the usual estimator for the mean of a multivariate distribution. *Proc. Third Berkeley Symp. Math. Statist. Prob.* 1, University of California Press, 197–206.

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

Stein, C. (1975): Estimation of a covariance matrix. Rietz Lecture, 39th Annual Meeting IMS. Atlanta, Georgia.

Takemura, A. (1984): An orthogonally invariant minimax estimator of the covariance matrix of a multivariate normal population. *Tsukuba J. Math.* 8, 367–376.

竹村彰通 (1991): 多変量推測統計の基礎, 共立出版 .

Tan, M. (1991): Improved estimators for the GMANOVA problem with application to Monte Carlo Simulation, *J. Multivariate Anal.* 38, 262–274.

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

Zheng, Z. (1986): On estimation of matrix of normal mean, *J. Multivariate Anal.* 18, 70–82.