非正規母集団における分散共分散行列の固有値・固有ベクトルに関する検定統計量

Criteria of hypothesis testing for latent root and vector of covariance matrix

塚田 真一*

概要

多変量統計解析において正値対称行列は様々な分析で用いられている. それらの固有値・固有ベクトルも重要な意味を持つ統計量になっており, 固有値・固有ベクトルに関する統計的仮説検定も重要である. 代表的な正値対称行列は分散共分散行列であり, 主成分分析において各主成分の寄与率を表すものが固有値で, 各変量への重みを与え, 主成分の意味づけに必要となるものが固有ベクトルである. 本論文では分散共分散行列の固有値・固有ベクトルに関する統計的仮説検定を取り上げ, 非正規母集団における検定統計量を提案する.

keywords: Hypothesis testing; Latent roots; Latent vectors; Wald criterion

1 はじめに

分散共分散行列の固有値や固有ベクトルは, 主成分分析における主成分の寄与率や主成分の各変量への重みを表している統計量である. これら固有値・固有ベクトルに関する分布論は多くの研究者により研究されている. 正規母集団での分散共分散行列の固有値・固有ベクトルの極限分布はAnderson[1]により得られており, また精密分布についてはSugiyam[10], [11]により研究されている. Sugiura[8], [9]では形式的な微分演算によりこれらの分布の漸近展開が導出されている.

しかし固有値・固有ベクトルに関する統計的仮説検定は研究されておらず, 過去に多変量正規分布を仮定して議論されているものが殆どである. 特に固有ベクトルに関する仮説検定問題は, 固有値に依存するということから研究が行われていない. そこで本論文では, 母集団分布に正規性を仮定せず分散共分散行列の固有値・固有ベクトルに関する統計的仮説検定問題を考

*TSUKADA, Shin-ichi [新潟国際情報大学情報文化学部情報システム学科] e-mail:tukada@nuis.ac.jp

え, それらの検定統計量を提案する.

2 検定問題

固有値・固有ベクトルに関する検定問題では母集団の状況によって次のような場合が考えられる.

(Ⅰ) 1母集団の場合は, 次のような検定問題が考えられる.

$$H_{01} \ : \ \lambda_\alpha = \lambda_0$$
$$H_{02} \ : \ \boldsymbol{\eta}_\alpha = \boldsymbol{\eta}_0$$
$$H_{03} \ : \ \lambda_\alpha = \lambda_0, \quad \boldsymbol{\eta}_\beta = \boldsymbol{\eta}_0$$

ここで, λ_0は定数, $\boldsymbol{\eta}_0$は既知ベクトル, $\boldsymbol{\eta}_\alpha$は母分散共分散行列 Σ の α 番目に大きい固有値 λ_α に対応する固有ベクトルとする. これらを検定するため, 母集団から$N=(n+1)$個の標本 $\{\boldsymbol{X}_1, \ldots, \boldsymbol{X}_N\}$ を得たとする. この標本から得られる標本分散共分散行列を

$$S = \frac{1}{n}\sum_{i=1}^{N}(\boldsymbol{X}_i - \bar{\boldsymbol{X}})(\boldsymbol{X}_i - \bar{\boldsymbol{X}})'$$

とし, α 番目に大きい固有値を l_α, 対応する固有ベクトルを $\boldsymbol{h}_\alpha$ で表すことにする.

(Ⅱ) 2母集団については,

$$H_{04} \ : \ \lambda_\alpha^{(1)} = \lambda_\beta^{(2)}$$
$$H_{05} \ : \ \boldsymbol{\eta}_\alpha^{(1)} = \boldsymbol{\eta}_\beta^{(2)}$$
$$H_{06} \ : \ \lambda_\alpha^{(1)} = \lambda_\beta^{(2)}, \quad \boldsymbol{\eta}_\alpha^{(1)} = \boldsymbol{\eta}_\beta^{(2)}$$

という検定問題を考える. ここで, $\boldsymbol{\eta}_\alpha^{(g)}(g=1,2)$ は1標本の場合と同様に母分散共分散行列Σ_gの α 番目に大きい固有値 $\lambda_\alpha^{(g)}$ に対応する固有ベクトルとする. また, それぞれの母集団から得た $N_g=(n_g+1)$ 個の標本から計算される統計量として, 標本分散共分散行列をS_g, この行列の α 番目に大きい固有値を$l_\alpha^{(g)}$, 対応する固有ベクトルを$\boldsymbol{h}_\alpha^{(g)}$とする.

3 これまでの研究概要

先ず帰無仮説 H_{02} については正規母集団の場合, Anderson[1]により検定統計量

$$n\left(l_\alpha \boldsymbol{\eta}_0' S^{-1}\boldsymbol{\eta}_0 + \frac{1}{l_\alpha}\boldsymbol{\eta}_0' S\boldsymbol{\eta}_0 - 2\right) \tag{1}$$

が提案されている. また正規母集団を仮定した場合のこの統計量の分布の漸近展開はHayakawa[3], 楕円母集団を仮定した場合の漸近展開は早川[14]により得られており, Mallows[5]による尤度比検定統計量との検出力比較も行っている. 検定統計量の n^{-1} のオーダーでBartlett補正をした検定統計量はSchott[6]により提案されている. Tsukada[12], 塚田・杉山[15]では帰無仮説 H_{02} に関して新しい検定統計量を提案し, それらの分布の漸近展開の導出と検出力の比較を行っている. 塚田・尾形[18]では正規母集団ではなく一般の母集団分布及び楕円母集団を仮定した場合のWald型検定統計量について提案している. この統計量の帰無分布の漸近展開を求め, それに基づいて検定を行うこともできるが, この報告では検定統計量にbootstrap法を用いて検定することを考察しています. また, Mosesのrank-like法を用いたノンパラメトリックな検定法は 牛沢[19]において研究されている.

2母集団における帰無仮説 H_{05} についてはKrzanowski[4], Flury[2], Schott[7]により研究され, それぞれ異なる検定統計量が提案されている. Krzanowski[4]は帰無仮説で用いられる固有ベクトルに対応する標本固有ベクトルによって張られるそれぞれの部分空間のなす角を用いて検定することを提案し, シミュレーション実験により, その有効性も調べている. Flury[2]はそれぞれの母集団に正規性を仮定して, 尤度比検定統計量を提案している. しかし, 尤度比検定統計量を求めるにあたり仮説の下での最尤推定量が必要になるが, 最尤推定量は解析的に得ることはできない. そのため近似的な最尤推定量を求め, それらを用いた近似的な検定統計量を提案し, その計算アルゴリズムも提案している. Schott[7]は固有値を用いた検定統計量を提案しており, その統計量の帰無仮説の下での極限分布も導出している.

正規母集団においてWald型検定統計量とその検出力について検討したものは, 塚田・小野・杉山[16]であり, 塚田・牛沢[17]では帰無仮説 H_{02} についての同時検定についてWald型検定統計量とほかに提案されている統計量の検出力比較を行った. 固有値に対するノンパラメトリック検定はUshizawa, Sato & Sugiyama[13]において研究されている.

4 検定統計量の構成について

本論文で取り上げている検定問題の検定統計量を構成する前に, 仮説検定における一般論を述べておく. 大標本の場合, 母数 $\boldsymbol{\theta}$ について次のような関数制約仮説を考える.

$$H_0 : \boldsymbol{a}(\boldsymbol{\theta}) = \mathbf{0}, \quad H_1 : \boldsymbol{a}(\boldsymbol{\theta}) \neq \mathbf{0}$$

ただし, $\boldsymbol{a}(\boldsymbol{\theta})$ は p 次元ベクトル値関数であり

$$A = \left(\frac{\partial a_i(\boldsymbol{\theta})}{\partial \theta_j}; i = 1, \ldots, p, \quad j = 1, \ldots, k \right)$$

なる $p \times k$ 行列の階数は p であるとする.

このような仮説検定に対して, 一般的に次のような検定統計量が構成できる. ここで $\lambda(\boldsymbol{X})$ は尤度比, $L(\boldsymbol{\theta}|\boldsymbol{X})$ は尤度関数, $I(\boldsymbol{\theta})$ はFisher情報行列とする. また, $\hat{\boldsymbol{\theta}}$ は対立仮説の下での最尤推定量, $\hat{\boldsymbol{\theta}}_0$ は帰無仮説の下での最尤推定量とする.

尤度比検定統計量

$$\begin{aligned} LR &= 2 \log \lambda(\boldsymbol{X}) \\ &= 2 \left\{ \log L(\hat{\boldsymbol{\theta}}|\boldsymbol{X}) - \log L(\hat{\boldsymbol{\theta}}_0|\boldsymbol{X}) \right\} \end{aligned}$$

Wald型検定統計量

$$W = \boldsymbol{a}(\hat{\boldsymbol{\theta}})' \left[AI(\hat{\boldsymbol{\theta}})A' \right]^{-1} \boldsymbol{a}(\hat{\boldsymbol{\theta}})$$

Lagrange乗数法検定統計量

$$LM = S(\hat{\boldsymbol{\theta}}_0)' I(\hat{\boldsymbol{\theta}}_0)^{-1} S(\hat{\boldsymbol{\theta}}_0),$$

$$S(\boldsymbol{\theta}) = \frac{\partial}{\partial \boldsymbol{\theta}} \log L(\boldsymbol{\theta}|\boldsymbol{X})$$

これら3つの検定統計量は標本数 N が無限大のとき仮説の下での分布が自由度 $(p-1)$ の χ^2 分布になり, 統計量として同値になることが知られている.

本論文で扱っている検定問題において, 帰無仮説の下におけるパラメータの最尤推定量を求めることは困難である. 上記の尤度比検定統計量やLagrange乗数法検定統計量を用いようとすると導出困難な帰無仮説の下でのパラメータの最尤推定量が必要となる. しかし, Wald型検定統計量は対立仮説の下での最尤推定量しか用いず, 本論文の検定統計量を構成するには有効であることが分かる. 従って, 本論文ではWald型検定統計量を検定統計量として用いる.

5 検定統計量の導出

ここでは固有値・固有ベクトルの漸近的な分散や共分散を求めることにより検定統計量を導出する. 1母集団においても2母集団においても基本的な考え方は同じであるので, まず初めにWald型検定統計量を求めるための一般的な導出法を説明し, 具体的に検定統計量を求めることにする.

いまベクトル $\sqrt{n}(\boldsymbol{h}-\boldsymbol{\eta})$ が平均 $\mathbf{0}$, 分散共分散行列 Θ の p 変量正規分布に漸近的に従っているとする. このときベクトル $\sqrt{n}\Theta^{-\frac{1}{2}}(\boldsymbol{h}-\boldsymbol{\eta})$ は平均 $\mathbf{0}$, 分散共分散行列 I の p 変量正規分布に漸近的に従う. このことから,

$$\left\{\sqrt{n}\Theta^{-\frac{1}{2}}(\boldsymbol{h}-\boldsymbol{\eta})\right\}'\left\{\sqrt{n}\Theta^{-\frac{1}{2}}(\boldsymbol{h}-\boldsymbol{\eta})\right\}=n(\boldsymbol{h}-\boldsymbol{\eta})'\Theta^{-1}(\boldsymbol{h}-\boldsymbol{\eta})$$

は自由度 $(p-1)$ の χ^2 分布に漸近的に従うことになり, Wald型検定統計量は

$$n(\boldsymbol{h}-\boldsymbol{\eta})'\hat{\Theta}^{-1}(\boldsymbol{h}-\boldsymbol{\eta})$$

となる. つまり, Wald型検定統計量を構成するためにはベクトル $\sqrt{n}(\boldsymbol{h}-\boldsymbol{\eta})$ の漸近的な分散共分散行列が分かれば良いということが分かる.

はじめに1母集団の場合を考える. いま母分散共分散行列 Σ を $\Sigma=\Gamma\Lambda\Gamma'$ と特異値分解し

$$\begin{aligned}\Gamma &= (\boldsymbol{\eta}_1,\ldots,\boldsymbol{\eta}_\alpha,\ldots,\boldsymbol{\eta}_p)\\ \Lambda &= \mathrm{diag}(\lambda_1,\ldots,\lambda_\alpha,\ldots,\lambda_p)\end{aligned}$$

とする. また母集団分布は4次モーメント

$$E[(X_i-\bar{X}_i)(X_j-\bar{X}_j)(X_k-\bar{X}_k)(X_l-\bar{X}_l)]=\kappa_{1111}^{ijkl}$$

まで存在すると仮定する. 標本分散共分散行列 $S=(s_{ij})$ の固有値 l_α, 固有ベクトル $\boldsymbol{h}_\alpha$ のTaylor展開は

$$\begin{aligned}l_\alpha &= \lambda_\alpha+c_{\alpha\alpha}+O_p(n^{-1})\\ \boldsymbol{h}_\alpha &= \boldsymbol{\eta}_\alpha+\Gamma B\boldsymbol{c}_\alpha+O_p(n^{-1})\end{aligned}$$

となる. ただし,

$$B = \alpha \begin{pmatrix} -\frac{1}{\lambda_1-\lambda_\alpha} & & \vdots & & \\ & \ddots & \vdots & & \\ \cdots & \cdots & 0 & & \\ & & & \ddots & \\ & & & & -\frac{1}{\lambda_p-\lambda_\alpha} \end{pmatrix}, \quad \boldsymbol{c}_i = (c_{1i}, c_{2i}, \ldots, c_{ii}, \ldots, c_{pi})', \; c_{ij} = s_{ij} - \delta_{ij}\lambda_i$$

（行列上部の α は第 α 列，左側の α は第 α 行を示す）

とする. 上記の展開式を用いて, 漸近的な分散や共分散を求める. まず初めに固有値の分散・共分散を求める.

$$E\left[\sqrt{n}(l_\alpha - \lambda_\alpha)\sqrt{n}(l_\beta - \lambda_\beta)\right] = E\left[\sqrt{n}c_{\alpha\alpha}\sqrt{n}c_{\beta\beta}\right] = \begin{cases} \kappa_{22}^{\alpha\beta} & (\alpha \neq \beta) \\ \kappa_4^\alpha + 2(\kappa_2^\alpha)^2 & (\alpha = \beta) \end{cases}$$

であり, 楕円母集団を仮定すると,

$$E\left[\sqrt{n}(l_\alpha - \lambda_\alpha)\sqrt{n}(l_\beta - \lambda_\beta)\right] = E\left[\sqrt{n}c_{\alpha\alpha}\sqrt{n}c_{\beta\beta}\right] = \begin{cases} \kappa\lambda_\alpha\lambda_\beta & (\alpha \neq \beta) \\ (3\kappa + 2)\lambda_\alpha^2 & (\alpha = \beta) \end{cases}$$

となる. ここで κ は楕円分布族の kurtosis parameter である. 固有ベクトルの分散・共分散は

$$\begin{aligned} & E\left[\sqrt{n}\,(\boldsymbol{h}_\alpha - \boldsymbol{\eta}_\alpha)\,\sqrt{n}\,(\boldsymbol{h}_\beta - \boldsymbol{\eta}_\beta)'\right] \\ = \; & \Gamma B \begin{pmatrix} \kappa_{1111}^{1\alpha1\beta} & \cdots & \kappa_{1111}^{1\alpha p\beta} \\ \vdots & & \vdots \\ \kappa_{1111}^{p\alpha1\beta} & \cdots & \kappa_{1111}^{p\alpha p\beta} \end{pmatrix} B\Gamma' + \Gamma B \begin{pmatrix} \lambda_1\kappa_{11}^{\alpha\beta} & & \\ & \ddots & \\ & & \lambda_p\kappa_{11}^{\alpha\beta} \end{pmatrix} B\Gamma' + \Gamma B \; \beta \begin{pmatrix} & \vdots & \\ \cdots & \lambda_\alpha\lambda_\beta & \cdots \\ & \vdots & \end{pmatrix} B\Gamma' \\ = \; & \begin{cases} \displaystyle\sum_{i,j\neq p}^{p} \frac{\kappa_{1111}^{i\alpha j\beta}}{(\lambda_i - \lambda_\alpha)(\lambda_j - \lambda_\beta)}\boldsymbol{\eta}_i\boldsymbol{\eta}_j' + \sum_{i\neq\alpha}^{p} \frac{\lambda_i\lambda_\beta}{(\lambda_i - \lambda_\beta)^2}\boldsymbol{\eta}_i\boldsymbol{\eta}_i' & (\alpha \neq \beta) \\ \displaystyle\sum_{i,j\neq p}^{p} \frac{\kappa_{1111}^{i\alpha j\alpha}}{(\lambda_i - \lambda_\alpha)(\lambda_j - \lambda_\alpha)}\boldsymbol{\eta}_i\boldsymbol{\eta}_j' & (\alpha = \beta) \end{cases} \end{aligned}$$

同様にして固有値と固有ベクトルの共分散は

$$E[\sqrt{n}(l_\alpha - \lambda_\alpha) \cdot \sqrt{n}(\boldsymbol{h}_\beta - \boldsymbol{\eta}_\beta)] = -\sum_{i\neq\beta}^{p} \frac{\kappa_{1111}^{i\beta\alpha\alpha}}{\lambda_i - \lambda_\beta}\boldsymbol{\eta}_i$$

であり, 特に楕円母集団を仮定すると $E[\sqrt{n}(l_\alpha - \lambda_\alpha) \cdot \sqrt{n}(\boldsymbol{h}_\beta - \boldsymbol{\eta}_\beta)] = \boldsymbol{0}$ となる.

以上から帰無仮説 H_{01} を検定するためのWald型検定統計量は

$$n\frac{(l_\alpha - \lambda_0)^2}{\hat{\kappa}_4^\alpha + 2(\hat{\kappa}_2^\alpha)^2}$$

ただし,

$$\hat{\kappa}_4^{\alpha} = \frac{1}{N}\sum_{i=1}^{N}(X_{\alpha i}-\bar{X}_{\alpha})^4$$

$$\hat{\kappa}_2^{\alpha} = \frac{1}{N}\sum_{i=1}^{N}(X_{\alpha i}-\bar{X}_{\alpha})^2$$

特に楕円母集団を仮定すると $\hat{\kappa}=\frac{1}{Np(p+2)}\sum_{i=1}^{N}\left\{(\boldsymbol{X}_i-\bar{\boldsymbol{X}})'S^{-1}(\boldsymbol{X}_i-\bar{\boldsymbol{X}})\right\}^2-1$ として

$$n\frac{(l_{\alpha}-\lambda_0)^2}{(3\hat{\kappa}+2)l_{\alpha}^2} \tag{CE1}$$

となる.

帰無仮説 H_{02} を検定するためのWald型検定統計量は

$$n(\boldsymbol{h}_{\alpha}-\boldsymbol{\eta}_0)'\hat{V}^{-1}(\boldsymbol{h}_{\alpha}-\boldsymbol{\eta}_0) \tag{CG2}$$

$$\hat{V} = \sum_{i,j\neq\alpha}^{p}\frac{\hat{\kappa}_{1111}^{i\alpha j\alpha}}{(l_i-l_{\alpha})(l_j-l_{\alpha})}\boldsymbol{h}_i\boldsymbol{h}_i'+\sum_{i\neq\alpha}^{p}\frac{l_il_{\alpha}}{(l_i-l_{\alpha})^2}\boldsymbol{h}_i\boldsymbol{h}_i'$$

$$\hat{\kappa}_{1111}^{i\alpha j\alpha} = \frac{1}{N}\sum_{k=1}^{N}(X_{ik}-\bar{X}_i)(X_{jk}-\bar{X}_j)(X_{\alpha k}-\bar{X}_{\alpha})^2$$

$$\bar{X}_k = \frac{1}{N}\sum_{i=1}^{k}X_{ki}$$

となり, 楕円母集団を仮定した場合には

$$\frac{n}{\hat{\kappa}+1}\left(l_{\alpha}\boldsymbol{\eta}_0'S^{-1}\boldsymbol{\eta}_0+\frac{1}{l_{\alpha}}\boldsymbol{\eta}_0'S\boldsymbol{\eta}_0-2\right) \tag{CE2}$$

となる.

帰無仮説 H_{03} を検定するためのWald型検定統計量は

$$n\begin{pmatrix} l_{\alpha}-\lambda_0 \\ B^{-1}\Gamma'(\boldsymbol{h}_{\beta}-\boldsymbol{\eta}_0)\end{pmatrix}'K^{-1}\begin{pmatrix} l_{\alpha}-\lambda_0 \\ B^{-1}\Gamma'(\boldsymbol{h}_{\beta}-\boldsymbol{\eta}_0)\end{pmatrix} \tag{CG3}$$

$$K=\begin{pmatrix} \hat{\kappa}_4^{\alpha}+2(\hat{\kappa}_2^{\alpha})^2 & \hat{\kappa}_{1111}^{1\beta\alpha\alpha} & \cdots & \hat{\kappa}_{1111}^{p\beta\alpha\alpha} \\ \hat{\kappa}_{1111}^{1\beta\alpha\alpha} & \hat{\kappa}_{1111}^{1\alpha1\beta} & \cdots & \hat{\kappa}_{1111}^{1\alpha p\beta} \\ \vdots & \vdots & & \vdots \\ \hat{\kappa}_{1111}^{p\beta\alpha\alpha} & \hat{\kappa}_{1111}^{p\alpha1\beta} & \cdots & \hat{\kappa}_{1111}^{p\alpha p\beta}\end{pmatrix}.$$

楕円母集団を仮定した場合には

$$n\frac{(l_{\alpha}-\lambda_0)^2}{(3\hat{\kappa}+2)^2l_{\alpha}^4}+\frac{n}{\hat{\kappa}+1}\left(l_{\beta}\boldsymbol{\eta}_0'S^{-1}\boldsymbol{\eta}_0+\frac{1}{l_{\beta}}\boldsymbol{\eta}_0'S\boldsymbol{\eta}_0-2\right). \tag{CE3}$$

次に2母集団の場合を考える. この場合も同様の方法によってWald型検定統計量が導出される.

$$E\left[\sqrt{n_k}\left(\boldsymbol{h}_\alpha^{(g)}-\boldsymbol{\eta}_\alpha^{(g)}\right)\sqrt{n_k}\left(\boldsymbol{h}_\alpha^{(g)}-\boldsymbol{\eta}_\alpha^{(g)}\right)'\right] = \sum_{i,j\neq\alpha}^{p}\frac{\kappa_{1111}^{(g)i\alpha j\alpha}}{(\lambda_i^{(g)}-\lambda_\alpha^{(g)})(\lambda_j^{(g)}-\lambda_\alpha^{(g)})}\boldsymbol{\eta}_i^{(g)}\boldsymbol{\eta}_j^{(g)\prime}\equiv\Phi_g$$

であるので, $r_k=N_k/N, (N=N_1+N_2)$として, それぞれのベクトルの差を考え仮説の下では

$$\sqrt{n_1r_2}\left(\boldsymbol{h}_\alpha^{(1)}-\boldsymbol{\eta}_\alpha^{(1)}\right)-\sqrt{n_2r_1}\left(\boldsymbol{h}_\alpha^{(2)}-\boldsymbol{\eta}_\alpha^{(2)}\right)=\sqrt{nr_1r_2}\left(\boldsymbol{h}_\alpha^{(1)}-\boldsymbol{h}_\alpha^{(2)}\right)$$

は平均 $\mathbf{0}$, 分散共分散行列 $r_2\Phi_1+r_1\Phi_2$ の正規分布に漸近的に従う. したがって, 帰無仮説 H_{05} のWald型検定統計量は

$$nr_1r_2\left(\boldsymbol{h}_\alpha^{(1)}-\boldsymbol{h}_\alpha^{(2)}\right)'\left(r_2\hat{\Phi}_1+r_1\hat{\Phi}_2\right)^{-1}\left(\boldsymbol{h}_\alpha^{(1)}-\boldsymbol{h}_\alpha^{(2)}\right) \tag{CG5}$$

となる.

帰無仮説 H_{04} のを検定するためのWald型検定統計量は

$$\frac{nr_1r_2(l_\alpha^{(1)}-l_\alpha^{(2)})^2}{r_1\left\{\hat{\kappa}^{(2)}{}_4^\alpha+2\left(\hat{\kappa}^{(2)}{}_2^\alpha\right)^2\right\}+r_2\left\{\hat{\kappa}^{(1)}{}_4^\alpha+2\left(\hat{\kappa}^{(1)}{}_2^\alpha\right)^2\right\}}. \tag{CG4}$$

ただし $\boldsymbol{\kappa}^{(g)}$で各母集団の4次モーメントを表すこととする.

帰無仮説 H_{06}のを検定するためのWald型検定統計量は

$$K_g=\begin{pmatrix}\hat{\kappa}^{(g)}{}_4^\alpha+2(\hat{\kappa}^{(g)}{}_2^\alpha)^2 & \hat{\kappa}^{(g)1\beta\alpha\alpha}_{1111} & \cdots & \hat{\kappa}^{(g)p\beta\alpha\alpha}_{1111}\\ \hat{\kappa}^{(g)1\beta\alpha\alpha}_{1111} & \hat{\kappa}^{(g)1\alpha1\beta}_{1111} & \cdots & \hat{\kappa}^{(g)1\alpha p\beta}_{1111}\\ \vdots & \vdots & & \vdots\\ \hat{\kappa}^{(g)p\beta\alpha\alpha}_{1111} & \hat{\kappa}^{(g)p\alpha1\beta}_{1111} & \cdots & \hat{\kappa}^{(g)p\alpha p\beta}_{1111}\end{pmatrix},\quad \hat{B}_g=\begin{pmatrix}-\frac{1}{l_1-l_\alpha} & & \\ & \ddots & \\ & & -\frac{1}{l_p-l_\alpha}\end{pmatrix},$$

$$H_g=(\boldsymbol{h}_1^{(g)},\ldots,\boldsymbol{h}_p^{(g)}),\quad M_g=\begin{pmatrix}1 & \mathbf{0}\\ \mathbf{0}' & H_g\hat{B}_g\end{pmatrix}$$

として,

$$nr_1r_2\begin{pmatrix}l_\alpha^{(1)}-l_\alpha^{(2)}\\ \boldsymbol{h}_\beta^{(1)}-\boldsymbol{h}_\beta^{(2)}\end{pmatrix}'(r_2M_1K_1M_1'+r_1M_2K_2M_2')^{-1}\begin{pmatrix}l_\alpha^{(1)}-l_\alpha^{(2)}\\ \boldsymbol{h}_\beta^{(1)}-\boldsymbol{h}_\beta^{(2)}\end{pmatrix}. \tag{CG6}$$

ここで導出した固有値・固有ベクトルに対する検定問題の検定統計量(CG1),(CE1)(CG4)は自由度1,(CG2),(CE2),(CG5)は自由度$(p-1)$,(CG3),(CE3),(CG6)は漸近的に自由度 p の χ^2分布に従う. 標本数が大きい場合には, これらの結果を使って検定を行うことができる.

6 今後の課題

ここで提案された検定統計量は漸近的に χ^2分布に従うが, χ^2分布への収束が早くなければ

使うことができず, その収束やパラメータ推定などの問題も残っている. また他に提案されている検定統計量との検出力の比較も必要である.

前章に書いた検定法はパラメトリックな方法による検定だが, 収束の問題も残されておりノンパラメトリックな検定法についても研究の必要がある.

参考文献

[1] Anderson, T. W. (1963). Asymptotic theory for principal component analysis, *Ann. Math. Statist.*, **34**, 122-148.

[2] Flury, B. (1987). Two generalizations of the common principal component model, *Biometrika*, **74**, 59-69.

[3] Hayakawa, T. (1978). The asymptotic expansion of the distribution of Anderson's statistic for testing a latent vector of a covariance matrix, *Ann. Inst. Statist. Math.*, Part A **30**, 51-55.

[4] Krzanowski, W.J. (1979). Between-groups comparison of principal components, *J. Amer. Statist. Assoc.*, **74**, 703-707.

[5] Mallows, C.L. (1961). Latent vectors of random symmetric matrices, *Biometrika*, **48**, 133-149.

[6] Nagao, H. (1973). On some test criteria for covariance matrix, *Ann. Statist.*, **1**, 700-709.

[7] Schott, J.R. (1987). An Improved chi-squared test for a principal component, *Statist. Probab. Lett.*, **5**, 361-365.

[8] Schott, J.R. (1988). Common principal component subspaces in two groups, *Biometrika*, **75**, 229-236.

[9] Sugiura, N. (1973). Derivatives of the characteristic root of a symmetric or a Hermitian matrix with two applications in multivariate analysis, *Commun. Statist.*, **1**, 393-417.

[10] Sugiura, N. (1976). Asymptotic expansions of the distributions of the latent roots and the latent vector of the Wishart and multivariate F matrices, *J. Multi. Anal.*, **6**, No.4, 500-525.

[11] Sugiyama, T. (1965). On the distribution of the latent vectors for principal component analysis, *Ann. Math. Statist.*, **36**, 1875-1876.

[12] Sugiyama, T. (1971). Tables of percentile points of a vector in principal component analysis, *J. Jpn. Statist. Soc.*, **1**, No.2, 63-68.

[13] Tsukada, S. (1997). Power comparison of hypothesis testing for an intermediate latent vector of covariance matrix, *J. Jpn. Soc. Computa. Statist.*, **10**, 73-88.

[14] Ushizawa, K., Sato, Y. & Sugiyama, T. (1998). Nonparametric test for equality of intermediate latent roots in non-normal distribution, *J. Jpn. Soc. Computa. Statist.*, **11**, 9-23.

[15] 早川毅(1996). 固有ベクトルと固有根の検定について, 平成8年度科研費シンポジウム「多変量解析の理論とその応用」, 137-144.

[16] 塚田真一・杉山高一(1997). 分散共分散行列の固有ベクトルの検定に関する3つの統計量の漸近帰無分布と検出力の比較, 計算機統計学, **10**, 19-35.

[17] 塚田真一・小野英夫・杉山高一(1998). 分散共分散行列の固有ベクトルに関する2標本問題の検出力, 第66回日本統計学会, 188-189.

[18] 塚田真一・牛沢賢二(1999). Power comparison of Flury criterion, Schott criterion and Wald criterion on the test of several latent vectors, 第67回日本統計学会, 363-364.

[19] 塚田真一・尾形唱子(2000). 非正規母集団での分散共分散行列の固有ベクトルの検定について, 第68回日本統計学会, 384-385.

[20] 牛沢賢二(1998). 主成分分析における固有ベクトルに関するノンパラメトリック検定法, 計算機統計学, **11**, 77-87.