因子分析と成分分析のどちらがより安定しているか
―直積構造を持つある斜交解の解析的結果―

小笠原　春彦　　小樽商科大学商学部

１．モデル

因子分析（ＦＡ）と（主）成分分析（ＣＡ）の解の安定性を比較するために、次のような直積構造を持つ斜交因子モデルをとりあげる。

$$\boldsymbol{\Sigma} = \boldsymbol{\Lambda\Phi\Lambda}' + \psi \mathbf{I}_q, \quad \psi > 0,$$

$$\boldsymbol{\Lambda} = \lambda \begin{bmatrix} \mathbf{1}_p & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{1}_p & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{1}_p \end{bmatrix}, \ (\lambda > 0),$$

$$\boldsymbol{\Phi} = \begin{bmatrix} 1 & & \phi \\ & \ddots & \\ \phi & & 1 \end{bmatrix}, \ (0 < \phi < 1)$$

ここで $\mathbf{I}_q$ は q 次の単位行列、$\mathbf{1}_p$ は p 個の 1 からなるベクトルである。パラメータは λ, ϕ, ψ の３個である。因子数を r とすると

$$\boldsymbol{\Phi} = r\phi \frac{\mathbf{1}_r}{\sqrt{r}} \frac{\mathbf{1}_r{}'}{\sqrt{r}} + (1-\phi)\mathbf{I}_r,$$

$$\boldsymbol{\Lambda\Phi\Lambda}' = \boldsymbol{\Phi} \otimes \left(p\lambda^2 \frac{\mathbf{1}_p}{\sqrt{p}} \frac{\mathbf{1}_p{}'}{\sqrt{p}} \right)$$

と書くことができる(q=rp)。

２．パラメータ間の関連

上の結果より $\boldsymbol{\Lambda\Phi\Lambda}'$ の最大固有値は $p\lambda^2(1+(r-1)\phi)$ で、他の０でない固有値は $p\lambda^2(1-\phi)$ である。λ, ϕ に対応するＣＡのパラメータを $\tilde{\lambda}, \tilde{\phi}$ とする。モデルより $\boldsymbol{\Sigma}$ の大きい方から r 個までの固有値は

$$m_1 = p\lambda^2(1+(r-1)\phi) + \psi = p\tilde{\lambda}^2(1+(r-1)\tilde{\phi})$$

$$m_2 = \ldots = m_r = p\lambda^2(1-\phi) + \psi = p\tilde{\lambda}^2(1-\tilde{\phi})$$

である。これらから次の関連が得られる。

$$\tilde{\phi} = \frac{\phi}{1 + \dfrac{\psi}{p\lambda^2}} < \phi, \quad \tilde{\lambda} = \sqrt{\lambda^2 + \frac{\psi}{p}} > \lambda$$

すなわちＣＡでは負荷が大きくなり、相関は小さくなることがわかる。また、$\tilde{\lambda}^2\tilde{\phi} = \lambda^2\phi$ であり、$\boldsymbol{\Sigma}$ の非対角ブロック部分はＦＡとＣＡで同一である。また、相違は p が大になると小さくなる。

標本共分散行列 $\mathbf{S}$ が与えられたときの推定値は ML/ULS 解のいずれも同一になり、

$$\hat{\lambda} = \sqrt{\frac{s_\mathrm{d} - \mathrm{tr}\,\mathbf{S}}{rp(p-1)}}, \quad \hat{\phi} = \frac{(p-1)\, s_\mathrm{off}}{p(r-1)(s_\mathrm{d} - \mathrm{tr}\,\mathbf{S})},$$

$$\hat{\psi} = \frac{\mathrm{tr}\,\mathbf{S}}{rp} - \hat{\lambda}^2,$$

$$\hat{\tilde{\lambda}} = \sqrt{\frac{s_\mathrm{d}}{rp^2}}, \quad \hat{\tilde{\phi}} = \frac{s_\mathrm{off}}{(r-1)s_\mathrm{d}}$$ である。

ここで $s_\mathrm{d} = \sum_{i=1}^{r} \mathbf{1}_p{}'\mathbf{S}_{ii}\mathbf{1}_p$， $s_\mathrm{off} = \sum_{i \neq j} \mathbf{1}_p{}'\mathbf{S}_{ij}\mathbf{1}_p$ であり、 $\mathbf{S}_{ij}$ は $\mathbf{S}$ の(i, j)ブロックである。これらより、母集団の場合と同様なパラメータ間の大小/等値関係が導かれる。

３．負荷の推定量の漸近標準誤差

ＦＡの負荷に関してはデルタ法を用いると

補題１.

$$\mathrm{avar}(\hat{\lambda}^2) = \frac{2}{rN}\left\{ \lambda^4 + \frac{2\lambda^2\psi}{p} + \frac{\psi^2}{p(p-1)} + (r-1)\lambda^4\phi^2 \right\}$$

が得られる。これと $\mathrm{ase}(\hat{\lambda}) = \mathrm{ase}(\hat{\lambda}^2)/(2\lambda)$ から

命題１.

$$\mathrm{ase}(\hat{\lambda}) = \sqrt{\frac{1}{2rN}\left\{ \lambda^2 + \frac{2\psi}{p} + \frac{\psi^2}{\lambda^2 p(p-1)} + (r-1)\lambda^2\phi^2 \right\}}$$

となる。一方、ＣＡについては

補題２.

$$\mathrm{avar}(\hat{\tilde{\lambda}}^2) = \frac{2}{rN}\left\{\lambda^4 + \frac{2\lambda^2\psi}{p} + \frac{\psi^2}{p^2} + (r-1)\lambda^4\phi^2\right\}$$

であり、

命題２.

$$\mathrm{ase}(\hat{\tilde{\lambda}}) = \sqrt{\frac{1}{2rN\left(1+\frac{\psi}{p\lambda^2}\right)}\left\{\lambda^2 + \frac{2\psi}{p} + \frac{\psi^2}{\lambda^2 p^2} + (r-1)\lambda^2\phi^2\right\}}$$

が得られる。以上より、

$\mathrm{ase}(\hat{\lambda}) > \mathrm{ase}(\hat{\tilde{\lambda}})$（注 $\tilde{\lambda}^2 - \lambda^2 = \psi/p$）,

$$\frac{\lambda}{\mathrm{ase}(\hat{\lambda})} < \frac{\tilde{\lambda}}{\mathrm{ase}(\hat{\tilde{\lambda}})},\quad \frac{\hat{\lambda}}{\mathrm{ase}(\hat{\lambda})} < \frac{\hat{\tilde{\lambda}}}{\mathrm{ase}(\hat{\tilde{\lambda}})},$$

$\dfrac{\hat{\lambda}}{\hat{\mathrm{ase}}(\hat{\lambda})} < \dfrac{\hat{\tilde{\lambda}}}{\hat{\mathrm{ase}}(\hat{\tilde{\lambda}})}$ であり、λ=0 または $\tilde{\lambda}$=0 の検定ではＣＡの方が仮説を棄却しやすい。

$p \to \infty$ とすると

$$\mathrm{ase}(\hat{\lambda}) \text{ と } \mathrm{ase}(\hat{\tilde{\lambda}}) \to \lambda\sqrt{\frac{1}{2rN}(1+(r-1)\phi^2)}$$

であり、さらに $r \to \infty$ とすると

$\mathrm{ase}(\hat{\lambda})$ と $\mathrm{ase}(\hat{\tilde{\lambda}}) \to \dfrac{\lambda\phi}{\sqrt{2N}}$ である。

４．相関パラメータの推定量の漸近標準誤差

$\hat{\phi}$ と $\hat{\tilde{\phi}}$ の漸近標準誤差は、デルタ法を用いると $\hat{\lambda}$ と $\hat{\tilde{\lambda}}$ と同様に得られる。しかし、結果は負荷のように簡単にならない。また、大小関係も不定となる。$p \to \infty$ とすると

$\mathrm{avar}(\hat{\phi})$ と $\mathrm{avar}(\hat{\tilde{\phi}})$

$$\to \frac{2}{r(r-1)N}\{1+2(r-2)\phi+(r^2-6r+6)\phi^2 - 2(r-1)(r-2)\phi^3+(r-1)^2\phi^4\}$$

となり、さらに $r \to \infty$ とすると

$$\mathrm{ase}(\hat{\phi}) \text{ と } \mathrm{ase}(\hat{\tilde{\phi}}) \to \sqrt{\frac{2}{N}}\,\phi(1-\phi)$$

となる。この最大値は $\phi = 0.5$ のときで $1/(2\sqrt{2N})$ である。

５．その他

$$\mathrm{ase}(\hat{\psi}) = \sqrt{\frac{2}{r(p-1)N}}\,\psi \quad \text{で} \quad p \to \infty$$

and/or $r \to \infty$ のとき０に近づく。

r=1 は１因子モデルの場合に対応し、$p \to \infty$ のとき $\mathrm{ase}(\hat{\lambda})$ と $\mathrm{ase}(\hat{\tilde{\lambda}}) \to \dfrac{\lambda}{\sqrt{2N}}$ となる。この極限値は $\mathrm{ase}(\sqrt{s_{ij}}),\ (i \neq j; i, j = 1, \dots, p)$ と同一である。

６．シミュレーション

p=3, 9; r=3, 10; λ=0.4, 0.6, 0.8; ϕ=0.3, 0.7; $\psi = 1-\lambda^2$ のすべての組み合わせで N=300、繰り返し数 1,000 のシミュレーションを行った。この結果、$\hat{\phi}$ の標準誤差は $\hat{\tilde{\phi}}$ のものより大であることが多いことなどが示された（表１－２参照）。

７．結論

採用したモデルの下ではＣＡの方がＦＡより平均的には安定している。この差は p が小さい場合により大きい。この結論は Velicer & Fava (1987)と Velicer & Jackson, (1990, p.106)の主張を支持する。

文献 Velicer, W. F., & Fava, J. L. (1987). *Multivariate Behavioral Research, 22,* 193-209.◆ Velicer, W. F., & Jackson, D. N. (1990). *Multivariate Behavioral Research, 25,* 1-114.

表１・２（略：表の主な著作権が The Psychometric Society にあり Web 上の公開を禁じているため。ただし、著者に請求して著者が個々に配布することは認めているので、希望する場合は著者に請求のこと。E-mail: hogasa@res.otaru-uc.ac.jp）

小笠原　春彦　　小樽商科大学商学部

因子分析と成分分析のどちらがより安定しているか
―直積構造を持つある斜交解の解析的結果―

因子分析と成分分析の解の安定性を比較するために、直積構造を持つ斜交因子モデルをとりあげ、パラメータの推定量の漸近標準誤差を求めた。因子負荷の標準誤差は成分負荷のものよりも大であることが解析的に示された。一方、因子／成分間相関パラメータの標準誤差は、因子分析の方が大であることが多かった。このようなモデルの下では全般的には成分分析の方が安定しているといえる。これらの相違は観測変量数が少ない場合はより大である。