

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

津熊 久幸 (千葉大学大学院 自然科学研究科)

今野 良彦 (千葉大学大学院 自然科学研究科)

1 はじめに

Loh (1991) では, 2つの多変量正規分布の共通の平均ベクトルの推定問題が議論されている. Loh (1991) は Graybill-Deal 型結合推定量を改良する Stein 型の結合推定量の導出を試み, 数値実験により 2乗損失の下で Stein 型結合推定量が Graybill-Deal 型結合推定量を改良することを示している. 本報告では, 2つの多変量楕円型分布の共通平均の推定問題を扱い, Loh (1991) の結果の拡張をおこなう.

いま, モデルを

$$Y_1 = \mathbf{1}_N \xi' + \epsilon_1, \quad Y_2 = \mathbf{1}_N \xi' + \epsilon_2 \quad (1)$$

とする. ただし, $Y_1, Y_2, \epsilon_1, \epsilon_2$ は $N \times p$ 確率行列, $\mathbf{1}_N$ はすべての成分が 1 である $N \times 1$ ベクトル, ξ は未知の $p \times 1$ ベクトルである. ここで, 確率行列 $\epsilon_i = (e_{i1}, e_{i2}, \dots, e_{iN})'$ ($i = 1, 2$) 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

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

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

ここで, 問題はモデル (1) の未知パラメータ ξ を損失関数

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

のもとで推定することであり, 具体的には Graybill-Deal 型結合推定量の代替推定量を求めることである. その方法としては, まずリスクの不偏推定量に似たりスクの評価式を求め, その評価式から代替推定量の候補を導出する, という手順でおこなう. しかし, 密度関数 (2) の下ではリスクの評価式の導出が困難であるため, 密度関数 (2) を

$$|\Sigma_1|^{-N/2} |\Sigma_2|^{-N/2} g\left(\sum_{i,j} e_{ij}' \Sigma_i^{-1} e_{ij}\right) = |\Sigma_1|^{-N/2} |\Sigma_2|^{-N/2} g(\text{tr}\{\Sigma_1^{-1} \epsilon_1' \epsilon_1 + \Sigma_2^{-1} \epsilon_2' \epsilon_2\}) \quad (3)$$

におきかえて考えていくことにする. ここで, g は $[0, \infty)$ 上の未知関数である.

2 正準形と推定量の族

$i = 1, 2$ に対し, Γ_i を $\Gamma_i \mathbf{1}_N = (\sqrt{N}, 0, \dots, 0)'$ を満たす $N \times N$ 直交行列とする. また, X_i と Z_i をそれぞれ, $p \times 1$ ベクトル, $n \times p$ 行列とし, $(X_i, Z_i)' = \Gamma_i Y_i$ とする ($n = N - 1$).

Γ_1 と Γ_2 による直交変換より, (3) の正準形を得る:

$$|\Sigma_1|^{-N/2} |\Sigma_2|^{-N/2} g \left(\sum_{i=1}^2 [\text{tr} \{ \Sigma_i^{-1} (\mathbf{X}_i - \boldsymbol{\theta})(\mathbf{X}_i - \boldsymbol{\theta})' + \Sigma_i^{-1} \mathbf{Z}_i' \mathbf{Z}_i \}] \right). \quad (4)$$

ただし, $\boldsymbol{\theta} = \sqrt{N} \boldsymbol{\xi}$ である. したがって, (3) の $\boldsymbol{\xi}$ の推定問題は (4) の $\boldsymbol{\theta}$ の推定問題におきかえることができ, $\boldsymbol{\theta}$ の推定問題は損失関数

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

のもとで考えていくことにする.

上記の記法を用いると, Graybill-Deal 型結合推定量は

$$\widehat{\boldsymbol{\theta}}^{GD} = (\mathbf{S}_1^{-1} + \mathbf{S}_2^{-1})^{-1} (\mathbf{S}_1^{-1} \mathbf{X}_1 + \mathbf{S}_2^{-1} \mathbf{X}_2)$$

と表すことができる. ただし, $\mathbf{S}_i = \mathbf{Z}_i' \mathbf{Z}_i$ ($i = 1, 2$) である. この推定量を改良するために, $\widehat{\boldsymbol{\theta}}^{GD}$ を一般化した推定量の族

$$\widehat{\boldsymbol{\theta}} = \mathbf{B}^{-1} \Phi \mathbf{B} \mathbf{X}_1 + \mathbf{B}^{-1} (\mathbf{I}_p - \Phi) \mathbf{B} \mathbf{X}_2 \quad (5)$$

を考える. ただし, $\mathbf{B}(\mathbf{S}_1 + \mathbf{S}_2)\mathbf{B}' = \mathbf{I}_p$, $\mathbf{B}\mathbf{S}_2\mathbf{B}' = \mathbf{F} = \text{diag}(f_1, \dots, f_p)$ ($f_1 \geq \dots \geq f_p$), Φ は各成分が \mathbf{F} の関数である対角行列である.

ここで, Graybill-Deal 型結合推定量は, 次のように表せることに注意しておく:

$$\widehat{\boldsymbol{\theta}}^{GD} = \mathbf{B}^{-1} \Phi^{GD} \mathbf{B} \mathbf{X}_1 + \mathbf{B}^{-1} (\mathbf{I}_p - \Phi^{GD}) \mathbf{B} \mathbf{X}_2, \quad \Phi^{GD} = \text{diag}(f_1, \dots, f_p).$$

3 リスクと Stein 型推定量

いま, U を $(\mathbf{X}_1, \mathbf{X}_2, \mathbf{Z}_1, \mathbf{Z}_2)$ の関数とし,

$$\mathbb{E}_G[U] = \int U \times \left(\prod_{i=1}^2 |\Sigma_i|^{-N/2} \right) G(C) d\mathbf{X}_1 d\mathbf{X}_2 d\mathbf{Z}_1 d\mathbf{Z}_2$$

と書くことにする. ただし, $G(x) = (1/2) \int_x^{+\infty} g(t) dt$,

$$C = \sum_{i=1}^2 [\text{tr} \{ \Sigma_i^{-1} (\mathbf{X}_i - \boldsymbol{\theta})(\mathbf{X}_i - \boldsymbol{\theta})' + \Sigma_i^{-1} \mathbf{Z}_i' \mathbf{Z}_i \}]$$

である.

上記の記法と Kubokawa and 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, 推定量(5) の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

$$\begin{aligned}
R(\hat{\boldsymbol{\theta}}; \boldsymbol{\theta}) &= \mathbb{E}_G \left[\sum_{j=1}^p [\mathbf{B}(\mathbf{X}_1 - \mathbf{X}_2)]_j^2 \left\{ \frac{n-p-1}{f_j} \phi_j^2 + 4 \frac{1-f_j}{f_j} \phi_j^2 + 4(1-f_j) f_j \phi_j \frac{\partial}{\partial f_j} \left(\frac{\phi_j}{f_j} \right) \right. \right. \\
&\quad + \sum_{k \neq j} \phi_j (\phi_j - \phi_k) \frac{1-f_k}{f_j - f_k} + \frac{n-p-1}{1-f_j} (1-\phi_j)^2 + 4 \frac{f_j}{1-f_j} (1-\phi_j)^2 \\
&\quad \left. \left. + 4f_j(1-f_j)(1-\phi_j) \frac{\partial}{\partial(1-f_j)} \left(\frac{1-\phi_j}{1-f_j} \right) + \sum_{k \neq j} (1-\phi_j)(\phi_j - \phi_k) \frac{f_k}{f_j - f_k} \right\} \right].
\end{aligned}$$

ただし, $[\mathbf{B}(\mathbf{X}_1 - \mathbf{X}_2)]_j$ は $\mathbf{B}(\mathbf{X}_1 - \mathbf{X}_2)$ の第 j 成分をあらわす.

ここで Theorem 1 の $\mathbb{E}_G[\cdot]$ 内の式に注目する. その式の微分を含む項を無視し, 形式的に ϕ_j で微分し評価することにより, 次のような Stein 型推定量 $\hat{\boldsymbol{\theta}}^{ST} = \mathbf{B}^{-1} \boldsymbol{\Phi}^{ST} \mathbf{B} \mathbf{X}_1 + \mathbf{B}^{-1} (\mathbf{I}_p - \boldsymbol{\Phi}^{ST}) \mathbf{B} \mathbf{X}_2$ を得る:

$$\begin{aligned}
\boldsymbol{\Phi}^{ST} &= \text{diag}(\hat{\phi}_1^{ST}, \dots, \hat{\phi}_p^{ST}), \quad \hat{\phi}_j^{ST} = \frac{\hat{\beta}_j^{ST}/(1-f_j)}{\hat{\beta}_j^{ST}/(1-f_j) + \hat{\alpha}_j^{ST}/f_j} \quad (j = 1, \dots, p), \quad (6) \\
\hat{\alpha}_j^{ST} &= n - p - 1 + 4(1-f_j) + 2 \sum_{k \neq j} \frac{f_j(1-f_k)}{f_j - f_k}, \\
\hat{\beta}_j^{ST} &= n - p - 1 + 4f_j - 2 \sum_{k \neq j} \frac{(1-f_j)f_k}{f_j - f_k}.
\end{aligned}$$

4 数値実験

数値実験では, 誤差項の各行 e_{ij} ($i = 1, 2, j = 1, 2, \dots, N$) の分布として, 次の2つの分布を考えた:

1. 多変量 t 分布:

$$\kappa_1 |\boldsymbol{\Sigma}_i|^{-1/2} (1 + \mathbf{e}'_{ij} \boldsymbol{\Sigma}_i^{-1} \mathbf{e}_{ij} / v)^{-(v+p)/2}.$$

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

2. 多変量 Kotz 型分布

$$\kappa_2 |\boldsymbol{\Sigma}_i|^{-1/2} \{ \mathbf{e}'_{ij} \boldsymbol{\Sigma}_i^{-1} \mathbf{e}_{ij} \}^{u-1} \exp[-r \{ \mathbf{e}'_{ij} \boldsymbol{\Sigma}_i^{-1} \mathbf{e}_{ij} \}^s],$$

ただし, $r > 0$, $s > 0$, $2u + p > 2$,

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

多変量 Kotz 型分布の乱数発生法については, Fang, Kotz, and Ng (1990) を参照のこと.

実験では, $\boldsymbol{\xi} = \mathbf{0}$ とし, $\boldsymbol{\Sigma}_2 \boldsymbol{\Sigma}_1^{-1}$ は対角行列とした. また $(N, p) = (8, 5), (13, 10)$ とし, 多変量 t 分布については $v = 3$, 多変量 Kotz 型分布については $(u, r, s) = (5, 0.5, 2)$ とした.

これらの分布に対する数値実験の結果は、それぞれ表 1, 2 に与えてある。Stein 型結合推定量については、 Φ^{ST} の成分 (6) を “isotonic regression” を用いて順序付けの修正を行った (Lin and Perlman (1985) を参照)。表内で、“ML”, “GD”, “ST” はそれぞれ $\hat{\theta}^{ML}$, $\hat{\theta}^{GD}$, $\hat{\theta}^{ST}$ をあらわし、“AV” は GD に対する ST のリスクの改良率を表している。ここで、 $\hat{\theta}^{ML}$ は Σ_1 と Σ_2 が既知の場合の最尤推定量である：

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

数値実験の結果をまとめると次のようになる：

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り、 $(N, p) = (13, 10)$ 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。
4. 推定量 ST は、密度関数 (3) のもとで導出したが、これら実験結果から密度関数 (2) の下でも、ST による GD の改良は有効であると思われる。

参考文献

- [1] Fang, K.T., Kotz, S., and Ng, K.W. (1990). *Symmetric multivariate and related distributions*. Chapman and Hall, New York.
- [2] Graybill, F.A. and Deal, R.B. (1959). Combined unbiased estimator. *Biometrics* **15**, 543–550.
- [3] Kubokawa, T. and Srivastava, M.S. (1999). Robust improvement in estimation of a covariance matrix in an elliptically contoured distribution. *Ann. Statist.* **27**, 600–609.
- [4] Kubokawa, T. and Srivastava, M.S. (2001). Robust improvement in estimation of a mean matrix in an elliptically contoured distribution. *J. Multivariate Anal.* **76**, 138–152.
- [5] Lin, S.P. and Perlman, M.D. (1985). *A Monte Carlo comparisons of four estimators for a covariance matrix*. In *Multivariate Analysis VI* (P.K. Krishnaiah, ed.) 411–429.
- [6] Loh, W.L. (1991). Estimating the common mean of two multivariate normal distributions. *Ann. Statist.* **19**, 297–313.
- [7] Tsukuma, H. and Konno, Y. (2002). Alternative estimators of the common regression matrix in two GMANOVA models under weighted quadratic losses. Technical Reports of Mathematical Sciences, Chiba University, Vol.18, No. 12. (<http://www.math.s.chiba-u.ac.jp/report/content.html>)

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)

$\Sigma_2 \Sigma_1^{-1}$ の固有値	ML	GD	ST	AV
$N = 8, p = 5$				
(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 %
(10, 10, 10, 0.1, 0.1)	14.140 (0.389)	29.514 (1.060)	29.140 (1.062)	1.27 %
(10, 1, 1, 1, 0.1)	13.800 (0.442)	26.304 (0.895)	24.916 (0.791)	5.28 %
(10, 10, 1, 0.1, 0.1)	14.687 (0.395)	29.810 (1.295)	29.729 (1.310)	0.27 %
(20, 5, 1, 0.5, 0.05)	15.162 (0.586)	28.074 (0.749)	27.495 (0.729)	2.06 %
($10^{10}, 5, 1, 0.5, 10^{-10}$)	14.042 (0.380)	27.115 (0.771)	26.693 (0.730)	1.56 %
(5, 2, 1, 0.5, 0.2)	15.789 (1.132)	30.963 (2.789)	30.059 (3.115)	2.92 %
(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 %
$N = 13, p = 10$				
(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 %
(10, 10, 10, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1)	28.492 (0.851)	67.595 (2.224)	60.339 (1.839)	10.73 %
(10, 10, 10, 10, 10, 0.1, 0.1, 0.1, 0.1, 0.1)	29.685 (0.861)	64.710 (2.408)	62.234 (2.232)	3.83 %
(10, 10, 10, 10, 10, 10, 10, 10, 0.1, 0.1)	31.737 (2.690)	82.026 (5.284)	69.922 (7.012)	14.76 %
(10, 9/2, 8/3, 7/4, 6/5, 5/6, 4/7, 3/8, 2/9, 1/10)	28.709 (0.788)	60.807 (1.733)	54.396 (1.567)	10.54 %
(10, 10, 10, 1, 1, 1, 1, 0.1, 0.1, 0.1)	30.111 (1.316)	61.654 (2.212)	57.437 (2.089)	6.84 %
(512, 256, 128, 64, 32, 16, 8, 4, 2, 1)	28.641 (1.594)	98.799 (4.157)	67.775 (2.573)	31.40 %
(32, 16, 8, 4, 2, 1, 1/2, 1/4, 1/8, 1/16)	27.964 (0.995)	58.601 (1.430)	54.855 (1.306)	6.39 %
(720, 360, 120, 30, 6, 1, 1/6, 1/30, 1/120, 1/360)	29.767 (0.949)	65.894 (2.429)	65.812 (2.423)	0.12 %
($10^{10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}$)	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)

$\Sigma_2 \Sigma_1^{-1}$ 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 %
(10, 10, 10, 0.1, 0.1)	2.452 (0.011)	3.861 (0.023)	3.775 (0.022)	2.22 %
(10, 1, 1, 1, 0.1)	2.462 (0.012)	3.760 (0.021)	3.552 (0.019)	5.54 %
(10, 10, 1, 0.1, 0.1)	2.446 (0.011)	3.804 (0.022)	3.775 (0.022)	0.76 %
(20, 5, 1, 0.5, 0.05)	2.464 (0.011)	3.799 (0.023)	3.735 (0.022)	1.70 %
($10^{10}, 5, 1, 0.5, 10^{-10}$)	2.453 (0.011)	3.768 (0.022)	3.720 (0.021)	1.27 %
(5, 2, 1, 0.5, 0.2)	2.452 (0.012)	3.748 (0.021)	3.473 (0.019)	7.35 %
(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 %
(10, 10, 10, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1)	2.899 (0.010)	5.313 (0.027)	4.809 (0.023)	9.49 %
(10, 10, 10, 10, 10, 0.1, 0.1, 0.1, 0.1, 0.1)	2.914 (0.010)	5.098 (0.024)	4.859 (0.022)	4.69 %
(10, 10, 10, 10, 10, 10, 10, 10, 0.1, 0.1)	2.925 (0.010)	5.609 (0.031)	4.688 (0.024)	16.42 %
(10, 9/2, 8/3, 7/4, 6/5, 5/6, 4/7, 3/8, 2/9, 1/10)	2.905 (0.010)	5.013 (0.022)	4.379 (0.018)	12.65 %
(10, 10, 10, 1, 1, 1, 1, 0.1, 0.1, 0.1)	2.911 (0.010)	5.070 (0.023)	4.679 (0.020)	7.72 %
(512, 256, 128, 64, 32, 16, 8, 4, 2, 1)	2.913 (0.009)	6.449 (0.048)	4.409 (0.028)	31.63 %
(32, 16, 8, 4, 2, 1, 1/2, 1/4, 1/8, 1/16)	2.907 (0.010)	5.093 (0.024)	4.760 (0.021)	6.54 %
(720, 360, 120, 30, 6, 1, 1/6, 1/30, 1/120, 1/360)	2.909 (0.010)	5.150 (0.025)	5.127 (0.025)	0.46 %
($10^{10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}$)	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 %