# Shrinkage estimators in multivariate normal distributions with block compound symmetry covariance structures(修正版)

今野　良彦　(大阪公立大学)*

## 1. 概要

多変量正規分布の分散共分散行列が block compound symmetry と呼ばれる制約下にあるときに, 平均ベクトルの推定問題を統計的決定理論の立場から考える. 適当な直交行列で標本空間の基底を変換することで, 標本空間の座標に作用する置換群の部分群に関して不変な構造をもつ分散共分散行列は, 成分が実, 複素, 四元数である Hermite 行列の直和で書けるできること [6] を援用して, 平均ベクトルの縮小型推定量を導出できることを報告する.

## 2. 問題の背景

多変量正規分布の分散共分散行列に不変性（対称性）によって構造化することができる. すなわち, 観測値の空間に作用する一般線型群のある部分群の作用に関して不変であるという制約を分散共分散行列にかすこと [1] によって特徴づけるすることである. この観点の背景にある事実として, 多変量正規分布の正則な分散共分散行列全体の成す空間は対称錐 (等質かつ自己双対な凸錐) の一例とみることができ, さらに, 対称錐は Euclid 型 Jordan 代数 (以下では EJA と記す) を用いて書き表せること [4] がある.

別の観点からの分散共分散行列の制約として, 分散共分散行列とその逆行列に関して線型的であることを制約とかすこともできる. このような構造をもつ分散共分散行列が EJA によって母数化されることが同値であること [7] が知られている. これも EJA の議論が背景にある.

さらに, [6] では, 適当な直交行列で標本空間の基底を変換することで, 標本空間の座標に作用する置換群の部分群に関して不変な構造をもつ分散共分散行列は, 成分が実, 複素, 四元数である Hermite 行列の直和で書けるできることを有限群の表現論を用いて証明している.

以上の一般論を踏まえて, 統一的な形で推測手法の性質を議論することが可能になることが予想される. そこで, 具体的なモデルである block compound symmetry 分散共分散行列をとりあげ, 不変構造をもつ分散共分散行列に関する一般論を踏まえると, この制約下での推定問題が見通しよく議論できることがわかる.

## 3. 問題設定

block compound symmetry 構造をもつモデルの中で最も簡単なモデル

$$\boldsymbol{Y} \sim \mathsf{N}_d\big(\widetilde{\boldsymbol{\xi}}, \boldsymbol{\Sigma}\big) \qquad \text{with} \qquad \boldsymbol{\Sigma} = \sigma^2\big\{(1-\rho)\mathbf{I}_d + \rho\mathbf{1}_d\mathbf{1}_d^\top\big\} \tag{1}$$

* 〒558-8585 大阪市住吉区杉本 3-3-138　大阪公立大学 大学院理学研究科
e-mail: konno@fc.jwu.ac.jp
web: https://mcm-www.jwu.ac.jp/~konno/index.html

を考える．ただし, $\widetilde{\boldsymbol{\xi}} \in \mathbb{R}^d$, $\mathbf{I}_d$ は $d$ 次の単位行列, $\mathbf{1}_d = (1, 1, \ldots, 1)^\top \in \mathbb{R}^d$, $0 < \sigma < \infty$, $-\dfrac{1}{d-1} < \rho < 1$ である．また, $(\cdot)^\top$ は転置作用素である．このモデルは intraclass covariance structures とも呼ばれる分散共分散構造をもつ多変量正規モデルである．

$n \geq 1$ とする．$\boldsymbol{Y}$ の $n+1$ 個の独立複製を $\boldsymbol{Y}_1, \boldsymbol{Y}_2, \ldots, \boldsymbol{Y}_{n+1}$ に対して

$$\overline{\boldsymbol{Y}} = \frac{1}{n+1}\sum_{i=1}^{n+1}\boldsymbol{Y}_i, \quad \boldsymbol{S} = \sum_{i=1}^{n+1}(\boldsymbol{Y}_i - \overline{\boldsymbol{Y}})(\boldsymbol{Y}_i - \overline{\boldsymbol{Y}})^\top.$$

とおく．すると, $\overline{\boldsymbol{Y}}$ と $\mathbf{S}$ は独立で

$$\overline{\boldsymbol{Y}} \sim \mathsf{N}_d(\widetilde{\boldsymbol{\xi}}, (n+1)^{-1}\boldsymbol{\Sigma}), \quad \boldsymbol{S} \sim \mathsf{W}_d(n, \boldsymbol{\Sigma}).$$

となる．$\sigma$ と $\rho$ を未知としたとき, 平均ベクトル $\widetilde{\boldsymbol{\xi}}$ の推定問題を不変な損失関数

$$\widetilde{\mathsf{L}}\big(\widetilde{\boldsymbol{d}}, \widetilde{\boldsymbol{\xi}} | \boldsymbol{\Sigma}\big) = (n+1)\big(\widetilde{\boldsymbol{d}} - \widetilde{\boldsymbol{\xi}}\big)^\top \boldsymbol{\Sigma}^{-1}\big(\widetilde{\boldsymbol{d}} - \widetilde{\boldsymbol{\xi}}\big)$$

のもとで考える．ただし, $\widetilde{\boldsymbol{d}}$ は, 観測 $\big(\overline{\boldsymbol{Y}}, \boldsymbol{S}\big)$ に基づく $\widetilde{\boldsymbol{\xi}}$ の推定量である．さらに, $\widetilde{\mathsf{R}}(\widetilde{\boldsymbol{d}}, \widetilde{\boldsymbol{\xi}} | \boldsymbol{\Sigma})$ を $\widetilde{\boldsymbol{d}}$ の危険関数とする．すなわち

$$\widetilde{\mathsf{R}}\big(\widetilde{\boldsymbol{d}}, \widetilde{\boldsymbol{\xi}} | \boldsymbol{\Sigma}\big) = \mathsf{E}\big[\widetilde{\mathsf{L}}\big(\widetilde{\boldsymbol{d}}, \widetilde{\boldsymbol{\xi}} | \boldsymbol{\Sigma}\big)\big]$$

である．ただし, $\mathsf{E}[\,\cdot\,]$ は $\big(\overline{\boldsymbol{Y}}, \boldsymbol{S}\big)$ の同時分布に関する期待値である．

まず, この推定問題の通常の canonical form を導出するために

$$\boldsymbol{X} = \sqrt{n+1} \times \overline{\boldsymbol{Y}} \sim \mathsf{N}_d\big(\boldsymbol{\xi}, \boldsymbol{\Sigma}\big) \quad \text{and} \quad \boldsymbol{S} \sim \mathsf{W}_d\big(n, \boldsymbol{\Sigma}\big) \tag{2}$$

とおく．ただし, $\mathsf{W}_d\big(n, \boldsymbol{\Sigma}\big)$ は自由度 $n$, スケール行列 $\boldsymbol{\Sigma}$ の $d$ 次の Wishart 分布で, $\boldsymbol{\xi} := \sqrt{n+1}\widetilde{\boldsymbol{\xi}}$ で, $\boldsymbol{\Sigma}$ は (1) で与えられる．ここで, $\boldsymbol{\Sigma}$ は正定置であるが, $n < d$ の場合には $\mathbf{S}$ は正則ではない (その逆行列が存在しない) ことに注意する．平均ベクトル $\widetilde{\boldsymbol{\xi}}$ の推定問題は, 平均ベクトル $\boldsymbol{\xi}$ を不変な損失関数

$$\mathsf{L}\big(\widehat{\boldsymbol{\xi}}, \boldsymbol{\xi} | \boldsymbol{\Sigma}\big) = \big(\widehat{\boldsymbol{\xi}} - \boldsymbol{\xi}\big)^\top \boldsymbol{\Sigma}^{-1}\big(\widehat{\boldsymbol{\xi}} - \boldsymbol{\xi}\big) \tag{3}$$

のもとで考えることと同一視できる．ただし, $\widehat{\boldsymbol{\xi}} = \sqrt{n+1}\widetilde{\boldsymbol{d}}$ は観測 $\big(\boldsymbol{X}, \boldsymbol{S}\big)$ に基づく推定量である．

**Canonical form 1**

> $\boldsymbol{X}$ と $\mathbf{S}$ は独立に (2) の分布に従い, 分散共分散は (1) で与えらえる．ただし, $\sigma > 0$, $-\dfrac{1}{d-1} < \rho < 1$ である．このとき, $\big(\boldsymbol{X}, \mathbf{S}\big)$ に基づき $\boldsymbol{\xi}$ の推定問題を不変な損失関数 (3) のもとで考える．

つぎに, 分散共分散構造 (1) を考慮に入れて canonical form を求めるために

$$\boldsymbol{Z} := \boldsymbol{U}\boldsymbol{X} \sim \mathsf{N}_d\big(\boldsymbol{\mu}, \boldsymbol{\Psi}\big) \quad \text{and} \quad \boldsymbol{W} = \boldsymbol{U}\boldsymbol{S}\boldsymbol{U}^\top \sim \mathsf{W}_d\big(n, \boldsymbol{\Psi}\big) \tag{4}$$

と変換する．ただし, $\boldsymbol{U}$ は $d \times d$ 直交行列 (観測者にとって既知) で第 1 列は $\dfrac{1}{\sqrt{d}}\mathbf{1}$ となるものである．すると分散共分散行列 $\boldsymbol{\Sigma}$ は対角化され

$$\boldsymbol{\mu} = \boldsymbol{U}\boldsymbol{\xi}, \quad \boldsymbol{\Psi} = \sigma^2 \mathrm{diag}\big(1 + (d-1)\rho, \underbrace{1-\rho, \ldots, 1-\rho}_{(d-1)\text{ 個}}\big) =: \mathrm{diag}\big(\psi_1, \underbrace{\psi_2, \ldots, \psi_2}_{(d-1)\text{ 個}}\big) \tag{5}$$

となることに注意する．ここで, $\mathrm{diag}(a_1, a_2, \ldots, a_d)$ は $d \times d$ の対角行列で, その $i$ 対角成分 (i=1, 2,. . . , d) は $a_i$ $(\in \mathbb{R})$ であるものを表し, $\psi_1 := \sigma^2(1+(d-1)\rho)$, $\psi_2 := \sigma^2(1-\rho)$ である．このとき, 損失関数 (3) は

$$\begin{aligned}\mathsf{L}\big(\boldsymbol{d}, \boldsymbol{U}^\top\boldsymbol{\mu} | \boldsymbol{U}^\top\boldsymbol{\Psi}\boldsymbol{U}\big) &= \big(\boldsymbol{U}\boldsymbol{d} - \boldsymbol{\mu}\big)^\top \boldsymbol{\Psi}^{-1}\big(\boldsymbol{U}\boldsymbol{d} - \boldsymbol{\mu}\big) = \big(\widehat{\boldsymbol{\mu}} - \boldsymbol{\mu}\big)^\top \boldsymbol{\Psi}^{-1}\big(\widehat{\boldsymbol{\mu}} - \boldsymbol{\mu}\big) \\ &=: \mathsf{L}\big(\widehat{\boldsymbol{\mu}}, \boldsymbol{\mu} | \boldsymbol{\Psi}\big)\end{aligned}$$

となる．ただし $\widehat{\boldsymbol{\mu}} = \boldsymbol{U}\boldsymbol{d}$ とおいた．

**Canonical form 2**

$\boldsymbol{Z}$ と $\mathbf{W}$ は独立に (4) の分布に従い, 分散共分散行列は (5) で与えられる．ただし, $0 < \psi_2 < \psi_1$ である．このとき, $\big(\boldsymbol{Z}, \mathbf{W}\big)$ に基づいて $\boldsymbol{\mu}$ の推定問題を損失関数

$$\mathsf{L}\big(\widehat{\boldsymbol{\mu}}, \boldsymbol{\mu} | \boldsymbol{\Psi}\big) = \big(\widehat{\boldsymbol{\mu}} - \boldsymbol{\mu}\big)^\top \boldsymbol{\Psi}^{-1}\big(\widehat{\boldsymbol{\mu}} - \boldsymbol{\mu}\big)$$

のもとで考える．ただし, $\widehat{\boldsymbol{\mu}}$ は $\big(\boldsymbol{Z}, \mathbf{W}\big)$ に基づく $\boldsymbol{\mu}$ の推定量である．

## 4. 縮小型推定量

次のような推定量の族

$$\widehat{\boldsymbol{\mu}}(c) = \left(1 - \frac{c}{\boldsymbol{Z}^\top \widehat{\boldsymbol{\Psi}}^{-1} \boldsymbol{Z}}\right) \boldsymbol{Z} \tag{6}$$

を考える．ただし, $c$ は非負定数で, $\mathbf{W} = \mathrm{diag}(w_{11}, \ldots, w_{dd})$ と記したとき

$$\widehat{\boldsymbol{\Psi}} = \mathrm{diag}\left(\frac{w_{11}}{n}, \underbrace{\frac{\sum_{i=2}^d w_{ii}}{(d-1)n}, \ldots, \frac{\sum_{i=2}^d w_{ii}}{(d-1)n}}_{(d-1)\,\text{times}}\right)$$

で与えられる．すると, 正規分布と $\chi^2$ 分布に関する部分積分の公式 (the Stein identity と the Stein-Haff identity) を用いることで

$$\begin{aligned}\mathsf{R}\big(\widehat{\boldsymbol{\mu}}(c), \boldsymbol{\mu} | \boldsymbol{\Psi}\big) &= \mathsf{E}\big[\mathsf{L}\big(\widehat{\boldsymbol{\mu}}(c), \boldsymbol{\mu} | \boldsymbol{\Psi}\big)\big] \\ &\leq \mathsf{E}\big[\big(\boldsymbol{Z} - \boldsymbol{\mu}\big)^\top \boldsymbol{\Psi}^{-1}\big(\boldsymbol{Z} - \boldsymbol{\mu}\big)\big] + (c^2 - 2(d-2)c)\mathsf{E}\left[\frac{1}{\boldsymbol{Z}^\top \widehat{\boldsymbol{\Psi}}^{-1} \boldsymbol{Z}}\right]\end{aligned}$$

がわかる．

**定理 1**

$d \geq 3$ かつ $n \geq 1$ とする．任意の $\boldsymbol{\mu} \in \mathbb{R}^d$ と $0 < \psi_2 < \psi_1 < \infty$ に対して

$$0 \leq c \leq 2(d-2) \ \Rightarrow \ \mathsf{R}\big(\widehat{\boldsymbol{\mu}}(c),\, \boldsymbol{\mu} \big|\, \boldsymbol{\Psi}\big) \leq \mathsf{R}\big(\boldsymbol{Z},\, \boldsymbol{\mu} \big|\, \boldsymbol{\Psi}\big).$$

ここで，$c = d-2$ とおくと，canonical form 1 の設定で，$\boldsymbol{\xi}$ の James-Stein 型推定量とその positive-part 推定量は

$$\widehat{\boldsymbol{\xi}}^{\mathrm{JS}} = \left(1 - \frac{d-2}{(d-1)n\boldsymbol{X}^\top\left\{(\mathrm{Tr}\,\boldsymbol{S} - \frac{1}{d}\mathbf{1}_d^\top \mathbf{S}\mathbf{1}_d)\boldsymbol{I}_d + \frac{1}{d}(\mathbf{1}_d^\top \boldsymbol{S}\mathbf{1}_d - \mathrm{Tr}\,\boldsymbol{S})\mathbf{1}_d\mathbf{1}_d^\top\right\}^{-1}\boldsymbol{X}}\right)\boldsymbol{X},$$

$$\widehat{\boldsymbol{\xi}}^{\mathrm{PJS}} = \left(1 - \frac{d-2}{(d-1)n\boldsymbol{X}^\top\left\{(\mathrm{Tr}\,\boldsymbol{S} - \frac{1}{d}\mathbf{1}_d^\top \mathbf{S}\mathbf{1}_d)\boldsymbol{I}_d + \frac{1}{d}(\mathbf{1}_d^\top \boldsymbol{S}\mathbf{1}_d - \mathrm{Tr}\,\boldsymbol{S})\mathbf{1}_d\mathbf{1}_d^\top\right\}^{-1}\boldsymbol{X}}\right)_+\boldsymbol{X}$$

で与えられる．ただし，$\mathrm{Tr}(\cdot)$ は行列のトレースで，$(a)_+ = \max\{0,\, a\}$ for $a \in \mathbb{R}$ である．

**定理 2**

任意の $\boldsymbol{\xi} \in \mathbb{R}^d$ と $\sigma > 0$, $-\dfrac{1}{d-1} < \rho < 1$ に対して

$$\mathsf{R}\big(\widehat{\boldsymbol{\xi}}^{\mathrm{PJS}},\, \boldsymbol{\xi} \big|\, \boldsymbol{\Sigma}\big) \leq \mathsf{R}\big(\widehat{\boldsymbol{\xi}}^{\mathrm{JS}},\, \boldsymbol{\xi} \big|\, \boldsymbol{\Sigma}\big) \leq \mathsf{R}\big(\boldsymbol{X},\, \boldsymbol{\xi} \big|\, \boldsymbol{\Sigma}\big).$$

## 5. Block compound symmetry covarinace matrix model

$dp \times 1$ 確率ベクトル $\boldsymbol{X}$ は多変量正規分布 $\mathsf{N}_{dp}\big(\widetilde{\boldsymbol{\xi}},\, \mathsf{Cov}(\boldsymbol{X})\big)$ に従うとする．ただし，$\widetilde{\boldsymbol{\xi}} \in \mathbb{R}^{dp}$ で

$$\mathsf{Cov}(\boldsymbol{X}) = \begin{pmatrix} \mathbf{\Sigma} & \mathbf{R} & \cdots & \mathbf{R} \\ \mathbf{R} & \mathbf{\Sigma} & \cdots & \mathbf{R} \\ \vdots & & \ddots & \vdots \\ \mathbf{R} & \mathbf{R} & \cdots & \mathbf{\Sigma} \end{pmatrix} =: \mathbf{I}_d \otimes \big(\mathbf{\Sigma} - \mathbf{R}\big) + \mathbf{1}_d\mathbf{1}_d^\top \otimes \mathbf{R}$$

である．ここで，$\mathbf{\Sigma}$ は $p \times p$ の正定置行列，$\mathbf{R}$ は $p \times p$ の対称行列で $-\dfrac{1}{d-1}\mathbf{\Sigma} < \mathbf{R} < \mathbf{\Sigma}$ をみたすと仮定する．すなわち，$-\dfrac{1}{d-1}\mathbf{\Sigma} < \mathbf{R} < \mathbf{\Sigma}$ により $\mathbf{\Sigma} - \mathbf{R}$ と $\mathbf{\Sigma} + (d-1)\mathbf{R}$ は共に $p \times p$ の正定置行列となることがわかる．$d = 1$ かつ $\mathbf{\Sigma} = \sigma^2$, $\mathbf{R} = \sigma^2\rho$ とすると (1) となることに注意せよ．

## 6. 終わりの注意

モデル (1) のもとでの数値実験による改良率と block compound symmetry covariance structures モデルへの拡張された結果，分散共分散行列の推定問題については当日報告する予定である．

# References


[1] Andersson, S. (1975). Invariant Normal Models. *Ann. Statist.* **3** 132-154.

[2] Coelho, C.A., Roy A. (2017). Testing the hypothesis of a block compound symmetric covariance matrix for elliptically contoured distributions. *Test* **26** 308-330.

[3] Dey, D.K., Gelfand, A.E. (1989). Improved Estimation of a patterned covariance matrix. *J. Multivariate Anal.* **31** 107-116.

[4] Faraut, J., Korányi, A.(1994). Analysis on symmetric cones. The Clarendon Press, Oxford University Press, New York.

[5] Fourdrinier, D., Strawderman, W.E., Wells, M.T. (2018). Shrinkage Estimation. Springer.

[6] Graczyk, P., Ishi, H., Kołodziejek, B., Massam, H. (2022). Model selection in the space of Gaussian models invariant by symmetry. *Ann. Statist.* **50** 1747-1774.

[7] Jensen, S.T. (1988). Covariance hypotheses which are linear in both the covariance and the inverse covariance. *Ann. Statist* **16** 302-322.

[8] Jensen, S. T., Madsen, J. (2004). Estimation of Proportional Covariances in the Presence of Certain Linear Restrictions. *Ann. Statist.* **32** 219-232.

[9] Liang, Y., Coelho, C.A., von Rosen, T. (2020). Hypothesis testing in multivariate normal models with block circular covariance structures. *Biometrical Journal* **64** 557-576.

[10] Loh, W.-L. (1991). Estimating covariance matrices. *Ann. Statist* **19** 283-296.

[11] Perlman, M.D. (1987). [A Review of Multivariate Analysis]: Cooment: Group Symmetry Covariance Models. *Statistical Scinece* **2** 421-425.

[12] Perlman, M.D. (2018). STAT 542 Notes, Mulitvariate Statistical Analysis. `https://sites.stat.washington.edu/people/mdperlma/542%20Multivariate%20Analysis%20Material/542%20Notes.pdf` (2025/05/07 accessed).

[13] Roy, A., Fonseca, M. (2012). Linear models with doubly exchangeable distributed errors. *Comm. Statist. Theory Methods* **41** 25452569.

[14] Srivastava, M.S., Singull, M. Testing sphericity and intraclass covariance structures under a growth curve model in high dimension. *Commun. Statist.-Simulation and Computation* **46** 5740-5751.

[15] Stein, C. (1981). Estimation of the mean of a multivariate normal distribution. *Ann. Statist.* **9** 1135-1151.

[16] Tsukuma, H., Kubokawa, T. (2015). A unified approach to estimating a normal mean matrix in high and low dimensions. *J. Multivariate Anal.* **139** 312-328.

[17] Tsukuma, H., Kubokawa, T. (2020). Shrinkage Estimation for Mean and Covariance Matrices. Springer.

[18] Zwiernik, P., Uhler,C., Richards, D. (2017). Maximum likelihood estimation for linear Gaussian covariance models. *J. R. Statist. Soc. Ser. B. Statist. Methodol.* **79** 12691292.