Test for block-diagonal covariance structure in high dimension under non-normality

専修大・経営　西山　貴弘
東京理科大・理・院　山田　雄紀
大阪府立大・工　兵頭　昌

近年, 高次元データに対する共分散構造に関する検定問題について多くの研究がされている（Chen, Zhang and Zhong (2010), Hyodo, et al. (2015) などを参照）. 本報告では, 特に共分散構造が "ブロック対角構造" を持つかどうかの検定について議論し, この問題に対して新たな検定統計量を提案する.

いま, $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N$ を平均ベクトル $\boldsymbol{\mu}$, 分散共分散行列 Σ の p 次元母集団からの互いに独立な N 個の観測ベクトルとし, $\boldsymbol{x}_i = \boldsymbol{\mu} + \Gamma \boldsymbol{z}_i$ $(i = 1, \ldots, N)$ を満たすものとする. ここで, Γ は $\Gamma\Gamma' = \Sigma$ を満たす $p \times m$ 行列であり, $\boldsymbol{z}_i = (z_{i1}, \ldots, z_{im})'$ は $\mathrm{E}[\boldsymbol{z}_i] = \boldsymbol{0}$, $\mathrm{Var}[\boldsymbol{z}_i] = I_m$ を満たす m 次元ベクトルである. また, $\boldsymbol{x}_i$, $\boldsymbol{\mu}$, Σ はそれぞれ次のように分割されるものとする.

$$\boldsymbol{x}_i = \begin{pmatrix} \boldsymbol{x}_i^{(1)} \\ \boldsymbol{x}_i^{(2)} \\ \vdots \\ \boldsymbol{x}_i^{(q)} \end{pmatrix}, \ \boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}^{(1)} \\ \boldsymbol{\mu}^{(2)} \\ \vdots \\ \boldsymbol{\mu}^{(q)} \end{pmatrix}, \ \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} & \cdots & \Sigma_{1q} \\ \Sigma_{21} & \Sigma_{22} & \cdots & \Sigma_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ \Sigma_{q1} & \Sigma_{q2} & \cdots & \Sigma_{qq} \end{pmatrix}.$$

ここで, $\ell, m = 1, \ldots, q$ に対して $\boldsymbol{x}^{(\ell)}$, $\boldsymbol{\mu}^{(\ell)}$ は p_ℓ 次元ベクトル, $\Sigma_{\ell m}$ は $p_\ell \times p_m$ 行列とする. このとき, $\Sigma_d = \mathrm{diag}(\Sigma_{11}, \Sigma_{22}, \cdots, \Sigma_{qq})$ とし, 次の仮説検定問題を考える.

$$H_0 : \Sigma = \Sigma_d \quad \text{vs.} \quad H_1 : \Sigma \neq \Sigma_d.$$

一般に, $p \leq n$ $(= N - 1)$ の場合, 共分散構造の検定問題に対して尤度比検定統計量が用いられるが, $p > n$ の場合は用いることが出来ない. そのため, ここでは Frobenius ノルムに基づいた検定を考え, 次の検定統計量を提案する.

$$T = \widehat{\mathrm{tr}\Sigma^2} - \widehat{\mathrm{tr}\Sigma_d^2}. \qquad (*)$$

ここで, $\widehat{\mathrm{tr}\Sigma^2}$, $\widehat{\mathrm{tr}\Sigma_d^2}$ はそれぞれ $\mathrm{tr}\Sigma^2$, $\mathrm{tr}\Sigma_d^2$ の不偏推定量である（詳細は Himeno and Yamada (2014) などを参照）.

本報告では，次元と標本サイズが共に大である場合の $(*)$ の極限分布を，いくつかの高次元枠組みの下で導出する．さらに, 得られた理論結果に対して，モンテカルロ・シミュレーションによって数値的に近似精度や検出力を評価する．詳細は当日報告する．

参考文献

[1] Chen, S. X., Zhang, L.-X. and Zhong, P.-S. (2010). "Tests for high-dimensional covariance matrices", *Journal of the American Statistical Association*, **105,** 810–819.

[2] Himeno, T. and Yamada, T. (2014). "Estimations for some functions of covariance matrix in high dimension under non-normality and its applications", *Journal of Multivariate Analysis*, **130,** 27–44.

[3] Hyodo, M., Shutoh, N., Nishiyama, T. and Pavlenko, T. (2015). "Testing block-diagonal covariance structure for high-dimensional data", *to appear in Statistica Neerlandica.*