

楕円型分布モデルの下での2標本問題における
共通平均の推定について

千葉大学大学院 自然科学研究科
津熊 久幸

千葉大学大学院 自然科学研究科
今野 良彦

内容

1. 先行研究について
2. モデルと問題
3. 正準形と推定量の族
4. Stein型推定量
5. 数値実験
6. 結論

先行研究 Loh, 1991

$$X_1 \sim \mathcal{N}_p(\theta, \Sigma_1), \quad X_2 \sim \mathcal{N}_p(\theta, \Sigma_2)$$
$$S_1 \sim \mathcal{W}_p(\Sigma_1, n), \quad S_2 \sim \mathcal{W}_p(\Sigma_2, n)$$

ただし, Σ_i は未知. 上記モデル (正準形) の共通平均 θ の推定問題を次の損失関数の下で議論 :

$$L(\hat{\theta}; \theta) = (\hat{\theta} - \theta)'(\Sigma_1^{-1} + \Sigma_2^{-1})(\hat{\theta} - \theta)$$

↓

$P : p \times p$ 正則行列, $d : p \times 1$ ベクトルについて, 変換群

$$\begin{cases} \theta \rightarrow P\theta + d \\ \Sigma_i \rightarrow P\Sigma_i P' \end{cases} \quad \begin{cases} X_i \rightarrow PX_i + d \\ S_i \rightarrow PS_i P' \end{cases} \quad (i = 1, 2)$$

に関する共変推定量が, 次で与えられることを示す:

(共変推定量)

$$\hat{\theta} = B^{-1}\Psi BX_1 + B^{-1}(I_p - \Psi)BX_2$$

● B : 正則行列, F : 対角行列

$$\begin{cases} B(S_1 + S_2)B' = I_p \\ BS_2B' = F = \text{diag}(f_1, \dots, f_p) \\ f_1 \geq \dots \geq f_p > 0 \end{cases}$$

● $\Psi = \text{diag}(\psi_1, \dots, \psi_p)$, ψ_i は F と $B(X_1 - X_2)$ の各成分の絶対値の関数

(先行研究の続き) Loh, 1991

共変推定量を

$$\begin{aligned}\hat{\theta} &= B^{-1}\Phi BX_1 + B^{-1}(I_p - \Phi)BX_2 \\ \Phi &\equiv \Phi(F) = \text{diag}(\phi_1, \dots, \phi_p)\end{aligned}$$

に制限 (結合比は, S_1 と S_2 にのみ依存)

↓

上記推定量のリスクの不偏推定量を計算

↓

リスクの不偏推定量から, **Graybill-Deal**型結合推定量

$$\begin{aligned}\hat{\theta}^{GD} &= \{(S_1/n)^{-1} + (S_2/n)^{-1}\}^{-1} \\ &\quad \times \{(S_1/n)^{-1}X_1 + (S_2/n)^{-1}X_2\} \\ &= (S_1^{-1} + S_2^{-1})^{-1}(S_1^{-1}X_1 + S_2^{-1}X_2)\end{aligned}$$

を改良する代替推定量の候補を導出

↓

数値実験で, リスク値の比較検討を行い, その代替推定量の有効性を示す

モデル

$$\begin{matrix} Y_1 & = & \mathbf{1}_N & \xi' & + & \epsilon_1 & , & Y_2 & = & \mathbf{1}_N & \xi' & + & \epsilon_2 \\ N \times p & & N \times 1 & 1 \times p & & N \times p & & N \times p & & N \times 1 & 1 \times p & & N \times p \end{matrix}$$

- $Y_1, Y_2, \epsilon_1, \epsilon_2$ は $N \times p$ 確率行列 ($N > p$)
- $\mathbf{1}_N$ はすべての成分が 1 である $N \times 1$ ベクトル
- ξ は未知の $p \times 1$ ベクトル
- $\epsilon_i = (e_{i1}, e_{i2}, \dots, e_{iN})'$ ($i = 1, 2$) の各行は独立に同一の楕円型分布に従うとし, その密度関数を

$$|\Sigma_i|^{-1/2} h(e_{ij}' \Sigma_i^{-1} e_{ij})$$
$$i = 1, 2, j = 1, \dots, N$$

とする. ただし, $h(\cdot)$ は $[0, \infty)$ 上の未知関数,
 Σ_i は未知の $p \times p$ 正定値対称行列

(問題)

共通平均 ξ を損失関数

$$\tilde{L}(\hat{\xi}; \xi) = N(\hat{\xi} - \xi)'(\Sigma_1^{-1} + \Sigma_2^{-1})(\hat{\xi} - \xi)$$

の下で推定 ($\hat{\xi}$ は ξ の推定量)

- Graybill-Deal 型結合推定量の改良を目標とする

(本報告におけるアイデア)

密度関数

$$\begin{aligned} & \prod_{i,j} |\Sigma_i|^{-1/2} h(e'_{ij} \Sigma_i^{-1} e_{ij}) \\ & = |\Sigma_1|^{-N/2} |\Sigma_2|^{-N/2} \prod_{i,j} h(e'_{ij} \Sigma_i^{-1} e_{ij}) \quad (\text{iid}) \end{aligned}$$

の下ではリスクの評価式の導出が困難であるため、次の密度に置き換え、その下で推定量の候補を導出する：

$$\begin{aligned} & |\Sigma_1|^{-N/2} |\Sigma_2|^{-N/2} g\left(\sum_{i,j} e'_{ij} \Sigma_i^{-1} e_{ij}\right) \\ & = |\Sigma_1|^{-N/2} |\Sigma_2|^{-N/2} g\left(\text{tr}\{\Sigma_1^{-1} \epsilon'_1 \epsilon_1 + \Sigma_2^{-1} \epsilon'_2 \epsilon_2\}\right) \\ & \hspace{15em} (\text{non-iid}) \end{aligned}$$

- $g(\cdot)$: $[0, \infty)$ 上の未知関数
- 一般に、 ϵ_i の各行は独立ではないが無相関
 - 正規分布の場合は独立 : (iid)=(non-iid)

↓

上記の方法で求められた Graybill-Deal 型結合推定量の代替推定量を、(iid) の下での数値実験を行い、リスクの挙動を調べる

正準形

記法 $i = 1, 2$ に対し

- $\Gamma_i : \Gamma_i \mathbf{1}_N = (\sqrt{N}, 0, \dots, 0)'$ を満たす $N \times N$ 直交行列
- $\mathbf{X}_i, \mathbf{Z}_i : (\mathbf{X}_i, \mathbf{Z}_i)' = \Gamma_i \mathbf{Y}_i \quad (n = N - 1)$
 $p \times 1 \quad n \times p$
- $\boldsymbol{\theta} = \sqrt{N} \boldsymbol{\xi}$

注意

- $\mathbf{X}_i = \mathbf{Y}_i' \mathbf{1}_N / \sqrt{N}$
- $\mathbf{Z}_i' \mathbf{Z}_i = \mathbf{Y}_i' (\mathbf{I}_N - \mathbf{1}_N \mathbf{1}_N' / N) \mathbf{Y}_i \equiv \mathbf{S}_i$
 $p \times p$

Γ_1 と Γ_2 による直交変換より, (non-iid) の正準形を得る:

(正準形)

$$|\Sigma_1|^{-N/2} |\Sigma_2|^{-N/2} g \left(\sum_{i=1}^2 [\text{tr} \{ \Sigma_i^{-1} \mathbf{Z}_i' \mathbf{Z}_i + \Sigma_i^{-1} (\mathbf{X}_i - \boldsymbol{\theta})(\mathbf{X}_i - \boldsymbol{\theta})' \}] \right)$$

- 共通平均 $\boldsymbol{\theta}$ の推定問題を損失関数

$$L(\hat{\boldsymbol{\theta}}; \boldsymbol{\theta}) = (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})' (\Sigma_1^{-1} + \Sigma_2^{-1}) (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})$$

の下で考察 ($\hat{\boldsymbol{\theta}} = \sqrt{N} \hat{\boldsymbol{\xi}}$)

推定量の族

Graybill-Deal型結合推定量：

$$\hat{\theta}^{GD} = (S_1^{-1} + S_2^{-1})^{-1}(S_1^{-1}X_1 + S_2^{-1}X_2)$$

↓ (一般化)

(推定量の族)

$$\hat{\theta} = B^{-1}\Phi BX_1 + B^{-1}(I_p - \Phi)BX_2$$

● B ：正則行列, F ：対角行列

$$\begin{cases} B(S_1 + S_2)B' = I_p \\ BS_2B' = F = \text{diag}(f_1, \dots, f_p) \\ f_1 \geq \dots \geq f_p > 0 \end{cases}$$

● $\Phi \equiv \Phi(F)$ ：対角行列

注意 Graybill-Deal型結合推定量：

$$\begin{cases} B(S_1 + S_2)B' = I_p \\ BS_2B' = F \end{cases} \Rightarrow \begin{cases} S_1^{-1} = B'(I_p - F)^{-1}B \\ S_2^{-1} = B'F^{-1}B \end{cases}$$

$$\begin{aligned} \hat{\theta}^{GD} &= (S_1^{-1} + S_2^{-1})^{-1}S_1^{-1}X_1 \\ &\quad + (S_1^{-1} + S_2^{-1})^{-1}S_2^{-1}X_2 \\ &= B^{-1}\Phi^{GD}BX_1 + B^{-1}(I_p - \Phi^{GD})BX_2, \end{aligned}$$

ただし, Φ^{GD} の第 j 成分は

$$\hat{\phi}_j^{GD} = \frac{1/(1 - f_j)}{1/(1 - f_j) + 1/f_j} = f_j$$

リスク

- $U \equiv U(X_1, X_2, Z_1, Z_2)$

- $G(x) = (1/2) \int_x^{+\infty} g(t) dt$

- $\mathbb{E}_G[U] = \int U \times \left(\prod_{i=1}^2 |\Sigma_i|^{-N/2} \right) \times G(C) dX_1 dX_2 dZ_1 dZ_2$

$$C = \sum_{i=1}^2 [\text{tr} \{ \Sigma_i^{-1} (X_i - \theta)(X_i - \theta)' + \Sigma_i^{-1} Z_i' Z_i \}]$$

上記の記法と Kubokawa & Srivastava (1999, 2001) による楕円型分布モデルの下での Stein-Haff identity を用いることにより、リスクはある正則条件のもとで、次のように表すことができる：

Theorem 1 (リスク)

$$\begin{aligned}
 R(\hat{\theta}; \theta) &= \mathbb{E}_G \left[\sum_{j=1}^p [B(X_1 - X_2)]_j^2 \right. \\
 &\quad \times \left\{ \frac{n-p-1}{f_j} \phi_j^2 + 4 \frac{1-f_j}{f_j} \phi_j^2 \right. \\
 &\quad \left. + 4(1-f_j) f_j \phi_j \frac{\partial}{\partial f_j} \left(\frac{\phi_j}{f_j} \right) \right. \\
 &\quad \left. + \sum_{k \neq j} \phi_j (\phi_j - \phi_k) \frac{1-f_k}{f_j - f_k} \right. \\
 &\quad \left. + \frac{n-p-1}{1-f_j} (1-\phi_j)^2 + 4 \frac{f_j}{1-f_j} (1-\phi_j)^2 \right. \\
 &\quad \left. + 4 f_j (1-f_j) (1-\phi_j) \frac{\partial}{\partial (1-f_j)} \left(\frac{1-\phi_j}{1-f_j} \right) \right. \\
 &\quad \left. + \sum_{k \neq j} (1-\phi_j) (\phi_j - \phi_k) \frac{f_k}{f_j - f_k} \right\} \\
 &\equiv \mathbb{E}_G [\hat{R}]
 \end{aligned}$$

ただし, $[B(X_1 - X_2)]_j$ は $B(X_1 - X_2)$ の第 j 成分

証明 省略

注意 一般に, 楕円型分布の下では, \hat{R} はリスクの不偏推定量ではない

Stein型推定量

直接, \hat{R} から Graybill-Deal 型結合推定量を改良する Φ の候補を, 解析的に求めるのは困難

↓

\hat{R} の微分を含む項を無視し, $\partial \hat{R} / \partial \phi_j = 0$ を計算. 次に ϕ_k を含む項を無視することにより, 次のような Stein 型推定量を得る:

$$\hat{\theta}^{ST} = B^{-1} \Phi^{ST} B X_1 + B^{-1} (I_p - \Phi^{ST}) B X_2$$

$$\Phi^{ST} = \text{diag} (\hat{\phi}_1^{ST}, \dots, \hat{\phi}_p^{ST})$$

$$\hat{\phi}_j^{ST} = \frac{\hat{\beta}_j^{ST} / (1 - f_j)}{\hat{\beta}_j^{ST} / (1 - f_j) + \hat{\alpha}_j^{ST} / f_j} \quad (j = 1, \dots, p)$$

$$\hat{\alpha}_j^{ST} = n - p - 1 + 4(1 - f_j) + 2 \sum_{k \neq j} \frac{f_j(1 - f_k)}{f_j - f_k}$$

$$\hat{\beta}_j^{ST} = n - p - 1 + 4f_j - 2 \sum_{k \neq j} \frac{(1 - f_j)f_k}{f_j - f_k}$$

注意 Graybill-Deal 型結合推定量 :

$$\hat{\theta}^{GD} = B^{-1} \Phi^{GD} B X_1 + B^{-1} (I_p - \Phi^{GD}) B X_2$$

$$\Phi^{GD} = \text{diag} (\hat{\phi}_1^{GD}, \dots, \hat{\phi}_p^{GD})$$

$$\hat{\phi}_j^{GD} = \frac{1/(1 - f_j)}{1/(1 - f_j) + 1/f_j} = f_j$$

数値実験

- 密度関数 (iid) の下で, Graybill-Deal 型結合推定量 $\hat{\theta}^{GD}$, Stein 型推定量 $\hat{\theta}^{ST}$ のリスクの挙動を調べる

仮定 : 誤差項の各行の分布

1. 多変量 t 分布 :

$$\kappa_1 |\Sigma_i|^{-1/2} (1 + e'_{ij} \Sigma_i^{-1} e_{ij} / v)^{-(v+p)/2}$$

$$v > 0, \kappa_1 = \Gamma[(v+p)/2] / \{(\pi v)^{p/2} \Gamma[v/2]\}$$

2. 多変量 Kotz 型分布

$$\begin{aligned} \kappa_2 |\Sigma_i|^{-1/2} \{e'_{ij} \Sigma_i^{-1} e_{ij}\}^{u-1} \\ \times \exp[-r \{e'_{ij} \Sigma_i^{-1} e_{ij}\}^s] \end{aligned}$$

$$r > 0, s > 0, 2u + p > 2,$$

$$\kappa_2 = \frac{s \Gamma[p/2] r^{\{u+p/2-1\}/s}}{\pi^{p/2} \Gamma[\{u+p/2-1\}/s]}$$

実験の設定

- $\xi = 0_p$
- $\Sigma_2 \Sigma_1^{-1}$: 対角行列 (固有根を考慮 ; 数種類)
- $(N, p) = (8, 5), (13, 10)$
- $v = 3 \dots$ (多変量 t 分布)
 $(u, r, s) = (5, 0.5, 2) \dots$ (多変量 Kotz 型分布)

Stein 型結合推定量

Φ^{ST} の成分を “isotonic regression” を用いて順序付けた

ML : 密度 (non-iid) の下で, Σ_1 と Σ_2 が既知の場合の最尤推定量

$$\hat{\theta}^{ML} = (\Sigma_1^{-1} + \Sigma_2^{-1})^{-1} (\Sigma_1^{-1} X_1 + \Sigma_2^{-1} X_2)$$

改良率 : GD に対する ST の改良率

$$AV = 100 \times (1 - \hat{R}^{*ST} / \hat{R}^{*GD})$$

$\hat{R}^{*ST}, \hat{R}^{*GD}$ は, それぞれ ST, GD のリスクの推定値

表の括弧内の値 : 標準誤差の推定値

表 1 : 多変量 t 分布の下でのリスクの推定値 (括弧内は標準誤差の推定値)

ML : 最尤推定量, GD : Graybill-Deal 型推定量
 ST : Stein 型推定量, AV : GD に対する ST の改良率

$\Sigma_2 \Sigma_1^{-1}$ の固有値	ML	GD	ST	AV
$N = 8, \quad p = 5$				
(1, 1, 1, 1, 1)	14.186 (0.504)	26.927 (1.272)	24.271 (1.423)	9.86 %
(10, 0.1, 0.1, 0.1, 0.1)	14.250 (0.551)	32.441 (2.726)	28.716 (2.292)	11.48 %
(16, 8, 4, 2, 1)	14.072 (0.475)	32.762 (1.664)	28.319 (1.450)	13.56 %
($10^{10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}$)	15.651 (1.094)	29.349 (2.148)	29.349 (2.148)	0.00 %
($10^8, 10^4, 1, 10^{-4}, 10^{-8}$)	14.912 (0.542)	29.434 (0.992)	29.784 (1.017)	-1.19 %
$N = 13, \quad p = 10$				
(1, 1, 1, 1, 1, 1, 1, 1, 1, 1)	30.478 (0.992)	62.767 (2.137)	50.855 (1.713)	18.98 %
(10, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1)	27.553 (0.970)	84.003 (2.670)	60.471 (1.803)	28.01 %
(512, 256, 128, 64, 32, 16, 8, 4, 2, 1)	28.641 (1.594)	98.799 (4.157)	67.775 (2.573)	31.40 %
($10^{10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}$)	28.401 (0.863)	47.988 (1.340)	47.988 (1.340)	0.00 %
($10^5, 10^4, 10^3, 10^2, 10, 1, 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}$)	28.307 (0.907)	63.451 (2.292)	63.348 (2.267)	0.16 %

表 2 : 多変量 Kotz 型分布の下でのリスクの推定値
(括弧内は標準誤差の推定値)

ML : 最尤推定量, GD : Graybill-Deal 型推定量
ST : Stein 型推定量, AV : GD に対する ST の改良率

$\Sigma_2 \Sigma_1^{-1}$ の固有値	ML	GD	ST	AV
$N = 8, p = 5$				
(1, 1, 1, 1, 1)	2.465 (0.012)	3.755 (0.021)	3.143 (0.017)	16.32 %
(10, 0.1, 0.1, 0.1, 0.1)	2.420 (0.010)	3.970 (0.027)	3.572 (0.023)	10.01 %
(16, 8, 4, 2, 1)	2.456 (0.011)	4.055 (0.027)	3.386 (0.020)	16.51 %
($10^{10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}$)	2.452 (0.010)	3.509 (0.029)	3.509 (0.029)	0.00 %
($10^8, 10^4, 1, 10^{-4}, 10^{-8}$)	2.435 (0.011)	3.846 (0.024)	3.878 (0.024)	-0.83 %
$N = 13, p = 10$				
(1, 1, 1, 1, 1, 1, 1, 1, 1, 1)	2.922 (0.010)	5.027 (0.022)	3.749 (0.015)	25.43 %
(10, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1)	2.927 (0.009)	6.151 (0.038)	4.484 (0.024)	27.11 %
(512, 256, 128, 64, 32, 16, 8, 4, 2, 1)	2.913 (0.009)	6.449 (0.048)	4.409 (0.028)	31.63 %
($10^{10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}$)	2.928 (0.008)	4.460 (0.064)	4.460 (0.064)	0.00 %
($10^5, 10^4, 10^3, 10^2, 10, 1, 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}$)	2.900 (0.010)	5.107 (0.024)	5.107 (0.024)	0.00 %

実験結果のまとめ

1. $\Sigma_2 \Sigma_1^{-1}$ の各固有値が近い値をとる場合は AV は大きくなる傾向にある. $(N, p) = (13, 10)$ のときは最大 30% の改良が見込める
2. 逆に, $\Sigma_2 \Sigma_1^{-1}$ の各固有値が散らばっている場合には, AV は小さくなる. また, その固有値が極端に散らばっている場合, AV が負になるときもあるが ($p = 5$), その大きさは -1% 程度である
3. $\Sigma_2 \Sigma_1^{-1}$ の各固有値が等しい場合に $N - p$ を固定したもとでは, p が増えるごとに AV も増え, 改良率が上がる

結論

- 推定量 ST は, 密度関数 (non-iid) のもとで導出したが, これら実験結果から密度関数 (iid) の下でも, ST による GD の改良は有効であると思われる

Stein's isotonic regression

$\{f_j/\hat{\alpha}_j^{ST}\}, \{(1-f_j)/\hat{\beta}_j^{ST}\}$ は順序が付いていたほうが望ましい \Rightarrow 以下のような修正をおこなう

$$\begin{array}{cc} f_1 & \hat{\alpha}_1^{ST} \\ f_2 & \hat{\alpha}_2^{ST} \\ \vdots & \vdots \\ f_p & \hat{\alpha}_p^{ST} \end{array}$$

(アルゴリズム : Lin & Perlman, 1985)

Step 1. すべての $\hat{\alpha}_j^{ST}$ を正值にするための Step

- (a) 下から負である $\hat{\alpha}_j^{ST}$ を探す. そのペアを $(f_j, \hat{\alpha}_j^{ST})$ とする
- (b) (a) のペア $(f_j, \hat{\alpha}_j^{ST})$ とこの上のペアを, $(f_j + f_{j-1}, \hat{\alpha}_j^{ST} + \hat{\alpha}_{j-1}^{ST})$ で置き換える
- (c) $\hat{\alpha}_j^{ST}$ がすべて正になるまで (a) と (b) を繰り返す

Step 2. 順序付け Step : まず, 比 $\{f_j/\hat{\alpha}_j^{ST}\}$ を計算

- (a) 下から順序の崩れているところを探す. そのペアを $(f_j, \hat{\alpha}_j^{ST})$ とする
- (b) (a) のペア $(f_j, \hat{\alpha}_j^{ST})$ とこの下のペアに関する比を $(f_j + f_{j+1})/(\hat{\alpha}_j^{ST} + \hat{\alpha}_{j+1}^{ST})$ で置き換える
- (c) (b) で置き換えた比の下のところから再度順序の崩れているところを探し, (b) を施す
- (d) $\{f_j/\hat{\alpha}_j^{ST}\}$ が正しく順序付けられるまで (c) を繰り返す