論文の内容の要旨

論文題目　Orthogonally equivariant estimation of the variance-covariance matrix of a normal distribution（正規分布の分散・共分散行列の直交共変推定）
氏　名　　椎　名　　洋

本論文では, 正規分布の分散共分散行列, あるいはその逆行列の推定を統計的決定理論の立場から扱う. 推定量には, いくつかのクラスがあるが, 本論文では, 直交行列による通常の変換に関して共変な推定量（以下, 直交共変推定量）に的を絞った研究の成果を述べる. 問題をより具体的に述べれば, 以下のようになる. $\boldsymbol{x}_i,\ i = 1, \ldots, n$ が, お互いに独立に, 同一の p 次元多変量正規分布 $N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ に従うと仮定する. この時, パラメーター $\boldsymbol{\mu}$ と $\boldsymbol{\Sigma}$ が, 共に未知の時は,

$$\boldsymbol{W} = \sum_{i=1}^{n} (\boldsymbol{x}_i - \bar{\boldsymbol{x}})(\boldsymbol{x}_i - \bar{\boldsymbol{x}})',$$

（但し, $\bar{\boldsymbol{x}} = \sum_{i=1}^{n} \boldsymbol{x}_i$）に基づいて, $\boldsymbol{\Sigma}$ 或いは $\boldsymbol{\Sigma}^{-1}$ の推定を行うことが多い. （$\bar{\boldsymbol{x}}$ もさらに利用して推定を行う方法も研究されているが, 本論文では扱っていない. ）また, $\boldsymbol{\mu}$ が 既知、例えば $\boldsymbol{\mu} = 0$ の場合は, 十分統計量

$$\widetilde{\boldsymbol{W}} = \sum_{i=1}^{n} \boldsymbol{x}_i {\boldsymbol{x}_i}'$$

を使って, 推定するのが合理的である. いずれにせよ, 問題は, p 次元ウィシャート分布に従う統計量 $\boldsymbol{W}$ があるとき, つまり

$$\boldsymbol{W} \sim \boldsymbol{W_p}\,(k, \Sigma)$$

の時, この $\boldsymbol{W}$ に基づいて, $\boldsymbol{\Sigma}$ 或いは $\boldsymbol{\Sigma}^{-1}$ をいかに（直交共変）推定するかという問題に帰着する.

直交共変推定量の定義を述べる．まず, $\boldsymbol{W}$ の直交分解

$$\boldsymbol{W} = \boldsymbol{HLH'}$$

を考える．但し, $\boldsymbol{L} = \mathrm{diag}(l_1, \ldots, l_p)$ であり, l_i は $\boldsymbol{W}$ の固有根で以下のように順序付けられているとする．

$$l_1 \geq l_2 \geq \ldots \geq l_p \geq 0. \tag{1}$$

また, $\boldsymbol{H} = (h_{ij})$ は, $p \times p$ 次元の直交行列である．さらに以下の記号を定義しておく．

$$\begin{aligned} \boldsymbol{l} &= (l_1, \ldots, l_p), \\ \mathcal{L} &= \{\boldsymbol{l} \mid 0 < l_p < \ldots < l_1\}, \\ \mathcal{O}(p) &: \ p \times p \text{ 次元直交行列のなす群.} \end{aligned}$$

さて, もし $\boldsymbol{\Sigma}$ の推定量 $\widehat{\boldsymbol{\Sigma}}$ が,

$$\widehat{\boldsymbol{\Sigma}}(\boldsymbol{OWO'}) = \boldsymbol{O}\widehat{\boldsymbol{\Sigma}}(\boldsymbol{W})\boldsymbol{O'}, \quad \forall \boldsymbol{O} \in \mathcal{O}(p),$$

を満たすならば, この推定量 $\widehat{\boldsymbol{\Sigma}}$ を直交共変な（$\boldsymbol{\Sigma}$ の）推定量と呼ぶ．同様に $\boldsymbol{\Sigma}^{-1}$ の推定量 $\widehat{\boldsymbol{\Sigma}^{-1}}$ が, 条件

$$\widehat{\boldsymbol{\Sigma}^{-1}}(\boldsymbol{OWO'}) = \boldsymbol{O}\widehat{\boldsymbol{\Sigma}^{-1}}(\boldsymbol{W})\boldsymbol{O'}, \quad \forall \boldsymbol{O} \in \mathcal{O}(p).$$

を満たすならば, $\widehat{\boldsymbol{\Sigma}^{-1}}$ を直交共変な（$\boldsymbol{\Sigma}^{-1}$ の）推定量と呼ぶ．直交共変な推定量は, 具体的には必ず次のような形をとることが知られている．

$$\widehat{\boldsymbol{\Sigma}}(\widehat{\boldsymbol{\Sigma}^{-1}}) = \boldsymbol{H\Psi H'}, \qquad \boldsymbol{\Psi} = \mathrm{diag}(\psi_1(\boldsymbol{l}), \ldots, \psi_p(\boldsymbol{l})). \tag{2}$$

ここで, $\psi_i(\boldsymbol{l}),\ i = 1, \ldots, p$ は, $\boldsymbol{H}$ に依存しないことに注意．

この論文では, 統計的決定理論の立場から以下のような損失関数を用いて, 推定量の良し悪しを吟味する．まず, $\boldsymbol{\Sigma}$ の推定では,

$$\begin{aligned} L_1(\widehat{\boldsymbol{\Sigma}}, \boldsymbol{\Sigma}) &= \mathrm{tr}(\widehat{\boldsymbol{\Sigma}}\boldsymbol{\Sigma}^{-1}) - \log|\widehat{\boldsymbol{\Sigma}}\boldsymbol{\Sigma}^{-1}| - p, \\ L_2(\widehat{\boldsymbol{\Sigma}}, \boldsymbol{\Sigma}) &= \mathrm{tr}(\widehat{\boldsymbol{\Sigma}}\boldsymbol{\Sigma}^{-1} - \boldsymbol{I}_p)^2 \end{aligned}$$

の 2 つの損失関数が, 研究の対象となることが多い．前者は, スタインの損失関数, あるいは尤度損失関数と呼ばれ, 後者は, 2 乗損失関数と呼ばれる．一方, $\boldsymbol{\Sigma}^{-1}$ の推定では,

$$\begin{aligned} L_3(\widehat{\boldsymbol{\Sigma}^{-1}}, \boldsymbol{\Sigma}^{-1}) &= \mathrm{tr}(\widehat{\boldsymbol{\Sigma}^{-1}}\boldsymbol{\Sigma}) - \log|\widehat{\boldsymbol{\Sigma}^{-1}}\boldsymbol{\Sigma}| - p, \\ L_4(\widehat{\boldsymbol{\Sigma}^{-1}}, \boldsymbol{\Sigma}^{-1}) &= \mathrm{tr}(\widehat{\boldsymbol{\Sigma}^{-1}}\boldsymbol{\Sigma} - \boldsymbol{I}_p)^2 \end{aligned}$$

の２つの損失関数が同様に, 考察の対象となることが多い. 最終的には, 推定量の優劣はこれらの損失関数に関するリスク $R_i(\widehat{\mathbf{\Sigma}}, \mathbf{\Sigma}),\ i=1,2$ や $R_i(\widehat{\mathbf{\Sigma}^{-1}}, \mathbf{\Sigma}^{-1}),\ i=3,4$ によって評価される. すなはち,

$$\begin{aligned} R_i(\widehat{\mathbf{\Sigma}}, \mathbf{\Sigma}) &= E[L_i(\widehat{\mathbf{\Sigma}}, \mathbf{\Sigma})], \quad i=1,2, \\ R_i(\widehat{\mathbf{\Sigma}^{-1}}, \mathbf{\Sigma}^{-1}) &= E[L_i(\widehat{\mathbf{\Sigma}^{-1}}, \mathbf{\Sigma}^{-1})], \quad i=3,4. \end{aligned}$$

とした時に, もし二つの推定量 $\widehat{\mathbf{\Sigma}}^{(1)},\ \widehat{\mathbf{\Sigma}}^{(2)}$ に,

$$\begin{aligned} R_i(\widehat{\mathbf{\Sigma}}^{(1)}, \mathbf{\Sigma}) &\le R_i(\widehat{\mathbf{\Sigma}}^{(2)}, \mathbf{\Sigma}), \quad \forall \mathbf{\Sigma}, \\ R_i(\widehat{\mathbf{\Sigma}}^{(1)}, \mathbf{\Sigma}) &< R_i(\widehat{\mathbf{\Sigma}}^{(2)}, \mathbf{\Sigma}), \quad \exists \mathbf{\Sigma}, \end{aligned}$$

という関係が成り立っているとすると, 明らかに $\widehat{\mathbf{\Sigma}}^{(1)}$ は, $\widehat{\mathbf{\Sigma}}^{(2)}$ より優れていることになる. これを統計的決定理論では, $\widehat{\mathbf{\Sigma}}^{(1)}$ は, $\widehat{\mathbf{\Sigma}}^{(2)}$ を, $L_i,\ i=1,2$ に関して「優越する」という. 同様にして, $\mathbf{\Sigma}^{-1}$ の推定においても, $R_i(\widehat{\mathbf{\Sigma}^{-1}}^{(1)}, \mathbf{\Sigma}^{-1})$ と $R_i(\widehat{\mathbf{\Sigma}^{-1}}^{(2)}, \mathbf{\Sigma}^{-1})$ に同様の関係が成り立てば, 前者は, 後者を $L_i,\ i=3,4$ に関して「優越する」という. 他の推定量に「優越されて」しまう推定量を「非許容的」という.

以上のような枠組みの中で, 直交共変な推定量がどんな性質をもつかについての研究結果が各章にまとめられている. 以下, 簡単に各章の内容を紹介する.

第２章では, 直交共変推定量の「順序保持性」を扱っている. (2) で与えられる $\mathbf{\Sigma}$ の直交共変推定量 $\widehat{\mathbf{\Sigma}}$ が, 次の性質を満たすとき,「順序を保持する」と呼ぶことにする.

$$\psi_1(\boldsymbol{l}) \ge \psi_2(\boldsymbol{l}) \ge \ldots \ge \psi_p(\boldsymbol{l}), \quad \forall \boldsymbol{l} \in \mathcal{L}. \tag{3}$$

この順序は (1) と同じであり, この順序が自然なものであることを L_1 に関して示したのがこの章の成果である. すなはち, 順序に関する条件 (3) を満たさない推定量は, 全てその推定量を改良してできた順序を保持する推定量によって優越されることを証明する. 改良の方法として, 順序統計量を使う方法と Isotonic Regression を使う方法の２つを示す.

第３章では, 前章と同じ問題を, $\mathbf{\Sigma}^{-1}$ の推定について, 考える. この場合 (2) で与えられる直交共変推定量 $\widehat{\mathbf{\Sigma}^{-1}}$ の自然な順序は

$$\psi_1(\boldsymbol{l}) \le \psi_2(\boldsymbol{l}) \le \ldots \le \psi_p(\boldsymbol{l}), \quad \forall \boldsymbol{l} \in \mathcal{L}$$

である. この順序を保持しない推定量について, 前章と同じ結果が成り立つことを $p=2,3$ の場合について証明する. 一般次元に関する結果はまだ得られていない.

第４章では, $\mathbf{\Sigma}$ の推定に関して, 直交共変な推定量のリスクの不偏推定量を, 新しいやり方で導出する. 直交共変な推定量に限らず, 一般的な推定量に関して, $R_i(\widehat{\mathbf{\Sigma}}, \mathbf{\Sigma}),\ i=1,2$ の不偏推定量, すなはち

$$R_i(\widehat{\mathbf{\Sigma}}, \mathbf{\Sigma}) = E_{\Sigma}[\hat{R}_i(\widehat{\mathbf{\Sigma}}, \boldsymbol{W})], \quad \forall \mathbf{\Sigma},$$

を満たす $\hat{R}_i(\widehat{\boldsymbol{\Sigma}}, \boldsymbol{W})$ を, $L_i,\ i=1,2$ に関するリスクの不偏推定量と呼ぶ．通常これらは, 部分積分法, あるいはストークスの定理を用いて導出されるが, 直交共変な推定量に関しては, より直裁な方法で導出できることを示したのが, この章の成果である．技術的な中身の章だが, 次章の成果はこの章の技術を用いて導かれている．

ここで, 第５章以下の中身に深く関わる直交共変推定量を紹介する．(2) と

$$\psi_i(\boldsymbol{l}) = l_i \delta_i, \quad i=1,\ldots,p$$

で定義される直交共変推定量で, 特に定数 δ_i が, $L_i,\ i=1,\ldots,4$ に関して最良な三角不偏推定量に使われる定数に一致している推定量を, それぞれ $\widehat{\boldsymbol{\Sigma}}_{oi},\ i=1,2,\ \widehat{\boldsymbol{\Sigma}_{oi}^{-1}},\ i=3,4$ で表す．これらに関しては, その基になっている最良三角不偏推定量を優越し, したがってミニマックスであるという予想がなされている．$\widehat{\boldsymbol{\Sigma}}_{o1}$ に関しては一般次元で, また $\widehat{\boldsymbol{\Sigma}_{o3}^{-1}}$ については, $p=2$ の場合にこの予想は証明されているが, そのほかの場合については, まだ証明がなされていない．第５章では, この予想が正しいことを, $\widehat{\boldsymbol{\Sigma}}_{o2}$ に関して $p=2$ の場合に, $\widehat{\boldsymbol{\Sigma}_{o3}^{-1}}$ に関して $p=3$ の場合に証明する．

第６章は, $\boldsymbol{\Sigma}$ の固有根が無限に拡散した場合の漸近論を扱っている．$\boldsymbol{\Sigma}$ の固有根を

$$\lambda_1 \geq \ldots \geq \lambda_p > 0$$

とした時, これらが無限に発散するとは,

$$\rho(\boldsymbol{\Sigma}) = \max\Big(\frac{\lambda_2}{\lambda_1}, \frac{\lambda_3}{\lambda_2}, \ldots, \frac{\lambda_p}{\lambda_{p-1}}\Big) \to 0$$

を意味する. 章の前半で基準化された $\boldsymbol{l}$ と $\boldsymbol{H}$ の漸近分布が導出されるが, その結果を一般の（直交共変とは限らない）推定量に当てはめると, $L_i,\ i=1,2$ に関して, 母集団固有根が無限に拡散した場合の, 唯一の合理的な推定量が $\widehat{\boldsymbol{\Sigma}}_{oi}$ であることが証明される．より正確に言えば, 次のようになる．母集団固有根が無限に拡散した状態の近傍を適当にとれば, その近傍で最良三角不偏推定量の（定数）リスクを常に下回る推定量を, Tail Minimax な推定量と呼ぶ．Tail Mimimax であるためには, その推定量が関数として依存している標本固有根を無限に拡散した時に, $\widehat{\boldsymbol{\Sigma}}_{oi},\ i=1,2$ に収束しなければならないという結果が得られる．すなはち, Tail Minimax であるための（第一の）必要条件が得られる．

第７章では, 第６章の漸近論をさらに高次で考える. 母集団の固有根が, 直線的に拡散する, すなはち,

$$\frac{\lambda_{j+1}}{\lambda_j} = a_j z, \quad 1 \leq j \leq p-1$$

として, $a_j,\ j=1,\ldots,p-1$ を固定し, z を上から 0 に限りなく近づける場合を考える．この時, 基準化された $\boldsymbol{l}$ と $\boldsymbol{H}$ の漸近分布と R_1 について, それぞれ z の第一次項まで計算する．この結果を利用することで, Tail Minimax であるための,（第二の）必要条件が得られる．この必要条件を当てはめることで, これまでミニマックスであるかどうかの証明がなされていない２つの推定量が, ミニマックスでないことを証明する．