（7）「数理統計学と計量心理学をつなぐ」に関する研究報告

CAUSAL EFFECTS ON THE VARIANCE IN LINEAR STRUCTURAL EQUATION MODELS

大阪大学　　黒木　学

1 はじめに

本稿では，変量間の因果関係およびそのデータ生成過程のそれぞれが非巡回的有向グラフと対応する線形構造方程式モデルで記述できるものとする．このとき，ある処理変量に対して外的操作を行ったときの反応変量の分散への因果的効果を推測する問題を扱う．黒木・宮川 (1999) はバックドア基準を満たす共変量が観測されているとき，分散への因果的効果の定式化をおこなった．本稿では，分散に対する因果的効果の性質について議論する．

2 因果ダイアグラム

非巡回的有向グラフ G は，頂点の集合 $V=\{X_1, X_2, \cdots, X_p\}$ とその直積 $V\times V$ の部分集合である矢線の集合 E によって，$G=(V,E)$ として表現される．因果ダイアグラムとは，変量間の因果関係を非巡回的有向グラフにより記述したものである．本稿では，その因果関係に基づくデータ生成過程を以下の線形構造方程式モデルで与えるものとする．

$$X_i = \sum_{X_j \in pa(X_i)} \alpha_{ij} X_j + \epsilon_i, \quad i=1,2,\ldots,p \tag{1}$$

ここに，$pa(X_i)$ は X_i の親からなる変量集合である．また，誤差変数 $\epsilon_1, \epsilon_2, \cdots, \epsilon_p$ は互いに独立とする．α_{ij} は X_j から X_i への直接的な因果関係の強さを表すものでパス係数と呼ばれる．なお，本稿では，$X_1, \cdots, X_p$ は平均 0，分散 1 に基準化されているものとする．

3 分散に対する因果的効果

本節では，観察研究において X に外的操作を行ったときの Y の分散 $\mathrm{Var}(Y|\mathrm{set}(X=x+\boldsymbol{a}'\boldsymbol{W}))$ を推測する問題を考える．ここに，$\mathrm{set}(X=x+\boldsymbol{a}'\boldsymbol{W})$ は外的操作により X の値を $\boldsymbol{W}$ の値に応じて $x+\boldsymbol{a}'\boldsymbol{W}$ とすることを意味する (Pearl, 2000)．また，$X, \boldsymbol{Y}, \boldsymbol{W}$ は観測変量であり，x と $\boldsymbol{a}$ はそれぞれ定数と定数ベクトルとする．なお，$\boldsymbol{W}$ は X の非子孫からなる変量集合とする．構造方程式モデルの観点からは，$\mathrm{set}(X=x+\boldsymbol{a}'\boldsymbol{W})$ は (1) 式において X に関する構造方程式を $X=x+\boldsymbol{a}'\boldsymbol{W}$ に置き換えることを意味する．

ここで，(1) 式を次のように記述する．

$$\begin{pmatrix} \boldsymbol{Y} \\ X \\ \boldsymbol{T} \end{pmatrix} = \begin{pmatrix} A_{yy} & A_{yx} & A_{yt} \\ \mathbf{0}_{xy} & 0 & A_{xt} \\ \mathbf{0}_{ty} & \mathbf{0}_{tx} & A_{tt} \end{pmatrix} \begin{pmatrix} \boldsymbol{Y} \\ X \\ \boldsymbol{T} \end{pmatrix} + \begin{pmatrix} \boldsymbol{\epsilon}_y \\ \epsilon_x \\ \boldsymbol{\epsilon}_t \end{pmatrix}. \tag{2}$$

ここに，$\boldsymbol{Y}$ と $\boldsymbol{T}$ はそれぞれ X の子孫からなる変量集合と非子孫からなる変量集合であり，$V=\boldsymbol{Y}\cup\{X\}\cup\boldsymbol{T}$，$\boldsymbol{W}\subset\boldsymbol{T}$ を満たすとする．また，$\epsilon_x, \boldsymbol{\epsilon}_y, \boldsymbol{\epsilon}_t$ はそれぞれ $X, \boldsymbol{Y}, \boldsymbol{T}$ に関する誤差変数である．さらに，A_{yt} は $\boldsymbol{T}$ の各要素から $\boldsymbol{Y}$ の各要素へのパス係数からなる行列であり，$\mathbf{0}_{ty}$ は零行列とする．他の行列についても同様に記す．(2) 式において，X に関する構造方程式を $x+\boldsymbol{a}'\boldsymbol{W}$ と置き換えたとき，$\boldsymbol{Y}$ の平均は，

$$E(\boldsymbol{Y}|\mathrm{set}(X=x+\boldsymbol{a}'\boldsymbol{W})) = (I_{yy}-A_{yy})^{-1}A_{yx}x = \boldsymbol{\tau}_{yx}x \tag{3}$$

である．ここに，I_{yy} は単位行列である．また，$\boldsymbol{\tau}_{yx}$ は X から $\boldsymbol{Y}$ の各要素への総合効果からなるベクトルである．

$\boldsymbol{Y}$ の分散は,

$$
\begin{aligned}
Var(\boldsymbol{Y}|\mathrm{set}(X = x + \boldsymbol{a}'\boldsymbol{W})) &= \Sigma_{yy\cdot x} + (\boldsymbol{\tau}_{yx} - B_{yx})(\boldsymbol{\tau}_{yx} - B_{yx})' \\
&+(I_{yy} - A_{yy})^{-1}(A_{yx}\boldsymbol{a}' + A_{yw} + A_{yz}B_{zw})\Sigma_{ww}(A_{yx}\boldsymbol{a}' + A_{yw} + A_{yz}B_{zw})'(I_{yy} - A'_{yy})^{-1} \\
&-(I_{yy} - A'_{yy})^{-1}(A_{yw} + A_{yz}B_{zw})\Sigma_{ww}(A_{yw} + A_{yz}B_{zw})'(I_{yy} - A'_{yy})^{-1} \qquad (4)
\end{aligned}
$$

となる．ここに，B_{zw} は $\boldsymbol{Z} = \boldsymbol{T}\backslash\boldsymbol{W}$ の各要素を目的変数，$\boldsymbol{w}$ を説明変数にした回帰モデルでの $\boldsymbol{w}$ の偏回帰係数からなる行列である．また，Σ_{ww} と $\Sigma_{yy\cdot x}$ はそれぞれ $\boldsymbol{W}$ の共分散行列と X を与えたときの $\boldsymbol{Y}$ の条件付き共分散行列である．他の係数ベクトルおよび相関行列についても同様に記す．

ここで，$A_{yx}\boldsymbol{a}' + A_{yw} + A_{yz}B_{zw} = \mathbf{0}_{yw}$ を満たす $\boldsymbol{a}$ を $\boldsymbol{a}^*$ とおくと，$\mathrm{set}(X = x + \boldsymbol{a}^{*\prime}\boldsymbol{W})$ は $X = x + \boldsymbol{a}'\boldsymbol{W}$ で与えられる外的操作の中で，$\boldsymbol{Y}$ の各要素の分散を最も小さくする外的操作となる．本稿では，$\mathrm{set}(X = x + \boldsymbol{a}^{*\prime}\boldsymbol{W})$ を最適な外的操作と呼ぶ．また，$(I_{yy} - A_{yy})^{-1}(A_{yw} + A_{yz}B_{zw}) = B_{yw} - \boldsymbol{\tau}_{yx}B_{xw}$ より

$$
Var(\boldsymbol{Y}|\mathrm{set}(X = x+\boldsymbol{a}^{*\prime}\boldsymbol{W})) = \Sigma_{yy\cdot x}+(\boldsymbol{\tau}_{yx}-B_{yx})(\boldsymbol{\tau}_{yx}-B_{yx})'-(B_{yw}-\boldsymbol{\tau}_{yx}B_{xw})\Sigma_{ww}(B_{yw}-\boldsymbol{\tau}_{yx}B_{xw})' \qquad (5)
$$

を得る．

(5) 式より，以下のことが考察できる．

(i) (5) 式の 第一項は X を与えたときの $\boldsymbol{Y}$ の条件付き共分散行列，第二項の () 内は X と $\boldsymbol{Y}$ の擬似相関と解釈できる．また，第三項は X を経由する効果を除いた $\boldsymbol{W}$ と $\boldsymbol{Y}$ の関連性の強さを意味している．

(ii) X, $\boldsymbol{Y}$, $\boldsymbol{T}$ が観測変数であることから，(5) 式に現われるパラメータは $\boldsymbol{\tau}_{yx}$ を除いてすべて識別可能である．すなわち，バックドア基準，フロントドア基準 (Pearl, 2000)，条件付き操作変数法 (Brito and Pearl, 2002) に代表される総合効果の識別可能条件は，分散への因果的効果を識別するために利用することができる．

(iii) (2) 式において，X に関する構造方程式を $X = x + \boldsymbol{a}^{*\prime}\boldsymbol{W}$ に置き換えた新たな構造方程式モデルにおいて，$\boldsymbol{W}$ と $\boldsymbol{Y}$ の共分散行列は $\mathbf{0}_{wy}$ となり，parametric cancellation (e.g. Cox and Wermuth, 1998) を引き起こしていることがわかる．したがって，最適な外的操作とは「$\boldsymbol{W}$ と $\boldsymbol{Y}$ の相関を $\mathbf{0}_{wy}$ とする (parametric cancellation を引き起こす) 外的操作」であると解釈することができる (Kuroki et al., 2003).

参考文献

[1] Brito, C. and Pearl, J. (2002). Generalized Instrumental Variables, *Uncertainty in Artificial Intelligence*, **18**, 85-93.

[2] Cox, D. R. and Wermuth, N. (1998). *Multivariate dependencies : models, analysis and interpretation*, Chapman and Hall.

[3] 黒木学・宮川雅巳 (1999). 適応制御における条件付き介入効果の定式化とその推定,「品質」, **29**, 476-486.

[4] Kuroki, M., Miyakawa, M. and Cai, Z. (2003). Joint Causal Effect in Linear Structural Equation Model and Its Application to Process Analysis, *AISTATS 2003*, 70-77.

[5] Pearl, J. (2000). *Causality : Models, Reasoning, and Inference.* Cambridge University Press.

On the Robustness Tuning Parameter in Covariance Selection

宮村 理，狩野 裕
大阪大学大学院人間科学研究科

1. はじめに 本発表では宮村・狩野 (2003) の提案した共分散選択の頑健な推定法において現れる任意パラメータの適当な値を数値実験によって考察し，リアルデータの分析例を通じてその妥当性を確認する．

グラフィカルモデリングは独立性・条件付独立性を利用し，多変量間の関係を整理する探索的モデル構築手法である．解析結果は頂点に確率変数を，辺に条件付独立性の有無を対応させたグラフとして表す．本稿で注目する量的変数についての共分散選択 (Dempster, 1972) は，カテゴリカル変数に対応した対数線形グラフィカルモデルと並んでグラフィカルモデリング実行の基本的なメソッドである．

2. 推定方法の頑健化 先行知見を持たずに探索的なモデル構築を行える点は大きな魅力であるが，最尤推定法を利用した多くの方法と同様に，共分散選択もまた外れ値に敏感である．共分散選択において推定値にバイアスの影響が現れるということは，モデルの誤特定につながる．

Basu et al. (1998) や Eguchi and Kano (2001) はサンプルをその尤度によって重み付けることで頑健化した最尤推定法を提案している．宮村・狩野 (2003) はこれを利用して頑健な共分散要素の推定方法を提案した．

条件付独立を仮定する変数対集合を I とする．

I の補集合を I^c とすれば β-divergence を最小にする $(i,j) \in I^c$ 要素に対応する分散（共分散）の頑健化された推定値は

$$\sigma_{ij} = \frac{\frac{1}{n}\sum_{k=1}^{n} e^{\beta z_k}(y_{ik}-\mu_i)(y_{jk}-\mu_j)}{\frac{1}{n}\sum_{k=1}^{n} e^{\beta z_k} - \beta/(\beta+1)^{1+p/2}} \qquad (1)$$

を満たす．ここで $z_k = -\frac{1}{2}(\boldsymbol{y}_k - \boldsymbol{\mu})^T\Sigma^{-1}(\boldsymbol{y}_k - \boldsymbol{\mu})$ である．β は頑健化の程度を規定することから **robustness tuning parameter** とよばれ，その値が大きいほど推定値は頑健になるが，同時に推定量の分散が大きくなり有効性が落ちる．$\beta = 0$ のとき (1) 式は通常の最尤推定方程式に一致する．

また偏相関係数をゼロと固定する制約をおいた要素 $(u,v) \in I$ の推定方程式は通常の共分散選択と同じ手続きで得られるが，(1) 式によって頑健化された推定値に基づく点が異なる．

3. robustness tuning parameter の選定 robustness tuning parameter β は，任意の非負の値をとることができる．$\beta = 0$ ならば頑健化は施されず，β の増加にともなって推定値は頑健化されるが，同時に推定値の分散も増加する．このことから提案手法の適用時には何らかの方法で適切な β を定める必要がある．β が持つトレードオフの性質から，同程度の頑健化が達成されるならばなるべく小さな値が望ましい．本節では数値実験によって β の値を見積り，ひとつの手がかりとする．

データとしてサンプルサイズ $n = 100$ で 3 変数 X_1, X_2, X_3 の基準化された多変量正規乱数を生成した．このとき外れ値として，X_1 と X_2 の平均をずらしたサンプルを混合比率 5% で混入した．具体的には $(\mu_1, \mu_2) = (0,0), (1,1), (1,2), (1,3), (2,2), (2,3), (3,3)$ の 7 パタンである．各パタンについてそれぞれ 1,000 データセットを作成し，偏相関係数 r^{12} の推定値を求めた．平均二乗誤差のプロットを図 1 に示す．

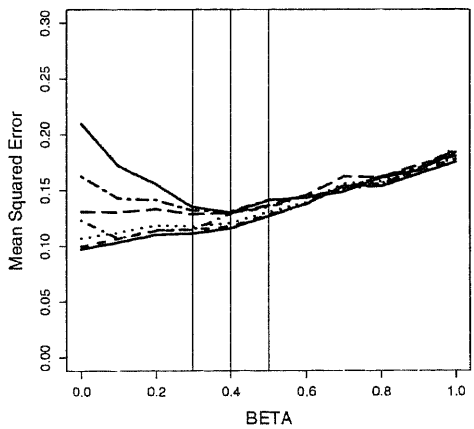


図 1: 偏相関係数 r^{12} の平均二乗誤差

バイアスの調整と分散の増加の影響を併せみる平均二乗誤差は，パタン (3, 3) に顕著に見られるように，$\beta = 0.3$ のあたりでバイアスの調整と分散の増加のトレードオフがもっとも小さくなる．これよりも大きな β になると，分散の増加によるデメリットがバイアス調整のメリットを打ち消してしまうと解釈できる．

以上の結果から多変量正規分布を仮定したデータについては 0.3 から 0.4 の β が適当と考えられる．

4. 親子身長データ　提案手法を用いて，親子身長データのグラフィカルモデリングを行った．

本データは 4 年制大学学生の女性 47 名に本人及び両親の身長を尋ねたものである．

一般に身長は遺伝によって 6 割程が規定され，とりわけ同性の親からの影響が強いと言われている (木村, 1979)．したがって女性であれば母親との結び付きが強いと考えられる．

従来法　まず頑健化を考慮していない通常の共分散選択を行った．標本相関・偏相関行列は以下のようになる．なお，上三角部分を相関係数，下三角部分を偏相関係数として記してある．以降この表現を用いる．

	Daughter	Father	Mother
Daughter	1.0000	0.1377	0.3182
Father	0.1340	1.0000	0.0339
Mother	0.3168	−0.0106	1.0000

(3, 2) の偏相関係数がごく小さく，父親と母親の間には身長についての関係がほとんどないことが伺える．そこでこのパスを切ったモデル，すなわち (3, 2) の偏相関係数をゼロと固定したモデルを作成する．つづいて父親と娘の間のパスを切ったモデル，全ての変数が互いに独立であるモデルと逐次探索した．その結果，独立モデルが逸脱度 5.924 , p 値 0.115 (df = 3) で受容された．このモデルは親子での身長の影響関係を全く反映しておらず，実質科学的に受け入れがたい．

モデルが誤特定された原因として，外れ値の影響を検討するために Mahalanobis 距離 (表 1) を求めたところ，ID 9 のサンプルが外れ値の候補として考えられることがわかった．そのデータをみると，両親，特に母親の身長が比較的高いけれども娘の身長が低く，両親と娘との関係の大きさを過小評価させる一種の外れ値として働いていると考えられる．

表 1: Mahalanobis 距離で昇順ソートしたデータセット（上位下位 5 つを示す）

No	ID	M. Dist.	D	F	M
1	9	10.25	148	168	163
2	5	8.91	170	180	165
3	4	7.28	151	163	163
4	8	6.80	165	156	158
5	18	6.70	155	176	145
⋮	⋮	⋮	⋮	⋮	⋮
43	7	0.38	160	172	158
44	33	0.30	160	172	157
45	23	0.22	159	169	154
46	35	0.17	160	170	158
47	1	0.10	159	170	155

D: Daughter; F: Father; M: Mother

実際，ID 9 を除いた 46 サンプルの標本相関，偏相関行列は

	Daughter	Father	Mother
Daughter	1.0000	0.1396	0.4220
Father	0.1360	1.0000	0.0387
Mother	0.4211	−0.0225	1.0000

であり，偏相関係数で 0.1 程度大きい．

47 サンプルの場合と同様にしてモデル構築を行った場合独立モデルは逸脱度 9.951, p-値 0.018 (df=3) で棄却され，娘－母親モデルが逸脱度 0.928, p-値 0.629 (df=2) で受容された．

頑健法　続いて提案手法を $\beta = 0.3$ で行った．このとき頑健化された標本相関・偏相関行列の推定値は

	Daughter	Father	Mother
Daughter	1.0000	0.0995	0.3992
Father	0.0996	1.0000	0.0204
Mother	0.3993	−0.0212	1.0000

である．頑健化されたことで ID9 の影響を取り除いた推定値を得ることに成功している．

5. おわりに　共分散選択における頑健な推定方法を提案し，このとき用いる robustness tuning parameter β として適当な値を数値実験を通して考察した．数値実験及びリアルデータの解析を通して，提案した手法及び β の有効性を確認した．今後，得られたモデルの評価を行うための適合度指標を開発する必要がある．

相対射影追跡法について

北海道大学大学院　工学研究科　弘　新太郎
北海道大学　情報基盤センター　水田　正弘

1　はじめに

近年，ゲノムデータやPOSデータのような高次元データが増加し，そのような高次元データの解析手法の必要性が増している．一般に，データ解析において，データの構造が高次元構造になるほど，有益な情報をデータ内から抽出することが困難になる．そこで，多変量データ解析では高次元データを解釈が容易な低次元に次元縮小し，有益な情報を引き出す手法が数多く研究されている．特に，Friedman and Tukey(1974)によって提案された射影追跡法 (Projection Pursuit) は，興味深い構造が表れる低次元空間を求めるという意味で，有効な次元縮小の方法である．射影追跡法では興味深さを数値化するための射影指標がいくつか提案されており，そのすべてが興味深い構造を正規分布から最も離れている分布と定義して低次元空間を探索している．

これに対して，解析者が参照とする標本を定義して，その標本の分布からもっとも離れている分布を示す射影方向を探索する新たな射影追跡法を相対射影追跡法をして提案する．相対射影追跡法の利点は，興味のない構造を持つと考えられる標本を予め得ている場合に，その標本の情報を生かして「興味深い」射影方向を検出できることにある．特に，実データの場合には，手元にあるデータの分布が正規分布に従うと仮定できない場合も多い．比較の参照とする標本を解析者が予め得ている状況を想定すれば，相対射影追跡法は有効であると考える．

本報告では，まず，従来の射影追跡法を拡張して，相対射影追跡法を提案すると共に，そのために必要な興味深さを測る相対射影指標もあわせて提案する．また，相対射影追跡法を人工データや実データへ適用し，目的とする特徴的な構造を検出できるかを評価し，その有効性について考察する．

2　相対射影追跡法

対象とするデータに対して参照とする標本を設定し，その標本の分布からの離れ具合を測る新たな射影指標を検討し，数学的な定式化を試みる．この新に提案する射影指標を以後相対射影指標と呼び，この指標を用いる射影追跡法を相対射影追跡法と呼ぶことにする．

相対射影指標の提案

本節では相対射影指標として従来のHallの射影指標の考え方に基づいたHall Type相対射影指標を作成し，提案する．Hallの指標は密度関数間の距離の差の2乗を測っている．従って，解析対象とする標本$X_i, i = 1, \cdots, n$を射影ベクトル$\boldsymbol{\alpha}$で1次元空間に射影したときの密度関数を$f_{\boldsymbol{\alpha}}(x)$とすると，2節で説明したように，1次元のHallの指標は

$$J \equiv \int_{-\infty}^{\infty} \{f_{\boldsymbol{\alpha}}(x) - \phi(x)\}^2 dx$$

と書くことができる．これを相対射影指標に拡張する．参照とする標本を$Y_j, j = 1, \cdots, m$とし，この標本を射影ベクトル$\boldsymbol{\alpha}$で1次元空間に射影したときの密度関数を$g_{\boldsymbol{\alpha}}(x)$とすれば，1次元のHall Type相対射影指標は

$$I(\boldsymbol{\alpha}) = \int_{-\infty}^{\infty} \{f_{\boldsymbol{\alpha}}(x) - g_{\boldsymbol{\alpha}}(x)\}^2 dx$$

と書くことができる．ここで，射影ベクトル$\boldsymbol{\alpha}$で射影したときの密度関数をバンド幅h，カーネル関数を正規分布の密度関数としたカーネル密度推定を用いて計算すると，1次元のHall Type相対射影指標は

$$I(\boldsymbol{\alpha}) = \frac{1}{2\sqrt{\pi}n^2 h_f} \sum_{i=1}^{n}\sum_{j=1}^{n} e^{-\frac{(X_i - X_j)^2}{4h_f^2}} + \frac{1}{2\sqrt{\pi}m^2 h_g} \sum_{i=1}^{m}\sum_{j=1}^{m} e^{-\frac{(Y_i - Y_j)^2}{4h_g^2}}$$

$$- \frac{\sqrt{2}}{\sqrt{\pi}nm\sqrt{{h_f}^2+{h_g}^2}} \sum_{i=1}^{n}\sum_{j=1}^{m} e^{-\frac{(X_i-Y_j)^2}{2({h_f}^2+{h_g}^2)}}$$

となる．

次に，2次元の Hall Type 相対射影指標の場合を考える．解析対象とする標本 $X_i, i=1,\cdots,n$ を射影ベクトル $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2$ で2次元空間に射影したときの密度関数を $f_{\alpha_1,\alpha_2}(x_1,x_2)$，参照とする標本 $Y_j, j=1,\cdots,m$ を射影ベクトル $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2$ で2次元空間に射影したときの密度関数を $g_{\alpha_1,\alpha_2}(x_1,x_2)$ とすれば，2次元の Hall Type 相対射影指標は

$$I(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\{f_{\alpha_1,\alpha_2}(x_1,x_2)-g_{\alpha_1,\alpha_2}(x_1,x_2)\}^2 dx_1 dx_2$$

で表される．ここで，射影ベクトル $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2$ で射影したときの密度関数 f_{α_1,α_2} の推定に用いるバンド幅をそれぞれ h_1, h_2，密度関数 g_{α_1,α_2} の推定に用いるバンド幅をそれぞれ b_1, b_2 とし，カーネル関数を正規分布の密度関数としてカーネル密度推定を行うと，2次元 Hall Type 相対射影指標は以下の式で示される．

$$\begin{aligned} I(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2) &= \frac{1}{4\pi n^2 h_1 h_2}\sum_{i=1}^{n}\sum_{k=1}^{n}\exp\left(-\frac{({\boldsymbol{\alpha}_1}^{\mathrm{T}}X_i-{\boldsymbol{\alpha}_1}^{\mathrm{T}}X_k)^2}{4{h_1}^2}-\frac{({\boldsymbol{\alpha}_2}^{\mathrm{T}}X_i-{\boldsymbol{\alpha}_2}^{\mathrm{T}}X_k)^2}{4{h_2}^2}\right) \\ &+ \frac{1}{4\pi m^2 b_1 b_2}\sum_{j=1}^{m}\sum_{k=1}^{m}\exp\left(-\frac{({\boldsymbol{\alpha}_1}^{\mathrm{T}}Y_j-{\boldsymbol{\alpha}_1}^{\mathrm{T}}Y_k)^2}{4{b_1}^2}-\frac{({\boldsymbol{\alpha}_2}^{\mathrm{T}}Y_j-{\boldsymbol{\alpha}_2}^{\mathrm{T}}Y_k)^2}{4{b_2}^2}\right) \\ &- \frac{1}{\pi nm\sqrt{({h_1}^2+{b_1}^2)({h_2}^2+{b_2}^2)}}\sum_{i=1}^{n}\sum_{j=1}^{m}\exp\left(-\frac{({\boldsymbol{\alpha}_1}^{\mathrm{T}}X_i-{\boldsymbol{\alpha}_1}^{\mathrm{T}}Y_j)^2}{2({h_1}^2+{b_1}^2)}-\frac{({\boldsymbol{\alpha}_2}^{\mathrm{T}}X_i-{\boldsymbol{\alpha}_2}^{\mathrm{T}}Y_j)^2}{2({h_2}^2+{b_2}^2)}\right). \end{aligned}$$

3 数値実験

各変数が混合比率 1：1 の混合正規分布 $N(-1.5,0.5^2)$, $N(1.5,0.5^2)$ に従う 10 変数 $(x_1,...,x_{10})$ のデータを 1000 個用意する．このデータに対して $\sin(x_1)+\cos(x_2)+\varepsilon$; $\varepsilon \sim N(0,0.2^2)$ の値を計算し，$-\frac{2}{3} < \sin(x_1)+\cos(x_2)+\varepsilon < \frac{2}{3}$ を満たす標本のみを取り出す．この取り出した標本に対して，参照とする標本を全てのデータと設定し，本報告で提案する相対射影追跡法を適用する．解析データの標本の分布は，変数 x_1, x_2 で張る空間において参照とする標本の分布とは異なる特徴を持つ．この真の x_1, x_2 で張る空間を検出できるかを評価し，有効性を示すことが本実験の目的であり，その評価として真の射影方向空間との重相関係数の2乗 $R(\cdot)$ を計算する．実験の結果，重相関係数の2乗の値が $R(\alpha_1)=0.998, R(\alpha_2)=0.998$ となり，真の射影方向を検出できた．

この実験のほかに実データへの適用例として AAUP Faculty Salary Data を利用した結果例を紹介し，興味のある低次元射影方向を検出できることを示した．

参考文献

Friedman, J.H.& Tukey, J.W.(1974). A projection pursuit algorithm for exploratory data analysis. *IEEE Trans. on Computer*, c-23, No.9, 881-890.

Hall,P.(1989). On polynomial-based projection indices for exploratory projection pursuit. *The Annals of Statistics*, Vol.17, No.2, 589-605.

Mizuta,M.(2002). Relative projection pursuit. *Data Analysis, Classification, and Related Methods*, (Edited by Andrzej Sokotowski and Krzysztof Jajuga) Cracow University of Economics, 131.

処置前処置後データ分析における回帰効果の再検討

鳥居稔　狩野 裕
大阪大学大学院人間科学研究科

0. はじめに　実験計画において何らかの処置効果を検証するために被験者を実験群と統制群に分け，実験群だけに何らかの処置を施し処置前後の特性値の変化パターンを両群で比較するという方法は文系・理系を問わずあらゆる学問分野で一般的によく用いられている．特に実験群と統制群（対応なし），処置前と処置後（対応あり）という4つの平均値の情報を用いる 2×2 分割法計画が最も典型的であろう．

先行研究によれば，このデザインにおいてはトランケーションなどによって処置前値が偏って観測されている場合は平均への回帰という現象が問題となり，そして，ある仮定のもとでその対処法として共分散分析が有効である．

本発表では処置前-処置後データにおいて回帰効果と処置効果が交絡する問題について再考する．また群間で共変量の回帰直線が異なる場合の共分散分析や，時点数が増えた際の潜在曲線モデルによる拡張についても述べる．

1. 共分散分析による回帰効果の調整　以下に、共分散分析によって回帰効果を調整した上で処置効果を検討することが可能であることを述べる．

処置前後の値を表す変数をX, Yとしそのデータ発生機構を以下のようにする．添え字として実験群であれば (e)，統制群であれば (c) を用いている．

$$Y_{(e)} = \mu_{Y(e)} + \beta_{(e)}(X_e - \mu_{X(e)}) + \epsilon_{(e)}$$
$$Y_{(c)} = \mu_{Y(c)} + \beta_{(c)}(X_c - \mu_{X(c)}) + \epsilon_{(c)}$$

ここで，Fを真の母集団分布とし，仮定i)として「処置前値の（トランケーションのない）全データでの期待値が各群で等しい」という仮定を置くと

$$\mu_X = E_F(X_{(e)}) = E_F(X_{(c)})$$
$$\mu_{Y(e)} = E_F(Y_e),\ \mu_{Y(c)} = E_F(Y_{(c)})$$

と書ける．ここで，実験群における $X = x$ における処置前後差（「$X = x$ における処置効果」+「$X = x$ における練習効果」+「$X = x$ における回帰効果」）は

$$\mu_{Y(c)} + \beta_{(e)}(x - \mu_X) - x$$

また統制群における $X = x$ における処置前後差（「$X = x$ における練習効果」+「$X = x$ における回帰効果」）は

$$\mu_{Y(c)} + \beta_{(c)}(x - \mu_X) - x$$

と書ける．そしてその差（「$X = x$ における処置効果」）は

$$(\mu_{Y(e)} - \mu_{Y(c)}) + (\beta_{(e)} - \beta_{(c)})(x - \mu_X) \quad (1)$$

となり，これが推定目標となる．今，何らかの理由で観測される X に偏りが生じ $X \sim G(\neq F)$ となるデータのみが観察されるとする．このとき，

$$\begin{aligned} E_G(Y_{(e)} - Y_{(c)}) &= \mu_{Y(e)} - \mu_{Y(c)} \\ &\quad +\beta_{(e)}(E_G(X_{(e)}) - \mu_X) \\ &\quad -\beta_{(c)}(E_G(X_{(c)}) - \mu_X) \end{aligned}$$

$$\begin{aligned} \mu_{Y(e)} - \mu_{Y(c)} &= E_G(Y_{(e)} - Y_{(c)}) \\ &\quad -\beta_{(e)}(E_G(X_{(e)}) - \mu_X) \\ &\quad +\beta_{(c)}(E_G(X_{(c)}) - \mu_X) \end{aligned}$$

これを(1)に代入すると

$$\begin{aligned} (1) &= E_G(Y_{(e)} - Y_{(c)}) - \beta_{(e)}(E_G(X_{(e)}) - \mu_X) \\ &\quad +\beta_{(c)}(E_G(X_{(c)}) - \mu_X) \\ &\quad +(\beta_{(e)} - \beta_{(c)})(x - \mu_X) \\ &= [E_G(Y_{(e)}) + \beta_{(e)}(x - E_G(X_{(e)}))] \\ &\quad -[E_G(Y_{(c)}) + \beta_{(c)}(x - E_G(X_{(c)}))] \end{aligned}$$

となる．

このとき，仮定ii)として「処置後値を処置前値に回帰したときの回帰係数が各群で等しい」という仮定が成り立つならば推定目標(1)は$\mu_{Y(e)}-\mu_{Y(c)}$，そしてそのとき

$$\begin{aligned}(1) &= (E_G(Y_{(e)})-E_G(Y_{(c)})) \\ &- \beta(E_G(X_{(e)})-E_G(X_{(c)}))\end{aligned}$$

となり，Xを共変量とした共分散分析によって回帰効果を調整した上で処置効果を検討できる．ただし，「処置後値を処置前値に回帰したときの回帰係数が各群で等しい」という仮定ii)は共分散分析における一般的な仮定であり，この仮定が成り立っていなければ回帰効果を調整できないということを意味しない．

2. 時点数が増えた場合の拡張　処置前-処置後データにおける共分散分析においては，従属変数を処置後値Yにしても処置前後差$Y-x$にしても実は群間差に関する推定結果は変わらない．処置前後差を従属変数とする共分散分析のモデル式は

$$Y-x=a+\beta x+\epsilon$$

と書けるが，左辺のxを移行すると

$$Y=a+(1+\beta)x+\epsilon$$

となり，処置後値Yを従属変数にする場合と比べると共変量の偏回帰係数が1変化するだけで群間差に関しては変化がない．つまり，2時点において処置前後差は傾きを表すので共分散分析は初期値(切片)→傾きをモデリングしていることになる．この，回帰効果を調整するには切片→傾きをモデリングすればよいという視点で考えると，時点数が2時点ではなく3時点，4時点あるいはそれ以上と増え，且つそのデータに何らかの関数が当てはまっている場合には回帰効果を調整する手段として，構造方程式モデリング(Structural Equation Modeling)による潜在曲線モデルが有効であると考えられる．

図1は3時点データで平均構造として直線が当てはまっている場合における潜在曲線モデルの模式図である．「切片」と「傾き」が潜在変数として与えられており，それぞれが「group(群)」

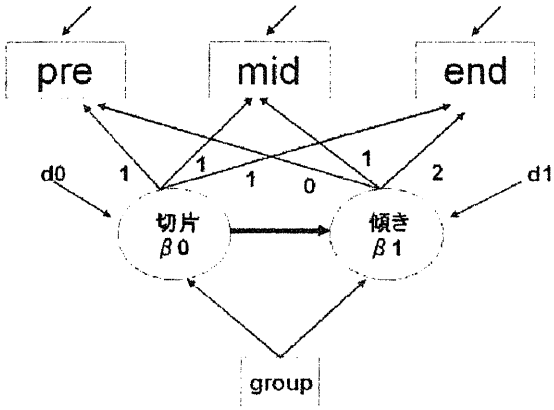


図1: 回帰効果を調整するための潜在曲線モデル

からパスを受けている．「group」はダミー変数であり，統制群を0，実験群を1などとするとパス係数は群間差の平均を表すことになる．

そして，最も重要な点は「切片」から「傾き」にパスを引いているということで，こうすることによって，初期値(切片)が高いと落差(傾き)が大きくなるということ，すなわち回帰効果をモデリングしていることになりそれを調整した上で群間差を検討することができる．群間差は「群」→「傾き」のパスの有意性によって検定することができる．こうしてみると，「傾き」と「切片」と「群」は「傾き」を従属変数とする重回帰モデルになっており，初期値(切片)の影響を取り除いた上で，群間差を検討しているという意味がわかりやすいであろう．

もちろん，曲線が当てはまっている場合は，「切片」から「1次の傾き」だけでなく「高次の傾き」にパスを引くことによって高次の回帰効果を調整することができ，その上で群の効果を検討することができる．

3. まとめ　回帰効果を調整した上で処置効果を検討するには共分散分析が有効であると考えられる．この分析において，最も本質的な仮定は処置前値の(トランケーションのない)全体の期待値が比較したい各群で共通ということである．この仮定は手元にあるデータがトランケーションのあるデータのみの場合検証が不可能である．したがって，この仮定が成り立っているかどうかについては実質科学的な考察が必要であろう．

Multi-indicator growth model for ordinal variable with missing observations

University of Illinois at Chicago Eisuke Segawa

1. Introduction

Growth of a scale score is usually analyzed in two steps, first creating a scale score at each time point by summing items (measurement step), then analyzing its growth over time (growth analysis step). Multi-indicator models for growth (growth of factor models or second order latent growth models) combine these two steps. Using modern SEM software the models can analyze data, with both item- and time-level missing, in a straightforward manner when observed variables are normal. Additionally, they can decompose variance into error and trait and investigate the measurement invariance property.

Despite such advantages, appropriate software is not yet available when data with missing observations are binary or ordinal. In order to analyze such data, we formulated special three-level hierarchical generalized linear models. They are special because they include a factor analytic model for the first and second nested data structures. Our models can analyze, not only data with item- and time-level missing observations, but also data whose time points are freely specified over subjects. Further, we implemented "autoregressive error degree one" structure for the trait residuals. Our approach is Bayesian and a Markov Chain and Monte Carlo computational method is used for its computation. Our sampling strategy is more efficient because we do not sample missing observations. The models can be seen as a combination of existing simpler models with some constraints.

2. Model

We model responses of y_{ijk}, assuming its underlying latent variable, y^{*}_{ijk}, by using the threshold parameter, α_{kc}. Let α_{kc} denote a threshold parameter of the upper limit of category c of item k. Then the model is,

$$y_{ijk} = \begin{cases} 1 \text{ if } -\infty \le y^{*}_{ijk} < \alpha_{k1} \\ 2 \text{ if } \alpha_{k1} \le y^{*}_{ijk} < \alpha_{k2} \\ \vdots \\ C \text{ if } \alpha_{k,C-1} \le y^{*}_{ijk} < \alpha_{kC} \end{cases} \quad (1)$$

,where y_{ijk} denotes a response to the k-th item at the j-th time point in subject i, for $i=1\ldots N$, $j=1\ldots N_i$, and $k=1\ldots K$. The category of item k is denoted as c, $c=1\ldots C$. The variable y^*_{ijk} distributes normally with mean m_{ijk} and variance 1. We assume that the responses are independent, given the vector of probabilities of response of a subject i at j-th time point to item k. A linear relationship between m_{ijk} and θ_{ij}, the trait latent variable of interest, is assumed,

$$m_{ijk} = \beta_{0k} + \beta_{1k}\theta_{ij}, \tag{2}$$

where β_{0k} and β_{1k} are intercept and slope parameters respectively. To identify m_{ijk}, the β_{0k} is fixed at 0. The growth model of θ_{ij} is

$$\theta_{ij} = \mathbf{t}_{ij}\mathbf{u}_i + \varepsilon_{\theta_{ij}}, \tag{3}$$

where $\mathbf{t}_{ij} = \begin{bmatrix} t_{ij}^0 & t_{ij}^1 & \cdots & t_{ij}^P \end{bmatrix}$ and $\mathbf{u}_i^T = \begin{bmatrix} u_{i0} & u_{i1} & \cdots & u_{iP} \end{bmatrix}$. The j-th time point for subject i is t_{ij}. We specify either *i.i.d.* or AR1 for residuals of traits. Subject-level variation is modeled using a normal linear regression,

$$\mathbf{u}_i = \mathbf{W}_i\mathbf{v} + \varepsilon_{\mathbf{u}_i} \text{ and } \varepsilon_{\mathbf{u}_i} \sim N(\mathbf{0}, \Sigma_{\mathbf{u}}), \tag{4}$$

where $\mathbf{W}_i$ is a subject level covariate matrix and ε_{u_i} is independent for all i. Finally, we assume that all disturbance terms, ε^*_{ijk} (error term for y^*_{ijk}), $\varepsilon_{\theta_{ij}}$, and $\varepsilon_{\mathbf{u}_i}$, are mutually independent. We use uniform and conjugate priors for α and other parameters respectively.

3. Data Analysis Result

We analyzed a simulated data set with no missing observation using both our own and Mplus programs. The differences in estimates and standard errors between the two programs are quite small. Estimates from both programs are close to the specified parameter values for the simulated data set as well. Our program also successfully analyzed real data with missing observations.

バランス理論と固有値分解

小杉考司 (関西学院大学社会学研究科)
藤澤隆史 (関西大学総合情報学研究科)

■バランス理論と固有値構造の相同性

Heider(1958) のバランス理論は、3 つの認知対象がどのような関係で結ばれていれば安定的であるか、を図 1 のように定義する。図 1 では、各対象を円で、ポジティブな関係を「+」で、ネガティブな関係を「−」で表現している。Heider の基本的な仮説は、**人はインバランスな関係よりもバランスな関係を好む**、というものである。すなわち、バランスされた状態になければ、人は緊張 (tension) あるいはバランスに向かう力 (force あるいは pressure) を感じるとするのがこの理論の骨子である。さて、図 1 にあるようなグラフ表現は、行列として表現できる。要素間関係を表現した行列に対して、代数的演算の基礎を与えたのは Abelson and Rosenberg(1958) である。彼らは、要素間の関係のセット $\boldsymbol{E}$ を肯定的 (positive、以下 p と略記)、否定的 (negative、同じく n と略記)、両価的 (ambivalent、同じく a)、無関係 (null、同じく o) の四種類とし、この四つの元に対して加法、乗法、交換法則、分配法則などの記号演算を定義した。その後 Phillips(1967) は、Abelson-Rosenberg モデルを代数的演算にまで拡張し、数値演算としての心理 論理体系を確立した。

この実数を用いた心理 論理演算では、対象に対する評価をひとつのベクトル $\boldsymbol{E}$ として表現すれば、バランスの取れた関係とは、評価ベクトル $\boldsymbol{E}$ を乗算した結果得られる正方行列である、ということができる。これを言い換えれば、関係行列 $\boldsymbol{R}$ が与えられたときに、これがひとつのベクトル $\boldsymbol{E}$ による乗算の形に変形できれば、それはバランスが取れているということになる。これは $\boldsymbol{E}' - k\boldsymbol{R}\boldsymbol{E}'$ とする行列 $\boldsymbol{R}$ に対する固有値問題であり、$\boldsymbol{E}$ は固有ベクトル、k は固有値の逆数として算出できる。

さて、ポジティブな関係を +1、ネガティブなものを −1 と表現して、実数の関係行列としてバランス・インバランス各状態を、3×3 の行列で表現すると、バランス状態の固有値は $\{3,0,0\}$、インバランス状態の固有値は $\{2,2,-1\}$ となる。インバランス状態において固有値が三つあるということは、この関係行列を解釈するのに三つのデータ次元が必要ということである。バランス状態は逆に、ひとつの固有値だけであるから、バランス関係を説明するのにはひとつの次元、一組のデータだけでよいことがわかる。すなわち、対象間関係をいちいち全て覚えておく必要はなく、要素それぞれに対する一組の評価と「敵の敵は味方である」というような心理 論理さえ把握してあれば、必然的に関係全体像を復元できる。このことが示すように、バランス状態は、

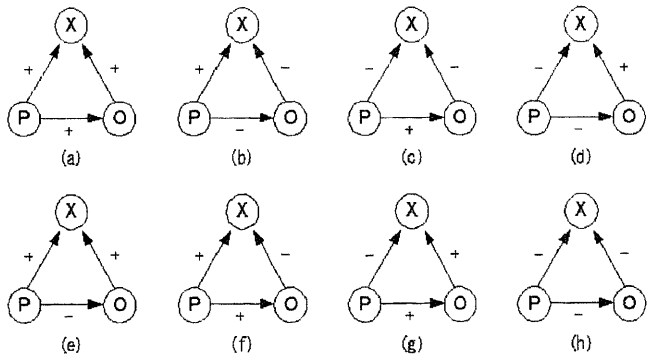


図 1　Heider の POX モデル。P は焦点となる人、O は他者、X は対象となる考えである。(a) から (d) がバランス状態、(e) から (h) がインバランス状態と定義される。

関係全体の情報を圧縮して覚えていられる状態のことである。

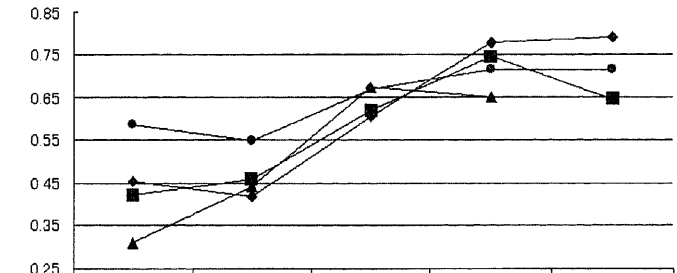
ここから逆に、「人は認知的経済性を求める」という命題をもとにした、より拡張的なバランス理論の定義が導かれるだろう。既に述べたように、バランス状態は全体として「よいもの」として捉えられる。これは、バランス状態にあれば、頭の中で把握しておくべき情報が少なくてすむ、すなわち心理的負荷が少ないからであり、心にとって効率的に外界を把握できる「よい」状態であるとされるからであろう。このような考え方は、認知心理学において特に「認知的経済性」と呼ばれるが、バランス状態はインバランス状態よりも認知的経済性が高いのである。

■人間関係へ応用するときの非対称性の問題

以上の議論から、固有値構造がネットワークにおけるバランスの程度を表現していることが確認された。しかし人間関係においては、この固有値の考え方だけでは対応できない現象がある。すなわち、A さんは B さんのことが好きなのに、B さんは A さんのことが嫌い、という非対称関係であり、人間関係においては非対称行列を固有値分解することを考えなければならない。

そこで、非対称 MDS の一種である、HFM(Hermitian Form Model; 千野,1997) を用いて、非対称行列の分解を考える。HFM では、従来の MDS や因子分析のように、固有値構造が次元数を、固有値に対応する固有ベクトルが対象間の相対的距離関係を表すと解釈できる。もちろんより少ない次元で説明できる方が、情報の圧縮度が高いといえるだろう。

具体的データへの応用例として、縦断的ソシオメトリックデータを HFM で分析し、第一固有値の推移を見たのが図 2 である。寄与率が右肩上がりに推移することから、時系列的に情報が圧縮され、認知的負荷が軽減されていくこと、すなわちバランスが取れた状態になっていくことが読みとれる。



	1st	2nd	3rd	4th	5th
G1	0.45	0.42	0.60	0.78	0.79
G2	0.59	0.55	0.67	0.72	0.72
G3	0.31	0.44	0.67	0.65	
G4	0.42	0.46	0.62	0.74	0.65

図 2　縦断的ソシオメトリックデータを HFM で分解した際の、第一固有値の寄与率の推移

上の定義を用いれば、このように非対称関係にもバランス理論を拡張することが可能である。この他にも、バランス理論を代数的演算の視点から考えると興味深いことは多い。例えば、対角項に当たる「自我」をどのようなものとして扱うかといったことは、大きな課題である。また、HFM による虚数部に統一的な解釈を与えるような、拡張された心理-論理体系の構築が必要であろう。

■引用文献

Abelson,R.P. & Rosenberg,M.J.1958. Symbolic Psyho-Logic:A model of attitudinal cognition. *Behavioral Science*,**3**,1-13. ／千野直仁. 1997. 『非対称多次元尺度構成法』 現代数学社. ／ Heider,F. 1958. *The psychology of interpersonal relations*. New York: John Wiley and Sons. =1978. 大橋正夫 (訳)『対人関係の心理学』誠信書房. ／ Phillips,J.L. 1967. A model for cognitive balance. *Psychological Review*.**74**,481-495.

主成分・正準相関・数量化分析の最小二乗基準と非等質性基準

立命館大学 文学部　足立 浩平

1. はじめに

前半の2節と3節では、記述的な多変量解析法の目的関数を、

非等質性基準;　$LH = \sum \| \underline{\text{個体パラメタ}} - \text{データ} \times \underline{\text{変数パラメタ}} \|^2$　(1)

最小二乗基準;　$LS = \| \text{データ} - \underline{\text{個体パラメタ}} \times \underline{\text{変数パラメタ}} \|^2$　(2)

に大別することによって、主成分分析(PCA)、正準相関分析(CCA)、多重対応分析(MCA)などを見渡す。(1)も最小二乗基準の範疇に分類されることがあるが、(1)と(2)の相違を4節で論じ、非等質性基準と距離との関わりを5節で論じる。なお、(1)をmeet loss、(2)をjoin-lossと呼ぶことがある(Gifi, 1990; Meulman, 1986)。

2. 記号の定義と基準の表現

個体×変数のデータ行列を $\mathbf{X}\,(n\times K) = [\mathbf{X}_1,\ldots,\mathbf{X}_m]$で表す。ここで、$\mathbf{X}_j\,(j = 1,\ldots, m)$は$K_j$個の変数をまとめた$n\times K_j$の行列である。$\mathbf{Q}_j'\mathbf{X}_j'\mathbf{X}_j\mathbf{Q}_j = n\mathbf{I}_{q_j}$となるように$\mathbf{X}_j$を直交化(無相関化)する行列$\mathbf{Q}_j$を考える。求めるべき個体スコアを$\mathbf{F}\,(n\times p)$、変数の成分負荷を$\mathbf{A} = [\mathbf{A}_1',\ldots,\mathbf{A}_m']'$と表す。ただし、$\mathbf{F}'\mathbf{F} = n\mathbf{I}_p$とする。非計量PCA(Gifi, 1990)の一般化のため、de Leeuw & van Rijckevorsel (1988)および村上(1999)は、それぞれ(1), (2)に対応する目的関数

$$LH_p(\mathbf{F},\mathbf{W}) = \sum_{j=1}^{m}\left\|\mathbf{F} - \mathbf{X}_j\mathbf{Q}_j\mathbf{A}_j\right\|^2 = \sum_{j=1}^{m}\left\|\mathbf{F} - \mathbf{X}_j\mathbf{W}_j\right\|^2 \tag{3}$$

$$LS_p(\mathbf{F},\mathbf{A}) = \sum_{j=1}^{m}\left\|\mathbf{X}_j\mathbf{Q}_j - \mathbf{F}\mathbf{A}_j'\right\|^2 = \left\|\mathbf{Z} - \mathbf{F}\mathbf{A}'\right\|^2 \tag{4}$$

を提示している。ここで、$\mathbf{Z} = [\mathbf{X}_1\mathbf{Q}_1,\ldots,\mathbf{X}_m\mathbf{Q}_m]$であり、$\mathbf{W} = [\mathbf{W}_1',\ldots,\mathbf{W}_m']'$は変数へのウェイト$\mathbf{W}_j = \mathbf{Q}_j\mathbf{A}_j$をまとめた行列である。$\mathbf{Q}_j$を未知としても(3), (4)は同等の解を与えるが(村上, 1999)、これが一般化非計量PCAである。以下では、$q_j = K_j$、$\mathbf{Q}_j = n^{1/2}(\mathbf{X}_j'\mathbf{X}_j)^{-1/2}$と限定する。

3. 正準相関・主成分・多重対応分析

(一般化)CCAは、$\mathbf{X}$を列中心化された行列として、$\mathbf{F}'\mathbf{F} = n\mathbf{I}_p$のもとで(3), (4)を最小化する方法と見なせる。SVDによる$n^{-1/2}\mathbf{Z}$の階数p近似を$n^{-1/2}\mathbf{Z} \cong \mathbf{K}_p\mathbf{\Lambda}_p\mathbf{L}_p'$と表すと、$\mathbf{F}, \mathbf{W}, \mathbf{A}$が

$$\hat{\mathbf{F}} = n^{1/2}\mathbf{K}_p,\quad \hat{\mathbf{A}} = \mathbf{L}_p\mathbf{\Lambda}_p,\quad \hat{\mathbf{W}} = \mathbf{Q}\hat{\mathbf{A}} \tag{5}$$

のとき、(3)ならびに(4)は最小化される。ここで、$\mathbf{Q} = \text{blockdiag}(\mathbf{Q}_1,\ldots,\mathbf{Q}_m)$である。(3)の最小化としてのCCAの定式化は柳井(1994)に見られる。

上記の定式化の中で$K_j = 1\,(j = 1,\ldots,m)$とすれば、PCAが導かれる(高根, 1995)。ただし、このPCAは、条件$\mathbf{F}'\mathbf{F} = \mathbf{I}_p$、および、$\mathbf{Z}$が標準得点の行列となることから、相関行列から正規化主成分を求めるPCAである。以上より、PCAをCCAの特殊ケースとする見方もできるが、(4)に着目して、$K_j = 1$と限定しないCCAを「変数の部分集合($\mathbf{X}_j$)内で直交化されたデー

タセット $\mathbf{Z}$ に PCA を適用すること」、さらにいえば、「集合内での直交化の結果、CCA は変数集合間の関係を表す主成分を抽出する」というイメージで CCA を捉えることもできる。

MCA は、$\mathbf{X}_j$ が個体×カテゴリーのダミー変数行列であるときに、(3)または(4)を最小にする個体スコア $\mathbf{F}$ とカテゴリースコア $\mathbf{W}_j$ を求めるものである。ここで、$\mathbf{X}_j'\mathbf{X}_j$ は各カテゴリーの頻度を対角要素とした対角行列になり、素データ $\mathbf{X}_j$ が既に「変数集合内で直交化されている」といえる。なお、MCA では、$\mathbf{X}_j$ は列中心化行列ではない。

4. 非適合度の性質と回転の不定性

LS と LH の重要な違いは、LS は、高次元解の方が非適合度(達成される基準の最小値)が小さい、つまり、$LS_p(\hat{\mathbf{F}},\hat{\mathbf{A}}) \le LS_{p-1}(\hat{\mathbf{F}},\hat{\mathbf{A}})$ であるのに対して、LH は

$$LH_p(\hat{\mathbf{F}},\hat{\mathbf{W}}) \ge LH_{p-1}(\hat{\mathbf{F}},\hat{\mathbf{W}}) \tag{6}$$

となって、通常の統計学の非適合度指標とは逆になる点である。(6)は次のように証明される。まず、$\mathbf{\Lambda}^2$ の第 l 要素(第 l 固有値)を λ_l^2 と表すと、$LH_p(\hat{\mathbf{F}},\hat{\mathbf{W}}) = n(mp - \sum_{l=1}^{p}\lambda_l^2)$ と表せる。ここで $p=1$ とすると $LH_1(\hat{\mathbf{F}},\hat{\mathbf{W}}) = n(m-\lambda_1^2) \ge 0$ が得られ、これが含意する $m \ge \lambda_1^2 \ge \lambda_p^2$ がを用いて、$LH_p(\hat{\mathbf{F}},\hat{\mathbf{W}}) - LH_{p-1}(\hat{\mathbf{F}},\hat{\mathbf{W}}) = n(m-\lambda_p^2) \ge 0$ が示される。

さて、LH, LS ともに直交回転に関する不定性を持つ。さらに、LS については、$\|\mathbf{Z}-\mathbf{M}\|^2$ を最小にする階数 p の行列 $\mathbf{M} = \mathbf{FA}'$ を求めることと(4)を捉え、条件 $\mathbf{F}'\mathbf{F} = n\mathbf{I}_p$ を例えば $\mathrm{diag}(\mathbf{F}'\mathbf{F}) = n\mathbf{I}_p$ のように緩めれば、LS は斜交回転を許容するが、LH は許容しない。

5. 非等質性基準と距離の関わり

$\mathbf{F}$, $\mathbf{W}_j$, $\mathbf{X}_j$ の各行を、それぞれ、$\mathbf{f}_i'$, $\mathbf{w}_{jk}'$, $\mathbf{x}_{ij}'$ と表すと、LH は $\mathbf{f}_i$ と $\mathbf{W}_j'\mathbf{x}_i$ の隔たりを両者間の二乗距離で定義して、$\mathbf{W}_j'\mathbf{x}_j (j = 1,..,m)$ の情報を $\mathbf{f}_i$ に集約させようとするアプローチといえる。MCA に限定すると、$\mathbf{x}_{ij}$ の要素 x_{ijk} は個体 i のカテゴリー k への該当 1・0 で表す 2 値変数であり、(3)は次のように書き換えられる。

$$LH_p(\mathbf{F},\mathbf{W}) = \sum_{i=1}^{n}\sum_{j=1}^{m}\sum_{k=1}^{K_j} x_{ijk}\left\|\mathbf{f}_i - \mathbf{w}_{jk}\right\|^2 . \tag{7}$$

個体のカテゴリーへの該当の有無を両者の類似・非類似に対応づけて、$-x_{ijk}$ を非類似性データと見なせば、(7)は−Σ非類似性×距離を最小にする $\mathbf{f}_i$, $\mathbf{w}_{jk}$ を求める MDS(展開法)の一種(数量化法 4 類, 林, 1993)と見なせ、この見方から $\mathbf{f}_i$ と $\mathbf{w}_{jk}$ の両者を点としてバイプロットすることが容認される。ただし、ポピュラーな MDS では、LS = Σ(非類似性−距離)^2 を目的関数とすることが多く、これは(7)と同等の解を与えない。

文 献

de Leeuw, J., & van Rijckevorsel, J. L. A. (1988). Beyond homogeneity analysis. In J. L. A. van Rijckevorsel & J. de Leeuw (Eds.), Component and correspondence analysis. Chichester, UK: Wiley.

Gifi, A. (1990). *Nonlinear multivariate analysis*. Chichester, UK: Wiley.

林 知己夫 (1993) 数量化. 朝倉書店.

Meulmann, J.J. (1986). *A distance approach to nonlinear multivariate analysis*. Leiden, Netherlands: DSWO Press.

村上 隆 (1999). カテゴリカル・データの主成分分析の心理計量学的研究. 平成 9-10 年度科研費研究成果報告書.

高根芳雄 (1995) 制約つき主成分分析法. 朝倉書店.

柳井晴夫 (1994). 多変量データ解析法. 朝倉書店.

複素力学系による小集団の分析ー計量から理論・予測に向けて

愛知学院大学心身科学部　千野直仁

（問題）

心理学や社会学の分野では、小集団の対人関係に関する縦断的データから対人関係の形成・変容過程を分析する方法は、これまで数多く提案されている。

Chino and Nakagawa (1990) は、縦断的ソシオマトリックスに対して計量心理学的方法、とりわけ MDS を応用し、さらに数学の力学系の定性理論とりわけ力学系の特異点の理論 (singularity theory) と同分岐理論 (bifurcation theory) を利用し、対人相互作用により生ずると考えられる対象相互の心理的距離の変容過程の定性的特徴 (特異点の種類、特異点の分岐のパターン) を、データから推定する方法 DYNASCAL (Dynamical System Scaling) を提案した。

ここでは、まず DYNASCAL の概要とその限界にふれる。つぎに、対人的距離を捉えるためになぜ複素空間が必要かについて述べる。それを説明するためにヤング・ハウスホールダーの定理 (Young & Householder, 1938) と千野・白岩の定理 (Chino & Shiraiwa, 1993) を紹介し、そのうえで複素力学系の理論の紹介とその対人相互作用の分析への応用可能性にふれる。最後に、ここで述べる対人相互作用に関する幾つかのモデルの可能性とパラメータ推定方法の可能性について紹介する。

（DYNASCAL とその限界）

DYNASCAL では、縦断的ソシオマトリックスの各々に対して Ramsay の MULTISCALE を施し、得られた各時点の成員の２次元ユークリッド布置を第２時点から逐次隣接布置のプロクラステス回転により回転していき、回転後の複数時点での布置に対して、２次元非線形微分方程式系を仮定し、１次元及び２次元スプライン関数により解軌道及び矩形格子点上のベクトル（解軌道の一次微分）を推定し、力学系の特異点の理論 (singularity theory) 及び構造安定性 (structural stability) の理論を用いて、各時点の微分方程式系により記述される力の場（ベクトル場）の定性的特徴（ベクトル場の特異点とそこでの基本的な解軌道）を特定しグラフに描く。

その結果、DYNASCAL では小集団の形成・変容過程に関する興味深い結果が得られるが、反面、以下のような短所（限界）もある：

1. プロクラステス回転の理論的根拠に乏しい。
2. 各ソシオマトリックスは非対称にもかかわらず、DYNASCAL では最終的には対象はユークリッド空間に位置づけられる。
3. ２次元非線形微分方程式系を仮定するので、特異点の多くやリミットサイクル (limit cycles) は推定できるが、構造不安定な特異点や分岐の瞬間、カオスは捉えられない。

（Young-Householder の定理及び Chino-Shiraiwa の定理）

紙面の都合上、省略する。

（複素力学系の理論の利用可能性）伝統的な MDS の基礎定理であるヤング・ハウスホールダーの定理と、その非対称 MDS の１つの基礎定理と考えられる千野・白岩定理を比較すれば、ソシオマトリックスなどの非対称データ行列では、１つの簡潔な対象の遠近関係の空間表現のために適した数学的空間が存在し、その１つが（有限次元複素）ヒルベルト空間であることがわかる。

この講演の主題である縦断的ソシオマトリックスを手にして、対象間の心理的遠近関係の形成・変容過程を記述したり予測したりするモデルを考える時、筆者の提案する複素力学系モデルを仮

定すると、複素力学系の理論を利用できる可能性が広がる点で、対象が相互作用をする心理空間（さらには状態空間）として複素空間を考えることのメリットは大きいと考える。

一般の非線形モデルは、

$$\boldsymbol{z}_{j,n+1} = \boldsymbol{z}_{j,n} + \sum_{m=1}^{r} \sum_{k \neq j}^{N} \boldsymbol{D}_{jk,n}^{(m)} \boldsymbol{f}^{(m)}(\boldsymbol{z}_{k,n} - \boldsymbol{z}_{j,n}), \quad j = 1, 2, \cdots, N. \tag{1}$$

モデルの詳細は、Chino (2002, 2003a) 等を参照のこと。

さて、上記モデルによる成員の動きについて、複素力学系の理論を応用する可能性について若干ふれる。例えば、うえの一般の非線形系モデルで、次元数は1、モデルの次数 m は2、かつ成員数は2であるとしよう。この場合、成員1及び成員2の時点 n での座標値を順に z_{1n}、z_{2n} と書けば、式は $z_{1,n+1} = P(z_{1n}) = az_{1n}^2 + c_1 z_{1n} + c_2$ のように書ける。ここで、$a = w_{12,n}^{(2)}$、$c_1 = 1 - w_{12,n}^{(1)} - 2w_{12,n}^{(2)} z_{2n}$、$c_2 = w_{12,n}^{(1)} z_{2n} + w_{12,n}^{(2)} z_{2n}^2$ この式から、もし成員2の位置を固定すれば、成員1の $n+1$ 時点での位置は、成員1自身の n 時点での位置の複素二次関数になっている。そこで、成員2の位置が固定されていると仮定した場合の成員1の動きがどのようになるかを、複素力学系の理論を用いてやれば、成員1の動きの定性的特徴を把握することが出来よう。また、この成員2の位置を変えた場合、成員1の動きはどうなるかについても、同様に複素力学系の理論が使えよう。

成員数を3以上にした場合も、同様な考察を行うことができよう。また、成員数にかかわらず、複素1次元非線形力学系でさえ、成員の振る舞いには実2次元非線形微分力学系では捉えられないカオティックな動きが理論の範疇に入ってくる。上記のような1次元の複素力学系では、固定点や周期点の近傍での系の振る舞いは、写像 $P(z_{1n})$ の一次微分である乗法因子 (the multiplier) により検討できる (Milnor, 2000)。一方、多次元の場合は複素ヤコビ行列を検討することになろう。

（モデルの誤差の取り扱いとパラメータ推定の方法）

紙面の都合上、省略する。詳細は Chino (2003b) を参照のこと。

引用文献

Chino, N. (2002). Complex space models for the analysis of asymmetry. In S. Nishisato, Y. Baba, H. Bozdogan, & K. Kanefuji (Eds.) *Measurement and Multivariate Analysis*, Tokyo: Springer. pp.107-114.

Chino, N. (2003a). Complex difference system models for the analysis of asymetry. In H. Yanai, A. Okada, K. Shigemasu, Y. Kano, & J. J. Meulman (Eds.) *New Developments in Psychometrics*, Tokyo: Springer. pp.479-486.

Chino, N. (2003b). Fitting complex difference system models to longitudinal asymmetric proximity matrices. *Paper presented at the 13th International Meeting of the Psychometric Society*, Cagliari, Italy

Chino, N., & Nakagawa, M. (1990). A bifurcation model of change in group structure. *The Japanese Journal of Experimental Social Psychology*, **29**, No.3, 25-38.

Chino, N., & Shiraiwa, K. (1993). Geometrical structures of some non-distance models for asymmetric MDS. *Behaviormetrika*, **20**, 35-47.

Milnor, J. (2000). *Dynamics in One Complex Variable*. 2nd edition. Wiesbaden:Vieweg & Sohn.

Young, G., & Householder, A. S. (1938). Discussion of a set of points in terms of their mutual distances. *Psychometrika*, **3**, 19-22.

探索的共分散構造分析：多変量データ解析の表と裏

筑波大学大学院経営システム科学専攻
椿　広計
電気通信大学システム工学科
椿　美智子

一般的な多変量データの線形関連性，すなわち相関構造には，因子，すなわち共通原因が潜在化していることに起因する表の構造と観測変数間の因果関係で説明される裏の構造とが混在している．従って，従来の探索的因子分析やグラフィカルモデリングだけでは，この構造を十分単純なモデルで近似することは難しい．そこで因子分析とパス解析の混在した共分散構造モデルをいかにデータ解析的に構築するニーズが生じる．本研究は，この種の構造を探索するための次のようなツールないしはノウハウを紹介する．

1）散布図行列と偏残差の散布図行列（裏の散布図行列）

2）対数固有値プロット（陸地（因子構造）＋大陸棚＋海溝（共線構造））

3）因子分析と偏残差データの因子分析（裏の因子分析）

4）因子モデルや多重指標モデルを飽和化したモデルからのパス減少手順

ここで，偏残差とは，ある変数からそれ以外の全ての変数の影響を回帰により引き去った残差である．

は，分析の前提に関わる直線的関連性や外れ値の非存在を確認するためにデータ解析の初動段階で行う

上記１）は，データに存在する因子数と共線関係数についての第一印象を解析者に与える，

上記２）は，通常の相関係数行列起点の主成分分析のスクリープロットを単に固有値を対数にするだけだなので，因子空間の次元数を決めることができるのは勿論だが，通常の方法よりは共分散構造に対するカイ二乗適合度を反映した方法になっている．さらに，小さな固有値のもつ急減少を観察することで共線構造を探索することができる．特に，理論的厳密性は欠くが，偏残差を起点と

する因子分析を実施し回転を行えば，共線構造についての印象を形成することも可能であり，これが３）で裏の因子分析と呼んだ方法である..

最終的な共分散構造を探索する方法としては，パス解析における手堅い方法としての，完全逐次モデルに対するサイモンブレイロック手順の適用あるいは，グラフィカルモデリングにおける共分散選択のように適切な飽和モデルからのパス消去手順が存在すると便利である．これを意識したのが４）である．通常，探索的因子モデル（斜交回転）を検証的因子モデルに変換し，因子間に因果順序を設定した多重指標モデルを構築すると適合度が著しく劣化したり，不適解が生じることが多い．その際，各因子の測定モデルが適合が悪い（内的適合度）のならば，それを改善すればよいのだが，外的適合度上の問題，すなわち，因子間の構造モデル以外の構造が存在する可能性は否定できない．そこで，観測変数間のパスあるいは観測変数に付随する誤差にパスを形式的に追加することで飽和モデルを構築する．これから，変数減少法を適用し，不適解などが生じる場合には，観測変数と因子の同一視などを通じて，因子構造を因果構造に変換するといった手続きを順じ行い，適合度上問題のない共分散構造モデルを探索するのである．もちろん，この初期飽和モデルの作成の仕方に恣意性があり，現象に対する固有知識の助けが必要となるが，これから導かれるモデルは，データ自身の共線構造と因子構造を反映したものとなる場合が多い．

本研究で示したノウハウは厳密な正当化に耐えるものではなく，対数スクリープロットと最終的に探索された共分散構造との関係性なども不明確なものに過ぎないが，この種の構造探索を通じて共分散構造モデリングで取り扱うモデルが定型的なものからより柔軟な構造になることが期待される．それを通じて，GFI などアドホックな適合度規準ではなく，適合度カイ二乗にもとづく判断（検定にこだわる必要がないが）が復権するのではないかと考えている．

参考文献

椿　広計(2002)狩野論文へのコメント－「尺度化＋回帰分析」の問題点に関する注意、行動計量学、Vol.29, No.2,pp.167-173.

椿　広計、椿美智子(1997)グラフィカルモデリングからの既成モデルの見直し，日本統計学会第 65 回発表要旨集 256－257.

"Connecting mathematical statistics and psychometrics" Conference supported by Grant-in-Aid for Scientific Research (A) No.14204006 hosted by Y. Kano and N. Chino, November 3-5, 2003. Osaka University, Suita, Japan

Asymptotic robustness of the asymptotic biases for some structural models

Haruhiko Ogasawara
Otaru University of Commerce

1. Assumptions and notation

Let

$$\mathbf{x} = \sum_{i=1}^{g} \mathbf{\Lambda}_i \mathbf{f}_i + \boldsymbol{\mu},\ \mathrm{Cov}(\mathbf{f}_i) = \mathbf{\Phi}_i,\ \mathrm{E}(\mathbf{x}) = \boldsymbol{\mu},$$

$$\mathbf{\Sigma} = \mathbf{\Sigma}(\boldsymbol{\theta}) = \sum_{i=1}^{g} \mathbf{\Lambda}_i \mathbf{\Phi}_i \mathbf{\Lambda}_i' \quad (p \times p),$$

$$\mathbf{\Lambda}_i = \mathbf{\Lambda}_i(\boldsymbol{\lambda}),\ \boldsymbol{\theta} = (\boldsymbol{\lambda}', \boldsymbol{\varphi}')' (q \times 1),$$

$$q = l^* + m^*,\ m^* = \sum_{i=1}^{g} m_i^* = \sum_{i=1}^{g} m_i(m_i+1)/2,$$

$$\boldsymbol{\varphi} = (\boldsymbol{\varphi}_1', \ldots, \boldsymbol{\varphi}_g')' = (\mathrm{v}'(\mathbf{\Phi}_1), \ldots, \mathrm{v}'(\mathbf{\Phi}_g))',$$

$\mathbf{s} = \mathrm{v}(\mathbf{S})$ and $\boldsymbol{\sigma} = \boldsymbol{\sigma}(\boldsymbol{\theta}) = \mathrm{v}(\mathbf{\Sigma}(\boldsymbol{\theta}))$.

It is also assumed that possibly nonnormally distributed $\mathbf{f}_i$ and $\mathbf{f}_j, (i \neq j)$ are mutually independent (not just uncorrelated) and that the nonduplicated elements of $\mathbf{\Phi}_i$ are mathematically independent or free parameters.

Then, it is known that

$$n\,\mathrm{acov}(\mathbf{s}) \equiv \mathbf{\Gamma}$$

$$= \mathbf{D}_p^+ \left\{ 2(\mathbf{\Sigma} \otimes \mathbf{\Sigma}) + \sum_{i=1}^{g} (\mathbf{\Lambda}_i \otimes \mathbf{\Lambda}_i) \times \mathbf{D}_{m_i} \mathbf{C}_i^* \mathbf{D}_{m_i}' (\mathbf{\Lambda}_i' \otimes \mathbf{\Lambda}_i') \right\} \mathbf{D}_p^{+\prime}$$

$$= \mathbf{D}_p^+ \left\{ 2(\mathbf{\Sigma} \otimes \mathbf{\Sigma}) + \sum_{i=1}^{g} \mathbf{\Delta}_{\phi_i} \mathbf{C}_i^* \mathbf{\Delta}_{\phi_i}' \right\} \mathbf{D}_p^{+\prime}$$

$$= 2\mathbf{D}_p^+ (\mathbf{\Sigma} \otimes \mathbf{\Sigma}) \mathbf{D}_p^{+\prime} + \mathbf{\Delta}_\phi \mathbf{C}^* \mathbf{\Delta}_\phi',$$

where $\mathbf{C}^*$ is the $m^* \times m^*$ forth-order cumulant matrix of $\mathbf{f}_i$'s.

2. Main result

Let $F_{\mathrm{LS}} = (1/2)(\mathbf{s}-\boldsymbol{\sigma})' \hat{\mathbf{V}} (\mathbf{s}-\boldsymbol{\sigma})$ with $\hat{\mathbf{V}} = \mathbf{V}(\mathbf{s})$. Then, we have

Theorem 1. *The NT asymptotic biases,* $\mathrm{abis}_{\mathrm{NT}}(\hat{\theta}_{\mathrm{LS}i}), (i = 1, \ldots, q)$, *with the associated assumptions are asymptotically robust against the violation of normality.*

Proof. It is known that $\mathbf{\Gamma}_{\mathrm{NT}} = 2\mathbf{D}_p^+ (\mathbf{\Sigma} \otimes \mathbf{\Sigma}) \mathbf{D}_p^{+\prime}$. Using this result and

$$\frac{\partial^2 \theta_i}{\partial \sigma_{ab} \partial \sigma_{cd}} = \frac{\partial (\mathbf{H}^{-1}\mathbf{\Delta}'\mathbf{V})_{i,ab}}{\partial \sigma_{cd}}$$

$$= \left\{ -\mathbf{H}^{-1} \frac{\partial \mathbf{H}}{\partial \sigma_{cd}} \mathbf{H}^{-1}\mathbf{\Delta}'\mathbf{V} + \mathbf{H}^{-1} \left(\frac{\partial \mathbf{\Delta}'}{\partial \sigma_{cd}} \mathbf{V} + \mathbf{\Delta}' \frac{\partial \mathbf{V}}{\partial \sigma_{cd}} \right) \right\}_{i,ab}$$

$$= \left\{ -\mathbf{H}^{-1} \frac{\partial \mathbf{\Delta}'}{\partial \sigma_{cd}} \mathbf{V}\mathbf{\Delta}\mathbf{H}^{-1}\mathbf{\Delta}'\mathbf{V} - \mathbf{H}^{-1}\mathbf{\Delta}' \frac{\partial \mathbf{V}}{\partial \sigma_{cd}} \mathbf{\Delta}\mathbf{H}^{-1}\mathbf{\Delta}'\mathbf{V} \right.$$

$$\left. - \mathbf{H}^{-1}\mathbf{\Delta}'\mathbf{V} \frac{\partial \mathbf{\Delta}}{\partial \sigma_{cd}} \mathbf{H}^{-1}\mathbf{\Delta}'\mathbf{V} + \mathbf{H}^{-1} \frac{\partial \mathbf{\Delta}'}{\partial \sigma_{cd}} \mathbf{V} + \mathbf{H}^{-1}\mathbf{\Delta}' \frac{\partial \mathbf{V}}{\partial \sigma_{cd}} \right\}_{i,ab},$$

$$\equiv (\mathbf{A}_{cd})_{i,ab},\ (i = 1, \ldots, q;\ p \geq a \geq b \geq 1;\ p \geq c \geq d \geq 1),$$

we have after some algebra

$$\mathrm{abis}(\hat{\theta}_{\mathrm{LS}i}) - \mathrm{abis}_{\mathrm{NT}}(\hat{\theta}_{\mathrm{LS}i})$$

$$= \frac{n^{-1}}{2} \sum_{a \geq b} \sum_{c \geq d} (\mathbf{A}_{cd})_{i,ab} (\mathbf{\Gamma} - \mathbf{\Gamma}_{\mathrm{NT}})_{ab,cd}$$

$$= \frac{n^{-1}}{2} \sum_{c \geq d} \left\{ -\mathbf{H}^{-1}\mathbf{\Delta}'\mathbf{V} \frac{\partial \mathbf{\Delta}_\phi}{\partial \sigma_{cd}} \mathbf{C}^* \mathbf{\Delta}_\phi' \right\}_{i,cd}$$

$$= -\frac{n^{-1}}{2} \sum_{e \geq f} (\mathbf{H}^{-1}\mathbf{\Delta}'\mathbf{V})_{i,ef} \times \sum_{k=l^*+1}^{q} \sum_{j=l^*+1}^{q} (\mathbf{C}^*)_{k-l^*, j-l^*} \frac{\partial^2 \sigma_{ef}}{\partial \theta_k \partial \theta_j}$$

$$= 0,\ (i = 1, \ldots, q),$$

where $$\frac{\partial^2 \sigma_{ef}}{\partial \theta_k \partial \theta_j} = \frac{\partial^2 \sigma_{ef}}{\partial \phi_{k-l^*} \partial \phi_{j-l^*}} = 0,$$

$$(k, j = l^*+1, \ldots, q;\ p \geq e \geq f \geq 1)$$

with $\phi_e = (\boldsymbol{\varphi})_e, (e = 1, \ldots, m^*)$ is used. Q.E.D.

3. An illustration

A simulation was carried out using normal and nonnormal data for a factor analysis model. The nonnormal data were randomly generated by independent chi-square distributions for factors. The results are in line with our theory.

Table . Theoretical and simulated biases of the raw-varimax solutions of Emmett's (1949) data (N=211) for unstandardized variables (Number of replications = 1,000,000)

			Bias×10^4							
	Parameter		Normal		(10)		Nonnormal (*df*) (3)		(1)	
		value	Th.	Sim.	Th.	Sim.	Th.	Sim.	Th.	Sim.
	1	.46	-54	-53	-54	-53	-54	-54	-54	-55
	2	.46	-54	-55	-54	-51	-54	-54	-54	-53
	3	.67	-77	-78	-77	-75	-77	-77	-77	-79
	4	.19	-30	-32	-30	-32	-30	-31	-30	-31
Ψ	5	.41	-46	-44	-46	-41	-46	-44	-46	-46
	6	.22	-34	-35	-34	-35	-34	-35	-34	-35
	7	.40	-49	-49	-49	-48	-49	-50	-49	-49
	8	.74	-85	-85	-85	-82	-85	-84	-85	-85
	9	.22	-33	-33	-33	-33	-33	-33	-33	-34
	1	.32	-8	-8	-11	-10	-18	-17	-38	-35
	2	.38	-10	-10	-13	-13	-21	-20	-44	-39
	3	.20	-3	-3	-5	-5	-9	-9	-23	-20
	4	.85	-11	-10	-17	-15	-32	-30	-75	-71
I	5	.74	-10	-13	-16	-21	-29	-31	-66	-65
	6	.85	-9	-10	-16	15	-31	-29	-73	-69
	7	.27	-7	-6	-9	-9	-16	-15	-34	-29
	8	.18	-2	-2	-4	-3	-8	-5	-20	-17
	9	.32	-10	-10	-13	-13	-20	-19	-42	-37
	1	.66	-13	-14	-18	-20	-29	-31	-61	-60
	2	.63	-14	-13	-18	-18	-29	-30	-59	-60
	3	.54	-10	-10	-14	-14	-23	-24	-49	-48
	4	.29	-11	-12	-14	-14	-19	-20	-34	-33
II	5	.21	-6	-4	-7	-5	-11	-9	-22	-21
	6	.23	-9	-8	-10	-12	-15	-15	-27	-26
	7	.73	-13	-13	-18	-18	-30	-31	-65	-64
	8	.47	-8	-9	-12	-11	-20	-21	-43	-41
	9	.82	-14	-14	-20	-19	-34	-35	-74	-71

Note. Th.=theoretical (asymptotic) values, Sim.=simulated values,
I=loadings of Factor I, II=loadings of Factor II.

学習指導要領に準拠した項目プール

統計数理研究所
柳本 武美　三好 美浩

1．問題の設定

受験者がある定められた範囲の知識を有するか否かを判定するのは、教育試験の極く標準的なスタイルである。実際の試験では、試験時間の制約・受験者の負担の限度から、多数の項目を使用するのは不可能で少数の項目に対する反応によって評価するしかない。当然、少数の項目により定められた範囲の知識が判定できるかの疑問が生じる。このとき注意しておくことは、試験の枠組み（制度）としての信頼性と結果としての信頼性とは異なることである。定められた範囲の知識を試していることが確信できる試験の枠組みの設計は、項目プールを作製して、項目プールから出題項目を無作為に抽出するしか方法はない。この主張を、項目プールの役割と位置づけを通じて、学習指導要領を例として考察した。

2．無作為化項目試験

演者らは、今日の新しい試験形態であるＣＢＴには、受験者の能力と項目の特性をより正確に求める統計学的側面と計算機の機能を活かす情報科学的側面とがあるとした(柳本・前田)。前者を無作為化項目試験と呼んだ。そこでは、１）ある範囲を明示的に細分化する、２）多数の項目を作製して、すべての細分化した内容をカバーするように項目プールを構築する、３）項目プールから無作為に選んだ項目を出題する、ことを枠組みとする。この枠組みから直ぐに分かることは、標本調査と酷似している事実である。調査対象を定義して、その名簿を作成して、その名簿から無作為に選んで調査する。もしも、枠組みとして信頼の置ける意見調査を行うならば、それは標本調査でしかない。逆に、無作為抽出を基礎とする標本調査以外に信頼の置ける意見調査の枠組みが構築できるならば、その枠組みから新しい信頼の置ける教育試験法を構築するヒントが容易に得られるに違いない。現状の試験では、試験実施者が少数の出題項目を巧みに選べることを前提としている。

項目プールなしに信頼の置ける試験を実施するためには、出題者の能力と熱意に依存するしかない。しかし、能力のある試験委員が熱意と時間をもってしても、求められる範囲すべてに目を配って出題したかを確かめることはできない。さらには、出題項目を互いに検討しあったり受験者に解いて貰うことができない。従って、制度としての保証は項目プールの構築によってなされる。制度としての保証は、科学的な根拠の質の高さと社会への説明のために欠かせない（柳本、印刷中)。

3．項目プールの不在

学習指導要領に準拠した項目が今日にも存在しないことには多くの理由がある。1つは、指導要領が国の教育基本政策に絡むので、関係者が避けたがることがある。しかし、その効用に対する認識の不足と技術的な側面がある事実を、冷静に指摘する必要がある。項目プールを構築することは大変な労力を要するが、その代替法はないことは前節でも考察したとおりである。もし公教育の充実のための労力を厭うとすれば、それは教育における重大な問題である。ところが、項目プールを構築するための基盤は確実に整備されてきているので、格別な大事業ではない。今日の通信機能向上がもたらす含意は計り知れないが、e-learning で代表されるように、教育充実への期待は大きい。この環境の変化が、項目プールの構築とその利用には大変好都合に働いている。全国の何処でも、家庭ではともかく学校ではこの変化の恩恵を享受できる。更に、項目反応理論を始め、項目プールを支える理論と技法が普及している。より具体的にも項目作製技法の向上と共に、項目プールのより高度な利用技術の急速に進んでいる。項目プールの構築と利用の基盤は整っている。

4．項目プールの効用

項目プールを作ると、意欲ある学習者が何時でも何処でも使用できる。質の高い項目を使って自らの学習到達度を確認できる。また、一組のセットにおける平均困難度を学習者のレベルに合わせられる。このことは、意欲ある学習者に刺激を与えることが出来て、公教育の充実に不可欠である。また、教師サイドでも項目プールを参考にして出題できる。さらに、様々な学習者が実際に解いた結果をデータにして解析することにより、悪問の追放・困難度のより正確な推定が可能になる。標準化が進むので、地域間・年度間の学習到達度の比較が容易になる。これまでの教育界での閉じたマターから、社会に対して説明可能な開かれた科学的証拠に基づいた教育の基盤となる。

5．双方向としての準拠

通常、学習指導要領に準拠しているとは逸脱していないことに主眼がおかれる。項目プールは、指導要領の内容を包含するから、現場からの再検証の機会を与える。項目プールを構築することは、指導要領を細部に亘り具体的に検証することでもある。そして、試験に出ない内容は学習しなくても良い内容となるから、実質的な指導要領となる。だから、学習指導要領と項目プールとの関係では、互いに刺激しあう対話の場になるので、相互の欠点を監視し合う関係になる。その結果、一方の改善が他方の改善となることが期待される。

参考文献

1）柳本 武美「科学的認識論の研究計画への含意― 質の高い証拠を得る要件 」科学哲学, 印刷中. 2）柳本武美, 前田忠彦「無作為化項目試験の枠組とその基盤」投稿中.

計量心理学から数理統計学への提言:多次元解析の枠組み[1]

西里 静彦

カナダ トロント大学 (snishisato@oise.utoronto.ca)

正規分布の母集団から無作為抽出した標本がデータであると言う数理統計学のお膳立ては、分布に関する情報、連続量の計算の利を得るが、同時に線型解析が主流になるという結果を招いているように思われる。

他方、計量心理学では、早くから人間各自の独特の能力、性格、態度などに関心をおき、多次元解析の必要性が認識され、線形解析だけでは不十分で、非線形模型へも関心が示されてきた。この立場は、動機の違いは別としても、統計学、心理学、社会学、生物学、生態学など広い領域における数量化理論、コレスポンデンスアナリシス、双対尺度法に代表される研究領域でもとられている。この立場では、分布、確率の話から離れ、データの説明が出発点となる。

双対尺度法のモデルを採用し、3個の選択肢の一つを選ぶという形で集める多肢選択データを考えると、反応型は(1,0,0), (0,1,0), (0,0,1)、これらを3次元空間の座標と考え、3種の反応を表現できる。被験者からの反応は、それら3点にのみ落ち、3点の位置は、データがあつまった段階で、各点から原点までの距離と各点におちた反応数の積が3点すべてに関して等しくなるように選ばれる。その様にして決められた3点を結ぶと、2次元空間に布置する三角形ができる。いま重心を基点にその3角形を回転した場合、3個の頂点を2本の直行軸に射影したときの3点の射影値の分散は、各50%ずつとなるという興味深い性質がある。

これを一般化すると、n個の選択肢(カテゴリー)をもつ変数は、(n-1)
次元の空間におけるn次元多面体として表現でき、多面体の頂点の各次元への寄与率はそれぞれ(100/n)% で、その完全な記述には (n-1)個の直行成分が必要
である。これを更に延長し、連続量も多面体として考えたい。連続量のデータの場合、カテゴリー数は、異なった数値の数と考える。この様な変数の多面体記述を通じて、非線形の関係も包括できる変数の形を導入し、同時に、連続変数、カテゴリー変数の統一的解析の枠組みを作ってはというのが、この論文の趣旨である。

先ず相関を考えよう。2個のカテゴリー変数で、かつそれぞれ3個のカテゴリーをもつ場合を例として考えると、それぞれの変数は多次元空間に同じ原点の周りにそれぞれ三角形として存在するので、一つの三角形を他の三角形に射影し、その射影された三角形と、それと同じ空間にある他の三角形との面積の比として相関を定義する。この定義では、カテゴリー間の線型、非線型の双方の関係を捉え得る。しかしカテゴリー数が3と4の変数の場合はどうなるか、更に増えると計算が困難

[1] この研究は Natural Sciences and Engineering Research Council of Canada からの研究費による。

になるのではないかという問題が残るので、上の定義は概念としてのもので、実用には、変数、成分間の相関の二乗という統計量を用いる。

射影という概念でデータの説明を考えると、従来一般に使われてきた固有値の総和を持ってデータの総情報量とする定義に問題が生ずる。5つの標準化された連続変数を考えよう。もし、5個の変数が完全な相関を持った場合、第一の固有値は5、他の4個の固有値はすべて0、固有値の総和は5。次に、すべての相関が0の場合を考えると、5個の固有値はすべて1となり、総和はこれも5である。この両極端のケースは、おなじ情報量を担っていると言えない。5個の変数が完全な相関を持つなら、一個の変数がわかれば、それで十分で、他の4個の変数はまったく冗長である。固有値の和として定義する情報量というのは、相対的な測度に過ぎず、データセットが多数あるとき、どのセットがより多くの情報を担っているかと言うような場合の答えとはならい。

固有値の和というのは、個々のエントロピーの和に類して考え、上述の問題に対し、同時エントロピーを情報量と定義してはどうであろうか。あるいは此れに相当するものとして、次式を提唱したい。

$$T = \Sigma\lambda_k - \Sigma r_{ij(k)} + \Sigma r_{ijm(k)} - \Sigma r_{ijmn(k)} + \ldots + (-1)^{(p-1)} r_{123\ldots p(k)}$$

但し、和は $i<j<m<n<\cdots<p$ に関してであり、$\Sigma\lambda_k$ は固有値の和、$\Sigma r_{ij(k)}$ は次元kにおける変数 i, j 間の相関の和、$\Sigma r_{ijm(k)}$ は次元kにおける変数 i,j,m 間の三変量間の相関、等々である (Nishisato, 2003). あるいは集合論で考えると、同時エントロピーに対応するものは、一変数、二変数、等々の独自の貢献の和と考えることが出来る。これを各成分がどのように説明するかが解析となる。

この研究は、「連続変量のデータの方が、カテゴリー化されたデータより多くの情報を荷なっている」という常識的なことへの回答を求めることに端を発した。データの幾何模型により連続量も多面体として取り上げることにより非線型の情報も捕らえられるし、其れによって上の質問への答えも自然の成り行きとして出てくる。カテゴリー数が多ければ多いほど情報量が大きい。その数が増えた場合のデータ解析の問題は、多くの困難を抱えていることが一目瞭然、其れは今後の課題としたい。

参考文献

Nishisato, S. (2003). Total information in multivariate data from dual scaling perspectives. *Alberta Journal of Educational Research, XLIX,* 244-251.

混合モデル、一般化線形モデル、GEE

— 心理学研究方法論に欠けているもの —

東京大学・医　　岸本 淳司

1　はじめに

統計学の方法論は、多変量間の相互依存関係を探索するもの (多変量解析) と因果関係を検証するもの (線形モデル・検定) とに大きく分けられる。心理学の研究でも双方の方法論が用いられるが、多変量解析系 (特に構造方程式モデル) の著しい発展に比較して、線形モデル系は新しい手法が取り入れられていないように思われる。具体的には、分散分析しか用いられていないといえる。

古典的な分散分析は、実際には制約が多いもので、適用範囲を拡張する努力が重ねられてきた。混合モデル・一般化線形モデル・GEE などがそれである。これらの方法論は、主に医学領域で活発な研究が行われているが、心理学の領域でも適用が可能であると思われる。

2　混合モデル

被験者内要因を含む実験計画を解析するとき、常識的には個人ごとに差のスコアを計算して母平均=0 の仮説に対応する一標本 t 検定 (対応のある t 検定) をするか、被験者をブロック要因とした分散分析を実行することになる。このとき、被験者の効果は固定効果として扱われていることになる。多くの場合、たまたま実験した特定の被験者について関心があるのではなく、そこからランダムサンプリングしたと想定している母集団について関心がある。そこで、被験者の効果はある分布 (たいていは正規分布) していると想定し、得られた効果はその実現値であるとみなすことがある。これがランダム効果である。固定効果とランダム効果を両方含むような線形モデルのことを混合効果モデルという。

混合効果モデルの重要な拡張点に、誤差分散に関する仮定を緩和がある。固定効果モデルでは、誤差分散は $\mathrm{var}(\boldsymbol{\varepsilon}) = \mathrm{N}(\mathbf{0}, \sigma^2\mathbf{I})$ という単純な形式しか許されなかった。現実には、個人を時間を追って測定したデータとか、家族のようなクラスターを形成するデータのように、測定値間の相関が自然に想定されるものがある。また、ランダム効果も相関があるかもしれない。そこで、混合効果モデルは次のように一般化される。

$$\begin{aligned} \boldsymbol{y} &= \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\gamma} + \boldsymbol{\varepsilon} \\ \boldsymbol{\gamma} &\sim \mathrm{N}(\mathbf{0}, \mathbf{G}) \\ \mathrm{var}(\boldsymbol{\varepsilon}) &= \mathbf{R} \\ \mathrm{var}(\boldsymbol{y}) &= \mathbf{Z}\mathbf{G}\mathbf{Z}' + \mathbf{R}. \end{aligned}$$

3 一般化線形モデル

古典的線形モデルは正規分布に対して適用されるが、現実にはさまざまなそれ以外のさまざまな分布が観測される。MuCullagh and Nelder(1989) は、正規分布以外のある種の分布 (指数型分布族と呼ばれる) に対して固定効果モデルを当てはめる画期的なモデル技法を導入した。それが一般化線形モデルである。実際によく使われるのは、2 値データに対応する二項分布と、頻度データに対応するポアソン分布である。すなわち、一般化線形モデルを使えば、2 値データあるいは頻度データを応答変数として要因効果を検討することができることになる。

一般化線形モデルでは、期待値にリンク関数と呼ばれる非線形関数を通して固定効果 $\mathbf{X}\boldsymbol{\beta}$ に関係づける。

$$g(\boldsymbol{\mu}) = \mathbf{X}\boldsymbol{\beta}$$

リンク関数には選択の余地があるが、分布ごとに決まる「正準リンク関数」を用いると都合がよい。これは二項分布ではロジット関数 ($\log(\mu/(1-\mu))$) であり、ポアソン分布のときは対数関数である。二項分布に対してロジットリンク関数を使った一般化線形モデルは、ロジスティック回帰分析と呼ばれる。

4 GEE

一般化線形モデルと混合効果モデルがあれば、それらを組合せたいというのは自然な要求である。これを一般化線形混合モデルという。モデルは次のように書ける。

$$\begin{aligned}
\boldsymbol{y} &= \boldsymbol{\mu} + \boldsymbol{\varepsilon} \\
g(\boldsymbol{\mu}) &= \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\gamma} \\
\boldsymbol{\gamma} &\sim \mathrm{N}(\mathbf{0}, \mathbf{G}) \\
\mathrm{var}(\boldsymbol{\varepsilon}) &= \mathbf{R}
\end{aligned}$$

$$\begin{aligned}
\mathrm{var}(\boldsymbol{y}) = \mathbf{V} &= \mathrm{var}(\boldsymbol{\mu}) + \mathbf{R} \\
&\approx \mathbf{BZGZ}'\mathbf{B} + \mathbf{R}
\end{aligned}$$

一般化線形混合モデルの推定にはさまざまな方法があるが、その中で安定した性能で人気がある方法に GEE がある。

5 おわりに

混合モデル・一般化線形モデル・GEE 等は、数理的な内容は難しいが、すべて SAS にとりこまれおり、すぐに実行することができる。心理学の分野でも広く利用されることを期待する。

共通因子数の決定とそれを援助するためのコンピュータ・プログラム

筑波大学心理学系　服部　環

1　目的

因子分析は観測変数の間に相関関係があるとき，観測変数の背後に少数個の因子を仮定して観測変数の共変動を説明する。この因子はもともと観測できない変数であるから，因子分析を実行する際には因子数の決定がまず問題となる。しかも，この因子数の決め方がモデル母数の推定問題よりも難しい（南風原,2002)。

従来から因子数を決定する種々の方法が提案されているが，主観的な判断に基づいて因子数を決定したり，統計ソフトウェアのデフォルト設定を利用して因子数を決定することが多い。そのように決めた因子数が常に最適であるとはいえないが，これは，主要な統計ソフトウェアが因子数を決定するために参考となる統計量を提供していないことが1つの理由であろう。そこで，本稿は以下の3点を目的とした。

(1) 因子数の決定方法を概観する。
(2) 複数の因子数を指定した上で因子分析を実行でき，因子数を決定するために必要となる主要な統計量を計算するコンピュータ・プログラムを紹介する。
(3) 実データで因子数を決定してみる。

2　因子数の決定

(1) 主に要約統計量に基づく方法
  (i) Ledermann(1937) の境界
  任意の観測変数の数に対して抽出できる最大の因子数を与える。
  (ii) Guttman(1954) 基準・Kaiser(1960) 基準
  観測変数の相関行列の1より大きい固有値の数を因子数とするが，大きすぎる傾向にある。
  (iii) Horn(1965) の平行分析
  実際の観測データと同じ大きさの乱数データを発生し，それから得られる相関行列の固有値と観測データから求めた相関行列の固有値を比較して，因子数を決める。
  (iv) Cattell(1966) のスクリー・テスト
  相関行列の固有値の減衰状況に注目する方法であり，Guttman 基準と同様に比較的よく利用されている。
  (v) SMC テスト
  観測変数の相関係数行列の対角線要素に共通性の推定値として SMC（重相関係数の2乗）を入れた行列の正の固有値の数を因子数とする。
  (vi) SAS で利用されている方法1
  SMC による共通性の初期推定値を相関係数行列の対角要素に代入し，その行列の固有値を求める。そして，固有値の大きいものから順に合計した和が，固有値の総和を超えたところまでを抽出因子数とする。
  (vii) SAS で利用されている方法2
  SMC（重相関係数の2乗）による共通性の初期推定値を1から引いて独自性の初期推定値とし，それを対角要素とする行列を独自分散行列 $\hat{\boldsymbol{\Psi}}$ とする。そして，$\hat{\boldsymbol{\Psi}}^{-\frac{1}{2}}(\boldsymbol{R}-\hat{\boldsymbol{\Psi}})\hat{\boldsymbol{\Psi}}^{-\frac{1}{2}}$ の固有値の大きいものから順に合計した和が，固有値の総和を超えたところまでを抽出因子数とする（狩野, 印刷中)。
  (viii) Velicer(1976) の MAP テスト
  第1主成分から第 $j(j \leq n-1)$ 主成分までを統制変数とする観測変数間の偏相関係数を求め，その2乗平均を最小とする j の値を抽出因子数とする。
  (ix) Zoski & Jurs(1996) の標準誤差スクリー法
  統計量を用いたスクリー・テストである。

(2) 適合性の指標に基づく方法

(i) χ^2 統計量

最尤法（ML）を用いたときに定義される検定統計量である。

(ii) 適合度統計量

適合度指標（Goodness of Fit Index;GFI），修正適合度指標（Adjusted Goodness of Fit Index;AGFI），RMSEA（Root Mean Square Error of Approximation）などがある。

(3) 情報量規準

複数のモデルの適合性を相対評価する情報量規準（AIC(Akaike's Information Criterion)，BIC(Schwarz's Bayesian Information Criterion)，CAIC(Consistent Akaike's Information Criterion)）が提案されている（Bentler,1995；Jöreskog & Sörbom,1996）。

3 プログラムの仕様

プログラムとマニュアルは服部 (2002) に公開されている。主な仕様は以下の通りである。

(1) コンパイラ　富士通 Fortran V2.0（Windows 版）

(2) 入力データ　素データあるいは相関係数行列あるいは共分散行列（下三角あるいは正方）

(3) 母数の推定　最尤法と最小 2 乗法（柳井晴夫・繁桝算男・前川眞一・市川雅教,1990），反復主因子法（共通性の初期推定値は SMC），主成分分析

(4) 因子の回転　最尤解と最小 2 乗解の基準化直交オーソマックス回転（芝,1979）およびプロマックス回転

(5) 出力　観測変数の相関係数行列・平均・標準偏差，因子パターンの初期解，回転解（因子パターン，因子間相関，因子構造），平行分析統計量，SMC テスト統計量，MAP テスト統計量，標準誤差スクリー・テスト統計量，先述した SAS で利用されている 2 つの方法で定義される累積共通分散説明率，適合性の統計量（χ^2 値，自由度，AIC，BIC，CAIC，RMSEA，GFI，AGFI，NFI，NNFI，CFI，RMSR）

4 計算例

5 つの実データセットを分析し，因子数を推定してみたところ，BIC と RMSEA が他の統計量に比べて比較的良い結果を示していた。高々 5 つの事例に過ぎないが，BIC と RMSEA が良い結果を示したことは新たな知見と言えよう。また，従来から指摘されているように，MAP テストは他の統計量よりも少ない因子数を示唆することが多かった。

芝 (1979) には，いくつの因子を抽出するかは数学的な問題というより，研究の内容に関わる問題であり，因子数を機械的に決めることに無理がある，という見解が紹介されている。芝（1979）は χ^2 統計量を用いた決定方法を念頭に置いているが，ここで分析した事例を見ると，他の方法も含めて因子数を機械的に決めることは難しいということを確認できた。複数の方法・統計量を用いて総合的に因子数を判断することになろう。

5 参考文献

Bentler, P.M. 1995 *EQS: Structural Equations Program Manual.* Encino, CA: Multivariate Software, Inc.

南風原朝和 2002 心理統計学の基礎 有斐閣 (p.345)

服部 環 2002 http://www.human.tsukuba.ac.jp/~hattori/

Jöreskog, K.G. & Sörbom, D. 1996 *LISREL 8: Structural Equation Modeling with the SIMPLIS Command Language.* Chicago: Scientific Software International.

狩野 裕 （印刷中） 因子分析（吉田光雄・狩野 裕・原田 章 印刷中 SAS による統計解析 科学技術出版）

芝 祐順 1979 因子分析法 第 2 版 東京大学出版会

柳井晴夫・繁桝算男・前川眞一・市川雅教 1990 因子分析–その理論と方法– 朝倉書店

因子数決定法の検討

Holzinger and Swineford(1939)の知能データをもとにして

香川大学経済学部　堀　啓造

使用データ

Holzinger & Swineford(1939)の知能テストのデータを分析した。Holzinger & Swineford(1939)のデータは2つある。イリノイ州のフォーレストパーク村の Pasteur 主学校 と シカゴの Grant-White 小学校 の7，8年生のデータである（日本の中学1．2年生)。それぞれ Pasteur 群，Grant-White 群と呼ぶ。Pasteur 群は平均知能が 94 と平均(100)よりやや劣る。Grant-White 群は数値は明らかでないが，学校では平均よりややよいとのことである。テストをすべて受けたものは Pasteur 群 156 名，Grant-White 群 145 名 である。明確な男女差はなかったので男女込みの分析をしている。Pasteur 群の親は外国生まれが多いし，家庭では母語を使っている。Pasteur 群 の両親とも外国生まれは 48%，両親とも米国生まれは 29%である。一方，Grant-White 群は両親とも外国生まれは 15%，両親とも米国生まれは 72%である。また，Grant -White 群のほとんどは学校の近くで生まれている。両群の知能差はこのテストにおいても確認された。

Pasteur 群は 24 テストをしている。Grant-White 群テストをしている。テスト 3,4 が難しかったので，同タイプの易しいテストをテスト 25,26 も行っている。

行ったテストは事前に空間、言語、記憶、数学能力の5因子を想定している。

表1にあるように5つのデータがある。(1)は Harman(1976)の相関行列である。(2)は Gorsuch(1983)の相関行列である。(3)～(6)は Holzinger and Swineford(1939)の素データである。

被験者の違い、少数の変数の入れ替えがあり、解の頑健性,指標の有効性をチェックするのに興味深いデータとなっている。

表1．使用データ

		使用テスト	除外テスト	被験者数	備考
(1)	Harman(1976)	test 1,2, 5-26	test 3,4 を除外	145	Grant-White A の相関行列（値が少し違う）
(2)	Gorsuch(1983)	test 1,2, 5-26	test 3,4 を除外	145	Grant-White A の相関行列（値が少し違う）
(3)	Grant-White A	test 1,2, 5-26	test 3,4 を除外	145	素データ
(4)	Grant-White B	test 1-24	test 25,26 を除外	145	素データ
(5)	Pasteur	test 1-24		156	素データ
(6)	全体	test 1-24	test 25,26 を除外	301	素データ

因子決定指標

服部(2003)の faccon.exe によって出力される指標を使用する。表2にある諸指標である。すべて faccon.exe に出力される略称を使っている。

結果

服部(2003)によって求めた結果を表2に示す。

極めて限定されたものであるが、探索的因子分析には耐えられない指標が明らかになった。推定因子数差からすると AGFI はまったく使えない。SE-SCREE, χ^2, GFI, NFI, 0.95 **を基準とする** NNFI(TLI) もよくない。PA-EIGEN-95, SMC-EIGEN, AIC, CAIC も信頼できない。推定因子数差が1の RAW-EIGEN, PA-SMC-M, BIC, 0.90 **基準の** NNFI(TLI)はどちらともいえない。推定因子数差が0の MAP, PA-EIGEN-M, PA-SMC-95, RMSEA, PGFI, RGFI, RMSR および.90 **を基準とする** CFI はこの事例においては安定した指標となっている。

表２．各種因子決定指標による因子数（適合度指標は最小因子数を示した）

test	1,2,5-26			1-24			
data	Harman	Gorsuch	Grant-White A	Grant-White B	Pasteur	全体	推定因子数差
個体数	145	145	145	145	156	301	
MAP-TEST	4	4	4	4	4	4	0
RAW-EIGEN	5	5	5	5	5	4	1
PA-EIG-M	4	4	4	4	4	4	0
PA-EIG95	3	3	3	2	4	4	2
SMC-EIGEN	13	14	13	13	12	12	2
PA-SMC-M	4	4	4	4	4	5	1
PA-SMC-95	4	4	4	4	4	4	0
SE-SCREE	7	5	7	5	6	4	3
CHI^2	7	7	6	4	5	6	3
AIC	5	4	4	5	5	6	2
BIC	3	3	3	3	3	4	1
CAIC	3	3	3	2	3	4	2
RMSEA	4	4	4	4	4	4	0
GFI	6	7	6	5	5	4	3
AGFI	-	-	-	-	-	6	++
PGFI	2	2	2	2	2	2	0
RGFI	3	3	3	3	3	3	0
RMSR	3	3	3	3	3	3	0
NFI	7	7	7	7	6	4	3
NNFI=TLI	7(4)	7(4)	6(4)	4(3)	5(4)	5(4)	3(1)
CFI	5(3)	5(3)	5(3)	4(3)	4(3)	5(3)	1(0)

推定因子数差０のうち、MAP, PA-EIGEN-M, PA-SMC-95, RMSEA が正しく４因子と推定した。PA-EIGEN-M の推定に問題のあることはすでに堀(2001)において示している。MAP, PA-SMC-95, RMSEA の３つが残ることになる。

今後は推定因子数差１以内の指標についてさらなる検討が必要である。特に, MAP と PA-SMC-95 の推定値に違いがある場合が往々にしてあるが，その間の因子数を決定するような指標が望まれる。今回の結果からはそのような指標を見つけ出すことは困難であることが推測される。

引用文献

Gorsuch, R. L. (1983). *Factor analysis*. 2nd ed. New Jersey; Erlbaum.

Harman, H. H. (1976). *Modern factor analysis*. 3rd ed. Illinois; The University of Chicago.

服部環 (2003). 共通因子数の決定とそれを援助するためのコンピュータ・プログラムの開発. *応用心理学研究, 28,* 135-144.

Holzinger, K. J. and Swineford, F. (1939). A study in factor analysis: The stability of a bi-factor solution. Supplementary Educational Monographs, No.48 . The University of Chicago.

堀　啓造(2001). Parallel analysis, http://www.ec.kagawa-u.ac.jp/~hori/yomimono/pa.html.

Maximum Propensity Score Weighted Likelihood Estimation and Its application to Structural Equation Modeling

Takahiro Hoshino Kazuo Shigemasu
Department of Cognitive and Behavioral Science,
The University of Tokyo

Estimation of differences between groups in observational studies often suffers from bias due to differences in distributions of observed covariates.For estimation of average treatment effects when the treatment variable is binary, Rosenbaum & Rubin (1983) defined propensity score and proposed an adjustment method for pre-treatment variables based on propensity score. Our aims in this study are to propose an extension of Horvitz-Thompson type propensity score adjustment method that can deal with complex models and to show asymptotic behavior of estimators.

THE MODEL AND THE ASSUMPTION

We assume that the number of treatments is J. Each unit (or individual/respondent) has theoretically J potential outcomes, but only one observation is obtained for one unit.

Let y_{ij} be the potential outcome of the i-th unit when the i-th unit is assigned to the j-th treatment (or the i-th unit is belongs to the j-th subpopulation). The missing indicator for the i-th unit of the j-th treatment is denoted by z_{ij} $(i = 1 \cdots N,\ j = 1 \cdots J+1)$ (s.t.,$z_{ij} = 1$ if the i-th unit is assigned to the j-th treatment, $z_{ij} = 0$ otherwise). For each unit, z is equal to one only for a certain treatment (i.e., $\sum_{j=1}^{J+1} z_{ij} = 1$).

Let also the marginal distribution of y_{ij} is $p(y_{ij}|\theta_j, \theta_c)$ $(i = 1 \cdots N)$, where θ_j is the parameter vector unique to the j-th treatment and θ_c is common to every treatment.

For simplification of notation, we define $\theta = (\theta_1, \cdots, \theta_J, \theta_c)$. Let $\theta^0 = (\theta_1^0, \cdots \theta_j^0, \theta_c^0)$ be the true value of θ.

We employ the "weak unconfoundedness" assumption proposed by Imbens (2000):

$$z_j \perp y_j | w_j(x) \quad for\ all\ j \tag{1}$$

where $w_j(x) = Pr(Z_j = 1|x)$ is the generalized propensity score for the j-th treatment. The usual maximum likelihood estimation using the observed portion of data would fail to find the true values of parameters except when y_j and x are mutually independent.

ASYMPTOTIC DISTRIBUTION OF MAXIMUM PROPENSITY SCORE WEIGHTED LIKELIHOOD ESTIMATORS

In this section we rewrite the sampling probability w_{ij} as $w_{ij}(\alpha_j)$ by regarding w_{ij} as a function of the unknown parameter α_j.

Let the joint distribution of $z = (z_{11}, \cdots, z_{ij}, \cdots z_{NJ})$ be $\prod_{j=1}^{J} \prod_{i=1}^{N} w_{ij}(\alpha_j)^{z_{ij}} (1 - w_{ij}(\alpha_j))^{1-z_{ij}}$.

Let α^0 be the true value of $\alpha = (\alpha_1, \cdots, \alpha_J)$.

We define the propensity score weighted log likelihood $L_N^W(y|\theta, \alpha)$:

$$L_N^W(y|\theta, \alpha^0) = \sum_{j=1}^{J} \sum_{i=1}^{N} \frac{z_{ij}}{w_{ij}(\alpha_j^0)} \log p(y_{ij}|\theta_j, \theta_c) \tag{2}$$

Let $\tilde{\theta}(\alpha^0)$ denote MWLE with the true value of α given.

We usually do not know the true value α^0, so we calculate the maximum likelihood estimator of α_j, $\hat{\alpha_j}$ and substitute $\hat{\alpha}$ for α^0.

We define $\tilde{\theta}(\hat{\alpha})$ that maximize the following function of θ:

$$L_N^W(y|\theta,\hat{\alpha}) = \sum_{j=1}^{J}\sum_{i=1}^{N} \frac{z_{ij}}{w_{ij}(\hat{\alpha_j})} \log p(y_{ij}|\theta_j,\theta_c), \tag{3}$$

and call it "the maximum propensity score weighted likelihood estimator (MPWLE)" with MLE of α given.

Theorem 1 $\tilde{\theta}(\hat{\alpha})$ is consistent and asymptotically normal with variance-covariance matrix $\frac{1}{N^2} I(\theta^0)^{-1}[\sum_{j=1}^{J} I_j(\theta^0) \sum_{i=1}^{N} \frac{1}{w_{ij}}] I(\theta^0)^{-1}$.

Proof Expanding $\partial L_N^W(y|\tilde{\theta}(\alpha^0),\alpha^0)/\partial\theta = 0$ about θ^0 yields

$$\sqrt{N}(\tilde{\theta}(\alpha^0) - \theta^0) = -\sqrt{N}(L_N^{''W}(y|\theta^0,\alpha^0))^{-1} \frac{\partial}{\partial\theta} L_N^W(y|\theta^0,\alpha^0) + o_p(1). \tag{4}$$

Differentiating with respect to α yields

$$\sqrt{N}\frac{\partial}{\partial\alpha}\tilde{\theta}(\alpha^0) = -\sqrt{N}(L_N^{''W}(y|\theta^0,\alpha^0))^{-1} \frac{\partial}{\partial\theta\partial\alpha^t} L_N^W(y|\theta^0,\alpha^0) + o_p(1). \tag{5}$$

The asymptotic distribution of $\tilde{\theta}(\hat{\alpha})$ follows from the Taylor approximation:

$$\sqrt{N}(\tilde{\theta}(\hat{\alpha}) - \theta^0) = \sqrt{N}(\tilde{\theta}(\alpha^0) - \theta^0) + \sqrt{N}\frac{\partial}{\partial\alpha}\tilde{\theta}(\alpha^0)(\hat{\alpha} - \alpha^0) + o_p(1). \tag{6}$$

As N increases,

$$\begin{aligned}\frac{1}{N}\left(\frac{\partial^2}{\partial\theta\partial\alpha^t} L_N^W(y|\theta^0,\alpha^0)\right) &\xrightarrow{p} -\frac{1}{N}\sum_{j=1}^{J}\sum_{i=1}^{N} E_{\theta^0,\alpha_j^0}\left[\frac{z_{ij}}{w_{ij}^2(\alpha_j^0)}\frac{\partial w_{ij}(\alpha_j^0)}{\partial\alpha_j}\frac{\partial}{\partial\theta}\log p(y_{ij}|\theta_j,\theta_c)\right] \\ &= -\frac{1}{N}\sum_{j=1}^{J}\sum_{i=1}^{N} E_{\alpha_j^0}\left[\frac{z_{ij}}{w_{ij}^2(\alpha_j^0)}\frac{\partial w_{ij}(\alpha_j^0)}{\partial\alpha_j}\right] E_{\theta^0}\left[\frac{\partial}{\partial\theta}\log p(y_{ij}|\theta_j,\theta_c)\right] = 0. \end{aligned} \tag{7}$$

Because $\tilde{\theta}(\alpha^0)$ is consistent and $\hat{\alpha}$ is asymptotically efficient, these two random variables are asymptotically independent (see e.g., Pierce, 1982). Then we get Theorem 1. □

It should be noted that the asymptotic distribution of $\tilde{\theta}(\hat{\alpha})$ is equal to $\tilde{\theta}(\alpha^0)$, indicating that the variation of the estimator of α does not influence that of θ asymptotically.

We can also show that the weighted lkelihood ratio test statistic is distributed asymptotically as a weighted sum of independent χ_1^2 random variables, also arise from pseudo-maximum likelihood estimation (Liang & Self, 1996).

References

Imbens, G.W. (2000). The role of the propensity score in estimating dose-response functions. *Biometrika,* **87**, 706–710.

Liang, K-Y., & Self, S.G. (1996). On the Asymptotic Behavior of the Pseudolikelihood Ratio Test Statistic. *Journal of the Royal Statistical Society, series B,* **58**, 785–796.

Rosenbaum, P.R., and Rubin, D.B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika,* **70**, 41–55.

非補償因子分析モデル

大学入試センター研究開発部・早稲田大学文学研究科
荘島宏二郎

1 導入

通常の因子分析 (factor analysis, FA) モデルは，ある次元の潜在得点の値が小さくても，他の潜在得点の値が高ければ，高い項目得点をとることができる．この関係を潜在次元の間に補償関係があると呼ぶことにする．この潜在次元の間の補償関係は，必ずしも成立しない場合もあるだろう．このような問題意識を出発点として，本研究では，潜在次元の間に補償関係が成立していない，いわば潜在特性間に非補償の関係がある因子分析モデルを提案する．

2 非補償因子分析モデル

通常の FA モデル (補償 FA モデル) は

$$h(x|\boldsymbol{\theta}) = \frac{\boldsymbol{\alpha}'\boldsymbol{\gamma}}{\sqrt{2\pi}} \exp\Big[-\frac{1}{2}\{\boldsymbol{\alpha}'(\boldsymbol{\theta}-\boldsymbol{\beta}-\boldsymbol{\gamma}x)\}^2\Big] = \frac{\boldsymbol{\alpha}'\boldsymbol{\gamma}}{\sqrt{2\pi}} \exp\Big[-\frac{1}{2}(\boldsymbol{\theta}-\boldsymbol{\beta}-\boldsymbol{\gamma}x)'\boldsymbol{A}(\boldsymbol{\theta}-\boldsymbol{\beta}-\boldsymbol{\gamma}x)\Big] \tag{1}$$

のように表現できる．ここで，$\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}$ は M 次元ベクトルである．また，$\boldsymbol{A} = \boldsymbol{\alpha}\boldsymbol{\alpha}'$ である．サイズが M の正方行列 $\boldsymbol{A}$ の mn 要素は $\alpha_m\alpha_n$ となっている．仮に $\boldsymbol{A}$ の非対角要素を 0 と置く．すると，(1) 式は

$$\begin{aligned} g^*(x|\boldsymbol{\theta}) &= \frac{\boldsymbol{\alpha}'\boldsymbol{\gamma}}{\sqrt{2\pi}} \exp\Big[-\frac{1}{2}\{\boldsymbol{\alpha}\odot(\boldsymbol{\theta}-\boldsymbol{\beta}-\boldsymbol{\gamma}x)\}'\{\boldsymbol{\alpha}\odot(\boldsymbol{\theta}-\boldsymbol{\beta}-\boldsymbol{\gamma}x)\}\Big] \\ &= \frac{\boldsymbol{\alpha}'\boldsymbol{\gamma}}{\sqrt{2\pi}} \exp\Big[-\frac{1}{2}\sum_m \alpha_m^2(\theta_m-\beta_m-\gamma_m x)^2\Big] \end{aligned} \tag{2}$$

となり，これを非補償因子分析モデルとして提案する ([1])．このモデルの条件付き得点の期待値は，Figure 1 のようになり，一方の潜在特性が高くても得点はある一定のところで頭打ちになる．高い得点を取るには，全ての潜在特性が高いことが要請される．また，このような反応は，非線形項，あるいは交互作用項と捉えることもできる.[1] なお，通常の因子分析モデルの条件付き得点の期待値は Figure 6 のようになる．

補償モデルと非補償モデルを分けるものは，$\boldsymbol{A}$ において，非対角要素を考えるか否かということが重要な問題となり，アナロジーとして，補償モデルはスキャナ・キャリッジ・アナロジー，非補償モデルはサーチライト・アナロジーとして説明が可能である ([2])．なお，完全に単純構造が成立している状況では，補償モデルと非補償モデルは同一のモデルとなる．

3 補償モデルと非補償モデルの混合モデル

m 番目と n 番目の潜在特性が補償か非補償かを評価するか否かを決めるダミー変数 d_{mn} を考えて，

$$d_{mn} = d_{nm} = \begin{cases} 1 & m \text{ 番目と } n \text{ 番目の潜在特性が「補償」の関係のとき} \\ 0 & m \text{ 番目と } n \text{ 番目の潜在特性が「非補償」の関係のとき} \end{cases} \tag{3}$$

とする．そして，このような d_{mn} を要素としてもつ対称行列 $\boldsymbol{D}$ を用意して，$\boldsymbol{A}$ の代わりに $\boldsymbol{D}\odot\boldsymbol{A}$ を用いることができる．

[1] この点につきましては，2003 年度の第 31 回行動計量学会 (於 名城大学) において狩野先生 (大阪大学) にご指摘いただきました．ありがとうございました．

4 補償度を連続的に考えるモデル

どの程度の補償関係 (非補償関係) があるかについて，$\boldsymbol{D}$ の各要素を [0,1] の範囲をもつ連続量と考えることができる．$\boldsymbol{D}$ の要素を連続的に操作すると，Figure 2-5 のような補償モデルと非補償モデルの中間的なモデルを考えることができる．

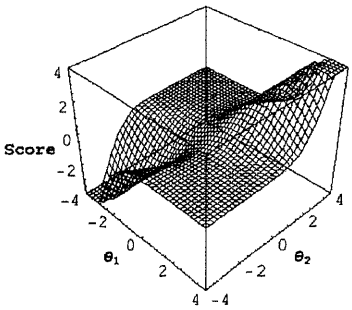


Figure 1: 非補償 FA(d=0.0)

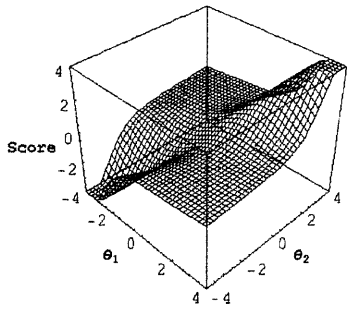


Figure 2: d=0.2

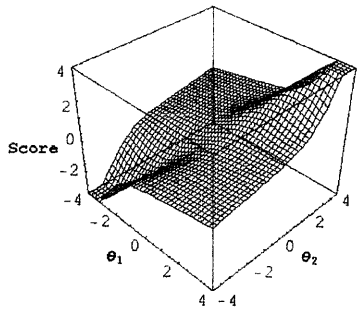


Figure 3: d=0.4

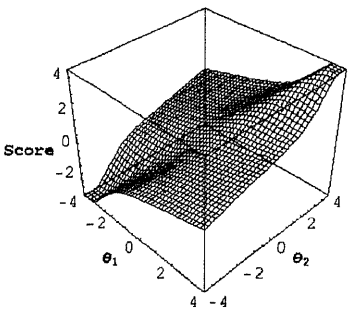


Figure 4: d=0.6

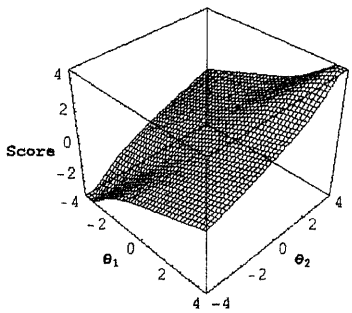


Figure 5: d=0.8

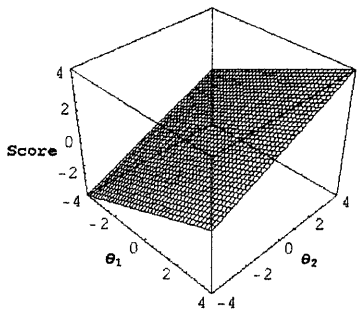


Figure 6: 通常の FA(d=1.0)

5 議論

本研究では，非補償因子分析モデルを提案し，当該モデルがもつ性質を通常の因子分析モデルの比較を通して述べた．そこでは，一方の因子得点が低いとき，もう一方の因子得点が高くても観測変数の値が大きくならない．このようなモデルは，潜在特性間に補償関係が成立しないであろう多くの応用場面で有効に働く可能性がある．また，補償度というものを導入し，潜在変数間の補償関係を連続的に評価することが可能であることが示された．

文献

[1] 荘島宏二郎 (2003) 非補償因子分析モデル 狩野裕・千野直仁 (編) 科学研究費シンポジウム「数理統計学と計量心理学をつなぐ」講演予稿集 pp.160–168.

[2] 荘島宏二郎 (2003 年 11 月 5 日) 非補償因子分析モデル 科学研究費シンポジウム「数理統計学と計量心理学をつなぐ」発表資料 (於 大阪大学)

心理学と独立成分分析

宮本 友介
大阪大学大学院 人間科学研究科

1. はじめに 独立成分分析 (independent component analysis; ICA) とは，複数の互いに独立な潜在変数が，その線形混合として観測されるという仮定の下で，観測データから「独立性の最大化」という基準のみによって元の潜在変数を復元することを目指す多変量解析手法である (Comon, 1994)．これは主として信号処理などの工学分野で用いられてきたが，近年では神経科学における脳磁図解析や，市場経済の時系列データの分析 (Kiviluoto & Oja, 1998) など，さまざまな分野での応用が試みられている．しかし，心理学の領域では，因子分析や主成分分析といった手法が取り挙げられる一方で，独立成分分析を適用した事例はいまだ数少ない．本研究では，独立成分分析を心理学データに適用する際の問題点を考え，それらの問題がどのようなところから起こるのか，またどのように対処すればよいのかということを検討することを目的としている．とくに，正規誤差が存在するもとでの独立成分分析モデルについて，信号の復元するだけではなく，混合行列を推定するための簡明な方法を提案する．正規誤差を含んだ独立成分分析モデルは，因子に非正規性を仮定した因子分析モデルと同等である．古典的な因子分析モデルでは識別性についての制約的な条件があるが，ここでは，因子の非正規性の仮定の下で，識別可能でない因子分析モデルの推定を可能にする方法を提案した．

2. 独立成分分析モデル 互いに独立な n 変量の確率ベクトル $\mathbf{s} = (s_1, \ldots, s_n)^T$ が未知の確率分布 $p(\mathbf{s})$ に従っており，それらが離散時間 $t = 1, 2, \ldots$ において観測されるという状況を考える．われわれは $\mathbf{s}$ を直接観測することができず，その線形混合である

$$\mathbf{x}(t) = A\,\mathbf{s}(t) \tag{1}$$

を通してのみ観測が可能とする．このとき，「確率ベクトルの各成分は互いに（統計的に）独立である」すなわち，$p(\mathbf{s}) = p_1(s_1) \cdots p_n(s_n)$ という情報だけを用いて，観測ベクトル $\mathbf{x}$ から元の $\mathbf{s}$ を復元することが独立成分分析の目的である．最も単純な状況では，A が $n \times n$ の正方行列のときであり，この場合，

$$\mathbf{y} = W\mathbf{x} \tag{2}$$

として観測値の線形結合を考え，$\mathbf{y} = (y_1, \ldots, y_n)^T$ の各成分が互いに独立になるような変換 W を求める問題となる．このとき W は成分のスケールと順序の不定性を除けば混合行列 A の逆行列の推定値となっており，

$$W = PDA^{-1} \tag{3}$$

と表すことができる．ここで，D は $\mathbf{y}$ の成分のスケールを調整する適当な対角行列，P は成分の順序を調整する置換行列である．一般的には，スケールに関する不定性を除去するには $\mathbf{s}$ の分散を 1 とするなどの制約が置かれる．主成分分析などとは対照的に，順序に対する不定性についてはあまり考慮されないことが多いが，便宜上，混合行列 A の列ベクトルのノルムや，目的関数である独立性の基準や Kolmogorov Smirnov 検定統計量によって順序づけがおこなえる．なお，これらの不定性を除去した後も，独立成分の符号については一意に定めることはできないという点には注意を要する (Comon, 1994)．

3. 心理学データと独立成分分析 多変量解析の手法の多くが極限分布としての正規性を暗黙のうちに仮定しているのに対して，独立成分分析は非正規性の高い成分を抽出しようとする，従来の方法とは異なる視点を与えるものであり，探索的なデータ解析手法としては大いに意義があるものである．しかし，独立成分分析は主として物理的信号処理という分野で理論的体系が整えられてきたという背景もあり，心理学的データに適用するに際していくつか問題となる点がある．具体的なものとしては，

- データの標本サイズが比較的小さい
- 「外れ値」の影響を受けやすい
- 測定に未知の誤差が含まれる
- 独立成分の「解釈」が困難である

といった点が挙げられるだろう．これらは，必ずしも互いに独立した問題ではない．たとえば，データの標本サイズが小さいと，パラメータ推定について過学習の問題が生じると考えられる．その場合，分析結果の再現性は低くなり，「外れ値」に対する頑健性も乏しい，といった状況に陥ることになる．また，そうした状況では，独立成分を解釈すること自体が無意味なことになってしまう．こうした問題に対して，Hyvärinen et al. (1999) は小標本データについては主成分分析による前処理で独立成分空間をあらかじめ削減しておくことで，過学習の問題を回避する方法を提案している．しかし，こうした方法は，正規誤差のないモデルの場合にはうまくいくが，心理学データではそのような仮定をおくことは困難であると考えられる．また，信号処理の分野では，これらの問題はセンサの性能に依存するものであったり，また研究の主眼が独立成分（信号源）$\mathbf{s}$ を復元することに置かれることが多いため，あまり活発な議論は交わされていないのが現状である．これに対して，心理学研究ではシステム A の理解が重要視されるので，上のような問題は不可避となる．心理学データでは，とくに非正規因子分析モデル（正規誤差を含む独立成分分析モデル）において，混合行列 A を推定するということを考えなければならない．

独立成分分析モデルが観測に誤差をともなうの場合，すなわち，

$$\mathbf{x} = A\mathbf{s} + \mathbf{e}, \quad \mathbf{e} \sim N(0, \sigma^2 I) \tag{4}$$

を考えるときには，因子分析モデルとの深い関連がある．ただし，因子分析では通常，共通因子に注目されるが，独立成分分析ではとくに共通因子と特殊因子の区別はしない．このモデルは非正規因子分析モデル，すなわち，共通因子が非正規的な分布にしたがうものと仮定した因子分析モデルとして捉えることができる．一般に，観測に誤差をともなう独立成分分析では，あらかじめ誤差（ノイズ）のフィルタリングをおこなった後，誤差のない独立成分分析モデルを当てはめるという手法がとられる．とくに音声データ

や画像データをあつかう信号処理の分野では，比較的ノイズの性質が明らかとなっているために，帯域透過フィルタなどでノイズを除去するという作業が有効である．しかし，この方法は正規誤差の分散が十分に小さい場合には有効であるが，そうした状況は現実のデータではあまり起こり得ないものである．また，データが正規分布にしたがう場合，主成分分析は非常に効率的な表現を与えるが，非正規データでは独立成分分析と全くことなる結果を示す．したがって，誤差分散の大きさについての情報がない状況で心理学データにこの方法を適用することはできない．

そこで，通常の因子分析によって誤差分散があらかじめ推定できるときには，因子分析を前処理として共分散行列を修正し，誤差なしの独立成分分析モデルに帰着させるという方法がとられる．ただし，「Ledermann の境界」などの通常の因子分析モデルが受ける制約は継承される．実際には，独立成分分析では高次統計量の情報も用いるため，独立成分の数には理論上の上限がないが，前処理として因子分析を用いることが隘路となってしまうのである．

Attias (1999) は，非正規因子分析モデルを一般化したアルゴリズムとして，「独立因子分析 (independent factor analysis; IFA)」を提案している．これは，非正規因子 (独立成分) の分布を混合正規分布で近似し，EM アルゴリズムでパラメータ推定をおこなうというものである．このモデルは非常に適用範囲が広いが，たとえば独立成分の分布をいくつの正規分布の混合で近似すべきかといった，新たな問題を生み出すことになる．結果として，推定するべきパラメータが多くなるため，過学習の問題を引き起こす可能性も高くなると考えられる．

これに対して，Akuzawa (2000) は 4 次のクロスキュムラントを基準として用いることで，正規誤差の影響を受けずに W を求めるアルゴリズムを提案している．しかし，観測変数が独立成分の数よりも多い場合には，混合行列 A を一意に定めることができないという問題がある．そこで，ここでは Akuzawa (2000) の挙げたクロスキュムラントタイプの推定方法を用い，任意の独立成分数のモデルでも動作するような拡張を加えることを検討した．まず混合行列として $m \times n$ 行列 A を考えたとき，A が正則な正方行列 ($m = n$) ならば，$A = W^{-1}$ により A を一意に推定することが可能である．そこで 混合行列 A が $m > n$ の場合には，観測に既知の $m \times (m-n)$ 行列 B と $m-n$ 次のベクトルからなる疑似変数 $B\mathbf{t}$ を加えることで，行列を平方化することが可能である．すなわち，

$$\begin{aligned}\mathbf{x}' &= \mathbf{x} + B\mathbf{t} = A\mathbf{s} + B\mathbf{t} + \mathbf{e} \\ &= [A\,B]\begin{bmatrix}\mathbf{s}\\ \mathbf{t}\end{bmatrix} + \mathbf{e} \end{aligned} \tag{5}$$

とすれば，$[A\,B] = W^{-1}$ により，A 順序の不定性を除けば解は一意に定まる．$\mathbf{t}$ として sin 関数などの特徴的な変数を選べば，復元された独立成分の中から擬似変数に対応するものを容易に特定することが可能である．このとき，行列 $[A\,B]$ は正則でなければならないので，擬似混合行列 B は未知の混合行列 A と一次従属とならないように選択する必要がある．シミュレーション実験の結果，わずかな従属性は問題にならないが従属性が高まると致命的な問題となり得ることがわかっている．こうした問題が起こったときには，B を (ランダムに) 選び直すか，モデルの独立成分の個数を再検討する必要がある．

4. まとめと展望 独立成分分析は，独立性最大化を基準とした潜在変数モデルの一つである．主成分分析と同様，外的基準を必要とせず，探索的な分析手法として有用である．

正規ノイズが存在するもとでの独立成分分析モデルにおいて，信号の復元だけでなく，混合行列をノンパラメトリック的に推定する簡明な方法を提案した．このモデルは因子に非正規性を仮定した因子分析モデルと同等である．古典的な因子分析モデルにおいては識別性について制約的な条件があるが，ここで提案する方法は，識別可能でない因子分析モデルの推定を，因子の非正規性の下で可能にする．

心理学では，何らかの先験的知識 (あるいは，仮説) をもってデータを扱うことが多い．独立成分分析に先験的知識による構造を取り込むことができれば，上述のような独立成分の取捨において有益であると考えられる．しかし信号処理分野では，独立成分分析はそうした事前情報がない状況でも適用できる手法であるという前提の下で研究されてきたため，先験知識を生かしたアプローチは未だ不十分であり，今後検討していく価値がある．

独立成分分析では，高次統計量を扱うため，従来の 2 次統計量に基づく因子分析モデルなどに比して，より多くの情報をデータから汲み取ることができる．しかし，高次の非線形相関については，その意味合いを解釈することが困難であることも事実である．また，これまでの独立成分分析の流れでは，モデルの適合度や推定値の信頼性を評価する方法はあまりとり上げられてこなかった．独立成分の数を選択するにも，明確な基準を挙げている研究 (e.g. Roberts, 1998) は少ないのが現状である．しかし最近になって，Himberg & Hyväinen (2003) などの研究がブートストラップ法やクラスタリングを用いて独立成分分析の推定結果を評価しようと試みており，期待が高まりつつある．

参考文献

Akuzawa, T. (2000). Extended quasi-Newton method for the ICA, Proceedings of the International Workshop on Independent Component Analysis and Blind Signal Separation: 521–525.

Attias, H. (1999). Independent factor analysis. *Neural Computation*, **11**(4): 803–851.

Comon, P. (1994). Independent component analysis, A new concept? *Signal Processing*, **36**(3): 287–314.

Friedman, J. H. & Tukey, J. W. 1974 A projection persuit algorithm for exploratry data analysis. *IEEE Transactions on Computers*,**C–23**: 881–890.

Himberg, J. & Hyväinen, A. (2003) Icasso: Software for investigating the reliability of ICA estimates by clustering and visualization. *A conference paper in Neural Networks for Signal Processing (NNSP2003)*.

Hyvärinen, A., Särelä, J. & Vigário, R. 1999 Spikes and bumps: Artefacts generated by independent component analysis with insufficient sample size. In *Proc. Int. Workshop on Independent Component Analysis and Signal Separation (ICA'99)*: 425–429.

Kiviluoto, K. & Oja, E. (1998). Independent component analysis for parallel financial time series. *Proc. Int. Conf. on Neural Information Processing ICONIP'98*, pp. 895–898.

Roberts, S. J. (1998). Independent component analysis: Source assessment & separation, a bayesian approach. *IEE Proceedings on Vision, Signal & Image Processing* **145**(3), pp. 149–154.