

Wishart 行列の固有ベクトルの分布の漸近展開

六島大 理 杉浦 成昭

1つ又は2つの Wishart 行列の固有ベクトルの分布の漸近展開を *Perturbation method* により求める。ただし対応する母集団での固有値は単根とする。固有ベクトルの *normalization* の仕方により漸近分布の *order* が変わることも示される。このことは標本共分散行列から主成分分析等を行うときに注意する必要があるように思われる。

§1. 一標本問題 S を Wishart 分布 $W_p(n, \Gamma)$, $\Gamma = \text{diag}(\lambda_1, \dots, \lambda_p)$ $\lambda_1 \geq \dots \geq \lambda_p > 0$ に従う行列とし, S の α 番目に大きい固有根に対応する固有ベクトルを $f_\alpha = (f_{1\alpha}, \dots, f_{p\alpha})'$ ($f_{\alpha\alpha} > 0$) とおく. $m = n - 2\Delta$ (Δ は補正項) とし $m \rightarrow \infty$ のときを考える. S/m は Γ に確率収束するから f_α の $\frac{1}{m}S = \Gamma$ の周りでの Taylor 展開は λ_α が単根のとき可能であり *Perturbation* により求めることができる. (Sugiura (1974))
 $a_{ij} = \Delta_{ij}/m - \lambda_i \delta_{ij}$ とおくと次のようになる.

$$\begin{aligned}
f_{h\alpha} (h \neq \alpha) = & -\frac{a_{h\alpha}}{\lambda_h - \lambda_\alpha} - \frac{a_{\alpha\alpha} a_{h\alpha}}{(\lambda_h - \lambda_\alpha)^2} + \frac{1}{\lambda_h - \lambda_\alpha} \sum_{l \neq \alpha} \frac{a_{hl} a_{l\alpha}}{\lambda_l - \lambda_\alpha} - \frac{a_{\alpha\alpha}^2 a_{h\alpha}}{(\lambda_h - \lambda_\alpha)^3} \\
& + \frac{a_{\alpha\alpha}}{(\lambda_h - \lambda_\alpha)^2} \sum_{l \neq \alpha} \frac{a_{hl} a_{l\alpha}}{\lambda_l - \lambda_\alpha} + \frac{a_{h\alpha}}{(\lambda_h - \lambda_\alpha)^2} \sum_{l \neq \alpha} \frac{a_{l\alpha}^2}{\lambda_l - \lambda_\alpha} + \frac{a_{\alpha\alpha}}{\lambda_h - \lambda_\alpha} \sum_{l \neq \alpha} \frac{a_{hl} a_{l\alpha}}{(\lambda_l - \lambda_\alpha)^2} \\
& - \frac{1}{\lambda_h - \lambda_\alpha} \sum_{\substack{l \neq \alpha \\ j \neq \alpha}} \frac{a_{hl} a_{lj} a_{j\alpha}}{(\lambda_l - \lambda_\alpha)(\lambda_j - \lambda_\alpha)} + \frac{1}{2} \frac{a_{h\alpha}}{\lambda_h - \lambda_\alpha} \sum_{l \neq \alpha} \frac{a_{l\alpha}^2}{(\lambda_l - \lambda_\alpha)^2} + (\text{高次の項})
\end{aligned} \tag{1.1}$$

$$f_{\alpha\alpha} = 1 - \frac{1}{2} \sum_{l \neq \alpha} \frac{a_{l\alpha}^2}{(\lambda_l - \lambda_\alpha)^2} - \sum_{l \neq \alpha} \frac{a_{\alpha\alpha} a_{l\alpha}^2}{(\lambda_l - \lambda_\alpha)^3} + \sum_{l \neq \alpha} \frac{a_{l\alpha} a_{lj} a_{j\alpha}}{(\lambda_l - \lambda_\alpha)^2 (\lambda_j - \lambda_\alpha)} + (\text{高次の項}).$$

f_α に対し $\|f_\alpha\| = 1$ とする. $a_{h\alpha} (h \neq \alpha)$ は漸近的に $N(0, \frac{\lambda_\alpha \lambda_h}{n})$ に従い, $a_{\alpha\alpha}$ は漸近的に $N(0, \frac{2\lambda_\alpha^2}{n})$ に従うから (1.1) より $\sqrt{n} f_{h\alpha} (h \neq \alpha)$ は漸近的に $N(0, \lambda_\alpha \lambda_h / (\lambda_h - \lambda_\alpha)^2)$ に従うことおよび $\sqrt{n} (f_{\alpha\alpha} - 1) = o_p(1)$ であることがわかる. これらは Anderson (1963) の結果と一致する. 更に $f_{\alpha\alpha}$ については $-2n(f_{\alpha\alpha} - 1)$ の極限分布が独立な χ^2 分布の weighted sum となること, すなわち $\sum_{\substack{l \neq \alpha \\ \alpha}} [\lambda_\alpha \lambda_l / (\lambda_l - \lambda_\alpha)^2] \chi_l^2$ となることがわかる. $e_\alpha = (0, \dots, 0 \mid 0, \dots, 0)$ とおけば (1.1) より $\sqrt{m} (f_\alpha - e_\alpha)$ の特性関数を次のように評価できる.

$$\exp \left\{ -\frac{1}{2} \sum_{l \neq \alpha} \frac{\lambda_\alpha \lambda_l t_l^2}{(\lambda_l - \lambda_\alpha)^2} \right\} \cdot \left[1 + \frac{g}{\sqrt{m}} + \frac{1}{m} \sum_{j=1}^3 i^{2j} h_{2j} + o\left(\frac{1}{m^{3/2}}\right) \right]$$

$$g = -\frac{i}{2} t_\alpha \sum_{l \neq \alpha} \frac{\lambda_\alpha \lambda_l}{(\lambda_l - \lambda_\alpha)^2} - \frac{i^3}{2} t_\alpha \sum_{l \neq \alpha} \frac{\lambda_\alpha^2 \lambda_l^2 t_l^2}{(\lambda_l - \lambda_\alpha)^4},$$

$$h_2 = -\Delta \sum_{l \neq \alpha} \frac{\lambda_\alpha \lambda_l t_l^2}{(\lambda_l - \lambda_\alpha)^2} + \sum_{l \neq \alpha} \frac{\lambda_\alpha \lambda_l t_l^2}{(\lambda_l - \lambda_\alpha)^4} (\lambda_\alpha^2 + \lambda_l^2 - \frac{1}{2} \lambda_\alpha \lambda_l)$$

$$\begin{aligned}
& + \frac{1}{2} \sum_{\substack{l \neq \alpha \\ l \neq \beta}} \frac{\lambda_\alpha \lambda_l \lambda_l t_l^2}{(\lambda_l - \lambda_\alpha)^2 (\lambda_l - \lambda_\alpha)} + \frac{3}{8} t_\alpha^2 \sum_{l \neq \alpha} \frac{\lambda_\alpha^2 \lambda_l^2}{(\lambda_l - \lambda_\alpha)^4} \\
& + \frac{t_\alpha^2}{8} \sum_{\substack{l \neq \alpha \\ l \neq \beta}} \frac{\lambda_\alpha^2 \lambda_l \lambda_l}{(\lambda_l - \lambda_\alpha)^2 (\lambda_l - \lambda_\alpha)^2}, \tag{1.2}
\end{aligned}$$

$$\begin{aligned}
h_4 &= \frac{1}{4} \lambda_\alpha^2 \left\{ \sum_{l \neq \alpha} \frac{\lambda_l t_l^2}{(\lambda_l - \lambda_\alpha)^2} \right\}^2 + \frac{\lambda_\alpha^3}{2} \left\{ \sum_{l \neq \alpha} \frac{\lambda_l t_l^2}{(\lambda_l - \lambda_\alpha)^3} \right\} \left\{ \sum_{l \neq \alpha} \frac{\lambda_l^2 t_l^2}{(\lambda_l - \lambda_\alpha)^3} \right\} \\
& + \frac{\lambda_\alpha^3}{4} t_\alpha^2 \sum_{\substack{l \neq \alpha \\ l \neq \beta}} \frac{\lambda_l \lambda_l^2 t_l^2}{(\lambda_l - \lambda_\alpha)^4 (\lambda_l - \lambda_\alpha)^2} + \frac{t_\alpha^2}{2} \sum_{l \neq \alpha} \frac{\lambda_\alpha^3 \lambda_l^3 t_l^2}{(\lambda_l - \lambda_\alpha)^6},
\end{aligned}$$

$$h_6 = \frac{t_\alpha^2}{8} \left\{ \sum_{l \neq \alpha} \lambda_\alpha^2 \lambda_l^2 t_l^2 / (\lambda_l - \lambda_\alpha)^4 \right\}^2.$$

これより $\sqrt{m}(f_\alpha - e_\alpha)$ は漸近的に平均 0, 共分散行列 $\text{diag}(\frac{\lambda_\alpha \lambda_1}{(\lambda_1 - \lambda_\alpha)^2}, \frac{\lambda_\alpha \lambda_2}{(\lambda_2 - \lambda_\alpha)^2}, \dots, 0, \dots, \frac{\lambda_\alpha \lambda_p}{(\lambda_p - \lambda_\alpha)^2})$ の退化した正規分布に従うことがわかる。特に $t_\alpha = 0$ とおくことにより次の漸近展開を得る。

定理 1.1. λ_α を単根とし $\Phi(x)$ を標準正規分布の分布関数 $\psi_j(x) = \Phi^{(j)}(x) / \Phi(x)$ とおくととき

$$P\left(\bigcap_{l \neq \alpha} \sqrt{m} \frac{|\lambda_l - \lambda_\alpha|}{\sqrt{\lambda_l \lambda_\alpha}} f_{l\alpha} < x_l\right) = \left\{ \prod_{l \neq \alpha} \Phi(x_l) \right\} \left[1 + \frac{1}{m} (h_2 + h_4) + O\left(\frac{1}{m\sqrt{m}}\right) \right]$$

$$\begin{aligned}
h_2 &= -\Delta \sum_{l \neq \alpha} \psi_2(x_l) + \sum_{l \neq \alpha} \frac{\psi_2(x_l)}{(\lambda_l - \lambda_\alpha)^2} (\lambda_\alpha^2 + \lambda_l^2 - \frac{1}{2} \lambda_\alpha \lambda_l) \\
& + \frac{1}{2} \sum_{\substack{l \neq \alpha \\ l \neq \beta}} \frac{\lambda_l}{\lambda_l - \lambda_\alpha} \psi_2(x_l), \tag{1.3}
\end{aligned}$$

$$\begin{aligned}
h_4 &= \frac{1}{4} \sum_{l \neq \alpha} \psi_4(x_l) + \frac{1}{4} \sum_{\substack{l \neq \alpha \\ l \neq \beta}} \psi_2(x_l) \psi_2(x_l) + \frac{1}{2} \sum_{l \neq \alpha} \frac{\lambda_\alpha \lambda_l \psi_4(x_l)}{(\lambda_l - \lambda_\alpha)^2} \\
& + \frac{1}{2} \sum_{\substack{l \neq \alpha \\ l \neq \beta}} \frac{\lambda_l \lambda_\alpha \psi_2(x_l) \psi_2(x_l)}{(\lambda_l - \lambda_\alpha)(\lambda_l - \lambda_\alpha)}.
\end{aligned}$$

4 一般の Wishart 分布

S が $W_p(n, \Sigma)$ に従うときは α 番目に大きい固有根に対応する固有ベクトルを \tilde{f}_α とおけば $H' \Sigma H = \Gamma$ となる直交行列 H を用いて $\tilde{f}_\alpha = H f_\alpha$ とかける. $\tilde{f}_\alpha = (\tilde{f}_{1\alpha}, \dots, \tilde{f}_{p\alpha})'$ とおけば (1.2) より h 番目の成分について次の漸近展開を得る.

定理 1.2. $H = (h_{ij})$, $\tau^2 = \sum_{l \neq \alpha} \{ \lambda_\alpha \lambda_l / (\lambda_l - \lambda_\alpha)^2 \} h_{j\alpha}^2$ とおき $\tau > 0$ とする.

$$P(\sqrt{m}(\tilde{f}_{j\alpha} - h_{j\alpha})/\tau < x) = \Phi(x) + \frac{1}{\sqrt{m}} \left\{ \frac{h_{j\alpha}}{2\tau} \Phi^{(1)}(x) \sum_{l \neq \alpha} \frac{\lambda_\alpha \lambda_l}{(\lambda_l - \lambda_\alpha)^2} \right. \\ \left. + \frac{1}{2} h_{j\alpha} \frac{\Phi^{(3)}(x)}{\tau^3} \sum_{l \neq \alpha} \frac{\lambda_\alpha^2 \lambda_l^2 h_{j\alpha}^2}{(\lambda_l - \lambda_\alpha)^4} + \frac{1}{m} \sum_{j=1}^3 h_{2j} \Phi^{(2j)}(x) / \tau^{2j} + O\left(\frac{1}{m^{3/2}}\right) \right\} \quad (1.4)$$

ただし $\tilde{f}_{j\alpha}$ の符号は $h_{j\alpha}$ の符号と同じものにとり h_2, h_4, h_6 は (1.2) 式の h において τ_l のところを $h_{j\alpha}$ ($l=1, 2, \dots, p$) で置き換えたものとする.

$p=2$ のときは Sugiyama (1971) により f_1 の % 点がいくつか計算されているので上の漸近展開の近似のよさを check することができる. $n=50$, $\lambda_1=1.4$, $\lambda_2=0.6$ のとき f_{21} の上側 2.5% 点は $\sin(0.35373)$, $n=50$, $\lambda_1=1.8$, $\lambda_2=0.2$ のとき f_{21} の上側 2.5% 点は $\sin(0.10734)$ であることが Sugiyama (1971) TABLE 2 よりわかる. (1.3) 式より $\Delta = 0$ ととり計算すれば

$$P(f_{21} > x) \quad n=50 \\ \lambda_1 = 1.4, \lambda_2 = 0.6 \quad \lambda_1 = 1.8, \lambda_2 = 0.2$$

$O(1)$ の項	0.0163	0.02168
$O(m^{-1})$ の項	0.0076	0.00326
近似値	0.0239	0.02494
正しい値	0.025	0.025

これらは比較的よい結果を与えているが、 λ_1, λ_2 がより接近すると近似が悪くなる。 λ_1 が単根という仮定より展開の係数分母に $\lambda_l - \lambda_1$ という項が含まれているからである。

さて上の結果は f_α の漸近分布が退化していて $f_{\alpha\alpha}$ の分布が別の形になることを示していた。退化していない極限分布が得られればその方が応用上使いやすいし又主成分分析等で係数の $f_{\alpha\alpha}$ のみ特別扱いをするというのは好ましくないであろう。このような要求を満すには固有ベクトルの長さを1にとらないで α 番目の S の固有根 λ_α を用いて $\|f_\alpha\|^2 = \lambda_\alpha$ と正規化するとよい。この方法はAnderson(1963)にも述べられている。このときTaylor展開は

$$f_{h\alpha} (h \neq \alpha) = -\frac{\sqrt{\lambda_\alpha} a_{h\alpha}}{\lambda_h - \lambda_\alpha} - \frac{(\lambda_h + \lambda_\alpha) a_{\alpha\alpha} a_{h\alpha}}{2\sqrt{\lambda_\alpha} (\lambda_h - \lambda_\alpha)^2} + \frac{\sqrt{\lambda_\alpha}}{\lambda_h - \lambda_\alpha} \sum_{l \neq \alpha} \frac{a_{hl} a_{l\alpha}}{\lambda_l - \lambda_\alpha} + (\text{高次の項}) \quad (1.5)$$

$$f_{\alpha\alpha} = \sqrt{\lambda_\alpha} + \frac{a_{\alpha\alpha}}{2\sqrt{\lambda_\alpha}} - \frac{a_{\alpha\alpha}^2}{8\lambda_\alpha\sqrt{\lambda_\alpha}} - \frac{1}{2\sqrt{\lambda_\alpha}} \sum_{l \neq \alpha} \frac{\lambda_l a_{l\alpha}^2}{(\lambda_l - \lambda_\alpha)^2} + (\text{高次の項})$$

これより $\sqrt{m} (f_\alpha - \sqrt{\lambda_\alpha} e_\alpha)$ の特性関数を次のように評価でき

る。

$$\exp\left\{-\frac{1}{2} \sum_{\ell \neq \alpha} \frac{\lambda_\alpha^2 \lambda_\ell t_\ell^2}{(\lambda_\ell - \lambda_\alpha)^2} - \frac{1}{4} \lambda_\alpha t_\alpha^2\right\} \cdot \left[1 + \frac{1}{\sqrt{m}} (i g_1 + i^3 g_3) + O\left(\frac{1}{m}\right)\right]$$

$$g_1 = \sqrt{\lambda_\alpha} t_\alpha \left\{ \Delta - \frac{1}{4} - \frac{1}{2} \sum_{\ell \neq \alpha} \frac{\lambda_\ell^2}{(\lambda_\ell - \lambda_\alpha)^2} \right\}, \quad (1.6)$$

$$g_3 = \lambda_\alpha \sqrt{\lambda_\alpha} \left\{ \frac{1}{24} t_\alpha^3 + \lambda_\alpha t_\alpha \sum_{\ell \neq \alpha} \frac{\lambda_\ell^2 t_\ell^2}{(\lambda_\ell - \lambda_\alpha)^3} - \frac{1}{2} \lambda_\alpha t_\alpha \sum_{\ell \neq \alpha} \frac{\lambda_\ell^3 t_\ell^2}{(\lambda_\ell - \lambda_\alpha)^4} \right\}$$

これより f_α の極限分布は退化しない正規分布であることおよび
 べきの漸近展開を得る。

定理 1.3. S は $W_p(n, P)$ に従い λ_α は単根, $\|f_\alpha\|^2 = \rho_\alpha$
 とする。

$$P\left(\bigcap_{\ell \neq \alpha} \frac{\sqrt{m} |\lambda_\ell - \lambda_\alpha|}{\sqrt{\lambda_\ell \lambda_\alpha}} f_{\ell\alpha} < x_\ell, \sqrt{\frac{2m}{\lambda_\alpha}} (f_{\alpha\alpha} - \sqrt{\lambda_\alpha}) < x_\alpha\right)$$

$$= \left\{ \prod_{\ell=1}^p \Phi(x_\ell) \right\} \left[1 + \sqrt{\frac{2}{m}} \left\{ \psi_1(x_\alpha) \left(-\Delta + \frac{1}{4} + \frac{1}{2} \sum_{\ell \neq \alpha} \frac{\lambda_\ell^2}{(\lambda_\ell - \lambda_\alpha)^2}\right) \right. \right.$$

$$\left. \left. - \frac{1}{12} \psi_3(x_\alpha) - \psi_1(x_\alpha) \sum_{\ell \neq \alpha} \frac{\lambda_\ell \psi_2(x_\ell)}{\lambda_\ell - \lambda_\alpha} + \frac{1}{2} \psi_1(x_\alpha) \sum_{\ell \neq \alpha} \frac{\lambda_\ell^2 \psi_2(x_\ell)}{(\lambda_\ell - \lambda_\alpha)^2} \right\} \right]$$

$$+ O\left(\frac{1}{m}\right) \quad (1.7)$$

定理 1.4. S は $W_p(n, \Sigma)$ に従い, λ_α は単根 $\|\tilde{f}_\alpha\|^2 = \rho_\alpha$
 で正規化する。

$$P(\sqrt{m}(\tilde{f}_{j\alpha} - h_{j\alpha})/\tau < x) = \Phi(x) - \frac{1}{\sqrt{m}} \left(\frac{g_1}{\tau} \Phi^{(1)}(x) + \frac{g_3}{\tau^3} \Phi^{(3)}(x) \right) + O\left(\frac{1}{m}\right)$$

ただし $\tau^2 = \sum_{\ell \neq \alpha} \lambda_\alpha^2 \lambda_\ell h_{j\ell}^2 / (\lambda_\ell - \lambda_\alpha)^2 + \frac{1}{2} \lambda_\alpha h_{j\alpha}^2$ であり g_1, g_3 は
 (1.6)式において t_ℓ に $h_{j\ell}$ ($\ell=1, 2, \dots, p$) を代入したものとす。

この場合退化しない正規分布を極限分布に持つのは利点で
 あるが, $O(m^{-1})$ の計算は (1.5) より明らかのように Taylor

展開の項が多くなり $\|f_\alpha\|^2 = 1$ の場合より一層面倒となる。又固有値と固有ベクトルの漸近的独立性もくずれてくる。

Sugiyama (1966), Khatri and Pillai (1969) による固有ベクトルの精密分布を基にして本節で求めた漸近分布が求められれば面白いと思われる。

§ 2. 二標本問題 $S_1 = (S_{ij}^{(1)})$, $S_2 = (S_{ij}^{(2)})$ を互いに独立で $W_p(n_1, P)$, $W_p(n_2, I)$ に従う Wishart 行列とする。ただし $P = \text{diag}(\lambda_1, \dots, \lambda_p)$ $\lambda_1 \geq \dots \geq \lambda_p > 0$ とする。 $(S_2/n_2)^{-1}(S_1/n_1)$ の α 番目に大きい固有値 λ_α に対応する固有ベクトルを $f_\alpha = (f_{1\alpha}, \dots, f_{p\alpha})'$ ($f_{\alpha\alpha} > 0$) とする。 $m_i = n_i - 2\Delta_i$ (Δ_i : 補正項) $m_i = p_i m$ ($p_i > 0$) $m = m_1 + m_2$ とおき $m \rightarrow +\infty$ のときを考慮する。 S_1/m_1 は P に S_2/m_2 は I に確率収束することに注意して $a_{ij} = S_{ij}^{(1)}/m_1 - \lambda_i \delta_{ij}$, $b_{ij} = S_{ij}^{(2)}/m_2 - \delta_{ij}$ とおくととき Perturbation により次の Taylor 展開が得られる。(Sugiura (1974)).

$$\begin{aligned}
 f_{r\alpha} (r \neq \alpha) = & -\frac{a_{r\alpha} - \lambda_\alpha b_{r\alpha}}{\lambda_r - \lambda_\alpha} - \frac{(a_{\alpha\alpha} - \lambda_\alpha b_{\alpha\alpha})(a_{r\alpha} - \lambda_r b_{r\alpha})}{(\lambda_r - \lambda_\alpha)^2} \\
 & + \sum_{l \neq \alpha} \frac{(a_{rl} - \lambda_\alpha b_{rl})(a_{l\alpha} - \lambda_\alpha b_{l\alpha})}{(\lambda_r - \lambda_\alpha)(\lambda_l - \lambda_\alpha)} - \frac{(a_{\alpha\alpha} - \lambda_\alpha b_{\alpha\alpha})^2 (a_{r\alpha} - \lambda_r b_{r\alpha})}{(\lambda_r - \lambda_\alpha)^3} \\
 & + \frac{b_{\alpha\alpha}(a_{\alpha\alpha} - \lambda_\alpha b_{\alpha\alpha})(a_{r\alpha} - \lambda_r b_{r\alpha})}{(\lambda_r - \lambda_\alpha)^2} + \frac{a_{\alpha\alpha} - \lambda_\alpha b_{\alpha\alpha}}{(\lambda_r - \lambda_\alpha)^2} \sum_{l \neq \alpha} \frac{(a_{rl} - \lambda_\alpha b_{rl})(a_{l\alpha} - \lambda_\alpha b_{l\alpha})}{\lambda_l - \lambda_\alpha} \\
 & + \frac{a_{r\alpha} - \lambda_r b_{r\alpha}}{(\lambda_r - \lambda_\alpha)^2} \sum_{l \neq \alpha} \frac{(a_{l\alpha} - \lambda_\alpha b_{l\alpha})^2}{\lambda_l - \lambda_\alpha} + \frac{a_{\alpha\alpha} - \lambda_\alpha b_{\alpha\alpha}}{\lambda_r - \lambda_\alpha} \sum_{l \neq \alpha} \frac{(a_{rl} - \lambda_\alpha b_{rl})(a_{l\alpha} - \lambda_l b_{l\alpha})}{(\lambda_l - \lambda_\alpha)^2}
 \end{aligned}$$

$$\begin{aligned}
& - \sum_{\substack{l \neq \alpha \\ j \neq \alpha}} \frac{(a_{rl} - \lambda_\alpha b_{rl})(a_{lj} - \lambda_\alpha b_{lj})(a_{ja} - \lambda_\alpha b_{ja})}{(\lambda_r - \lambda_\alpha)(\lambda_l - \lambda_\alpha)(\lambda_j - \lambda_\alpha)} \quad (2.1) \\
& - \frac{a_{\alpha\alpha} - \lambda_\alpha b_{\alpha\alpha}}{\lambda_r - \lambda_\alpha} \sum_{l \neq \alpha} \frac{b_{ll}(a_{l\alpha} - \lambda_\alpha b_{l\alpha})}{\lambda_l - \lambda_\alpha} + \frac{1}{2} \frac{a_{r\alpha} - \lambda_\alpha b_{r\alpha}}{\lambda_r - \lambda_\alpha} \sum_{l \neq \alpha} \frac{(a_{l\alpha} - \lambda_\alpha b_{l\alpha})^2}{(\lambda_l - \lambda_\alpha)^2} + (\text{高次の項}) \\
f_{\alpha\alpha} &= 1 - \frac{1}{2} \sum_{l \neq \alpha} \frac{(a_{l\alpha} - \lambda_\alpha b_{l\alpha})^2}{(\lambda_l - \lambda_\alpha)^2} - \sum_{l \neq \alpha} \frac{(a_{\alpha\alpha} - \lambda_\alpha b_{\alpha\alpha})(a_{l\alpha} - \lambda_\alpha b_{l\alpha})(a_{l\alpha} - \lambda_l b_{l\alpha})}{(\lambda_l - \lambda_\alpha)^3} \\
& + \sum_{\substack{l \neq \alpha \\ j \neq \alpha}} \frac{(a_{l\alpha} - \lambda_\alpha b_{l\alpha})(a_{lj} - \lambda_\alpha b_{lj})(a_{ja} - \lambda_\alpha b_{ja})}{(\lambda_l - \lambda_\alpha)^2(\lambda_j - \lambda_\alpha)} + (\text{高次の項}).
\end{aligned}$$

これより $\sqrt{n} f_{r\alpha} (r \neq \alpha)$ は漸近的に $N(0, \tilde{\lambda}_r \lambda_\alpha / (\lambda_r - \lambda_\alpha)^2)$, $\tilde{\lambda}_r = \lambda_r / p_1 + \lambda_\alpha / p_2$ に従い $-2n(f_{\alpha\alpha} - 1)$ は漸近的に χ^2 分布の一次結合 $\sum_{l \neq \alpha} \{ \tilde{\lambda}_l \lambda_\alpha / (\lambda_l - \lambda_\alpha)^2 \} \chi_l^2$ に従うことがわかる。ただし $\chi_1^2, \dots, \chi_p^2$ は互いに独立な自由度1の χ^2 分布に従う確率変数とする。 $\sqrt{m}(f_\alpha - e_\alpha)$ の特性関数の漸近的な評価が(2.1)よりでき、その第1項は

$$\exp \left\{ -\frac{1}{2} \sum_{l \neq \alpha} \lambda_\alpha \tilde{\lambda}_l t_l^2 / (\lambda_l - \lambda_\alpha)^2 \right\} \quad (2.2)$$

となる。更に次の漸近展開が得られる。

定理 2.1. S_1, S_2 が互いに独立に $W_p(n_1, I), W_p(n_2, I)$ に従い λ_α は単根とする。

$$\begin{aligned}
P\left(\bigcap_{r \neq \alpha} \frac{\sqrt{m} |\lambda_r - \lambda_\alpha|}{\sqrt{\lambda_\alpha \tilde{\lambda}_r}} f_{r\alpha} < \chi_r\right) &= \left\{ \prod_{l \neq \alpha} \Phi(\chi_l) \right\} \left\{ 1 + \frac{h_2 + h_4}{m} + o\left(\frac{1}{m\sqrt{m}}\right) \right\} \\
h_2 &= -\Delta \sum_{l \neq \alpha} \psi_2(\chi_l) - \left(\frac{1}{p_1^2} - \frac{1}{p_2^2}\right) \sum_{l \neq \alpha} \frac{\lambda_\alpha \lambda_l \psi_2(\chi_l)}{\tilde{\lambda}_l (\lambda_l - \lambda_\alpha)} + \sum_{l \neq \alpha} \frac{\psi_2(\chi_l)}{(\lambda_l - \lambda_\alpha)^2} \\
&\cdot \left(\frac{\lambda_l^2}{p_1} + \frac{2\lambda_\alpha \lambda_l}{p_1 p_2} + \frac{\lambda_\alpha^2}{p_2} - \frac{3}{2} \lambda_\alpha \tilde{\lambda}_l \right) + \frac{1}{2p_1} \sum_{\substack{l \neq \alpha \\ r \neq l}} \frac{\lambda_r \tilde{\lambda}_l \psi_2(\chi_r)}{\tilde{\lambda}_r (\lambda_l - \lambda_\alpha)}, \quad (2.3)
\end{aligned}$$

$$\begin{aligned}
R_4 = & \frac{1}{4} \sum_{l \neq \alpha} \frac{\psi_4(x_l)}{\tilde{\lambda}_l^2} \left(\frac{\lambda_l^2}{P_1^2} + \frac{\lambda_\alpha^2}{P_2^2} \right) + \frac{1}{4} \sum_{\substack{k \neq \alpha \\ l \neq \alpha}} \frac{\psi_2(x_k) \psi_2(x_l)}{\tilde{\lambda}_k \tilde{\lambda}_l} \left(\frac{\lambda_k \lambda_l}{P_1^2} + \frac{\lambda_\alpha^2}{P_2^2} \right) \\
& + \frac{\lambda_\alpha}{2P_2} \left\{ \sum_{l \neq \alpha} \frac{\psi_4(x_l)}{\lambda_l - \lambda_\alpha} + \sum_{\substack{k \neq \alpha \\ l \neq \alpha}} \frac{\psi_2(x_k) \psi_2(x_l)}{\lambda_l - \lambda_\alpha} \right\} + \frac{1}{2P_1 P_2} \sum_{l \neq \alpha} \frac{\lambda_\alpha \lambda_l \psi_4(x_l)}{(\lambda_l - \lambda_\alpha)^2} \\
& + \frac{1}{2P_1 P_2} \sum_{\substack{k \neq \alpha \\ l \neq \alpha}} \frac{\lambda_\alpha \lambda_k \tilde{\lambda}_l \psi_2(x_k) \psi_2(x_l)}{\tilde{\lambda}_k (\lambda_k - \lambda_\alpha) (\lambda_l - \lambda_\alpha)}.
\end{aligned}$$

S_1, S_2 が互いに独立で一般の Wishart 分布 $W_p(n_1, \Sigma_1)$, $W_p(n_2, \Sigma_2)$ に従うときは、適当な正則行列 $P = (p_{ij})$ をとり、 $P' \Sigma_1 P = I$, $P' \Sigma_2 P = I$ とすれば問題は前の場合に帰着される。

このとき $\sqrt{m}(\bar{f}_\alpha - P e_\alpha)$ の特性関数は極限が

$$\exp \left[-\frac{1}{2} \sum_{i,j=1}^p t_i t_j \left\{ \sum_{l \neq \alpha} \frac{p_{il} p_{jl} \lambda_\alpha \tilde{\lambda}_l}{(\lambda_l - \lambda_\alpha)^2} \right\} \right]$$

となり退化した正規分布である。 \bar{f}_α の主成分の分布の漸近展開も一標本のときと同様にできる。又もし極限分布が退化しないようにするためには $\|f_\alpha\|^2 = \lambda_\alpha$ と $(S_2/n_2)^{-1}(S_1/n_1)$ の固有根を用い正規化すればよい。このとき $\sqrt{n} f_{k\alpha}$ $k=1, 2, \dots, p$ は漸近的に互いに独立で $\sqrt{n} f_{k\alpha} (k \neq \alpha)$ は $N(0, \tilde{\lambda}_k \lambda_\alpha^2 / (\lambda_k - \lambda_\alpha)^2)$ に $\sqrt{n}(f_{\alpha\alpha} - \sqrt{\lambda_\alpha})$ は $N(0, \frac{\lambda_\alpha}{2P_1 P_2})$ に従うことが示され $O(m^{-1/2})$ の漸近展開が求められる。(Sugiura (1974))。この他にも二標本問題については $f_\alpha' S_2 f_\alpha = m_2$ で $\|f_\alpha\|$ を正規化することもできる。Anderson (1951), Konishi (1973) 等は二の正規化により固有値, 固有ベクトルの極限分布を求めている。このとき Taylor 展開は

$$f_{h\alpha}(h \neq \alpha) = -\frac{a_{h\alpha} - \lambda_\alpha b_{h\alpha}}{\lambda_h - \lambda_\alpha} + \frac{(a_{\alpha\alpha} - \lambda_\alpha b_{\alpha\alpha})b_{h\alpha}}{\lambda_h - \lambda_\alpha} + \frac{b_{\alpha\alpha}(a_{h\alpha} - \lambda_\alpha b_{h\alpha})}{2(\lambda_h - \lambda_\alpha)}$$

$$- \frac{(a_{\alpha\alpha} - \lambda_\alpha b_{\alpha\alpha})(a_{h\alpha} - \lambda_\alpha b_{h\alpha})}{(\lambda_h - \lambda_\alpha)^2} + \sum_{l \neq \alpha} \frac{(a_{hl} - \lambda_\alpha b_{hl})(a_{l\alpha} - \lambda_\alpha b_{l\alpha})}{(\lambda_l - \lambda_\alpha)(\lambda_h - \lambda_\alpha)} + (\text{高次の項})$$

$$f_{\alpha\alpha} = 1 - \frac{1}{2}b_{\alpha\alpha} + \frac{3}{8}b_{\alpha\alpha}^2 + \sum_{l \neq \alpha} \frac{b_{l\alpha}(a_{l\alpha} - \lambda_\alpha b_{l\alpha})}{\lambda_l - \lambda_\alpha} - \frac{1}{2} \sum_{l \neq \alpha} \frac{(a_{l\alpha} - \lambda_\alpha b_{l\alpha})^2}{(\lambda_l - \lambda_\alpha)^2} + (\text{高次の項})$$

これより次の漸近展開を得る。

定理 2.2. $f'_\alpha S_2 f_\alpha = m_2$ で $\|f_\alpha\|$ を定めるとさ定理 2.1 と同じ条件の下で

$$P\left(\bigcap_{h \neq \alpha} \frac{\sqrt{m} |\lambda_h - \lambda_\alpha|}{\sqrt{\lambda_\alpha \tilde{\lambda}_h}} f_{h\alpha} < \chi_h, \sqrt{2P_2 m} (f_{\alpha\alpha} - 1) < \chi_\alpha\right)$$

$$= \left\{ \prod_{l=1}^p \Phi(\chi_l) \right\} \cdot \left[1 + \sqrt{\frac{2P_2}{m}} \left\{ \psi_1(\chi_\alpha) \left(\Delta - \frac{3}{4P_2} + \frac{1}{2P_2} \sum_{l \neq \alpha} \frac{\lambda_\alpha}{\lambda_l - \lambda_\alpha} + \frac{1}{4} \sum_{l \neq \alpha} \frac{\lambda_\alpha \tilde{\lambda}_l}{(\lambda_l - \lambda_\alpha)^2} \right) \right. \right.$$

$$- \frac{5}{12P_2} \psi_3(\chi_\alpha) - \frac{\psi_1(\chi_\alpha)}{2P_2^2} \sum_{l \neq \alpha} \frac{\lambda_\alpha}{\tilde{\lambda}_l} \psi_2(\chi_l) - \frac{\psi_1(\chi_\alpha)}{2P_2} \sum_{l \neq \alpha} \psi_2(\chi_l)$$

$$\left. \left. + \frac{2\lambda_\alpha}{P_2} \sum_{l \neq \alpha} \frac{\psi_1(\chi_\alpha) \psi_2(\chi_l)}{\lambda_l - \lambda_\alpha} + \frac{1}{2} \sum_{l \neq \alpha} \frac{\lambda_\alpha \tilde{\lambda}_l}{(\lambda_l - \lambda_\alpha)^2} \psi_1(\chi_\alpha) \psi_2(\chi_l) \right\} + O\left(\frac{1}{m}\right) \right].$$

重根のある場合の固有ベクトルの漸近分布に関する Konishi (1973) の結果より少し計算すれば定理 2.2 の第 1 項と同じものが求まる。Konishi (1973) の結果は条件付き一般の場合の極限分布を求めているのでそれを条件なしでの極限分布に直す必要がある。(単根の場合可能)

References

- [1] Anderson, T.W. (1951). The asymptotic distribution

of certain characteristic roots and vectors. Proc. Second Berkeley Symp. on Math. Stat. and Prob. 103-30.

- [2] Anderson, T. W. (1963). Asymptotic theory for principal component analysis. Ann. Math. Statist. 34, 122-48.
- [3] Khatri, C. G. and Pillai, K. C. S. (1969). Distributions of vectors corresponding to the largest roots of three matrices. Multivariate Analysis II Academic Press, N.Y.
- [4] Konishi, S. (1973). 2つの Wishart 行列の固有値固有ベクトルの漸近分布 数理解析研講究録 197.
- [5] Sugiura, N. (1974). Asymptotic expansions of the distributions of the latent roots or latent vector of Wishart matrices. In preparation.
- [6] Sugiyama, T. (1966). On the distribution of the largest latent root and the corresponding latent vector for principal component analysis. Ann. Math. Statist. 37, 995-1001.
- [7] Sugiyama, T. (1971). Tables of percentile points of a vector in principal component analysis. J. Jap. Statist. Soc. 1, 63-68.