

Likert尺度と数量化理論： 多重対応分析を主成分分析として使う

村上 隆

(中京大学・現代社会学部)

本講演の内容は、

村上 隆 (2016). 多重対応分析の変換としての正規直交主成分分析 —Likert型項目の探索的分析のための新たな手続き— 「データ分析の理論と応用」, Vol.5, 印刷中に基づく。

Likert型項目（質問文＋順序のあるカテゴリー尺度（例））

あなたは、過去3ヶ月の間に、以下の感情を各場面でどのくらい経験しましたか？それぞれ最もあてはまると思う数字を、○印で囲んでください。

		全く感じなかった	あまり感じなかった	少し感じました	かなり感じました	とてもよく感じました	
		学業					
1. うれしい気持ち		1	2	3	4	5	
2. 満足な気持ち		1	2	3	4	5	
3. いらいらした気持ち		1	2	3	4	5	
4. 楽しい気持ち		1	2	3	4	5	

5. 心配な気持ち*

6. 面白い気持ち

7. 悲しい気持ち*

8. 充実した気持ち

9. 憂うつな気持ち*

10. 安心した気持ち

11. 満たされない気持ち*

12. 腹立たしい気持ち*

Assessment, Screening のために個人差測定尺度として用いられることも多い。

問題・目的：MCAの心理学的利用

3つの問い

Q1. 順序尺度に過ぎない Likert 型項目を間隔尺度として扱っていいのか？

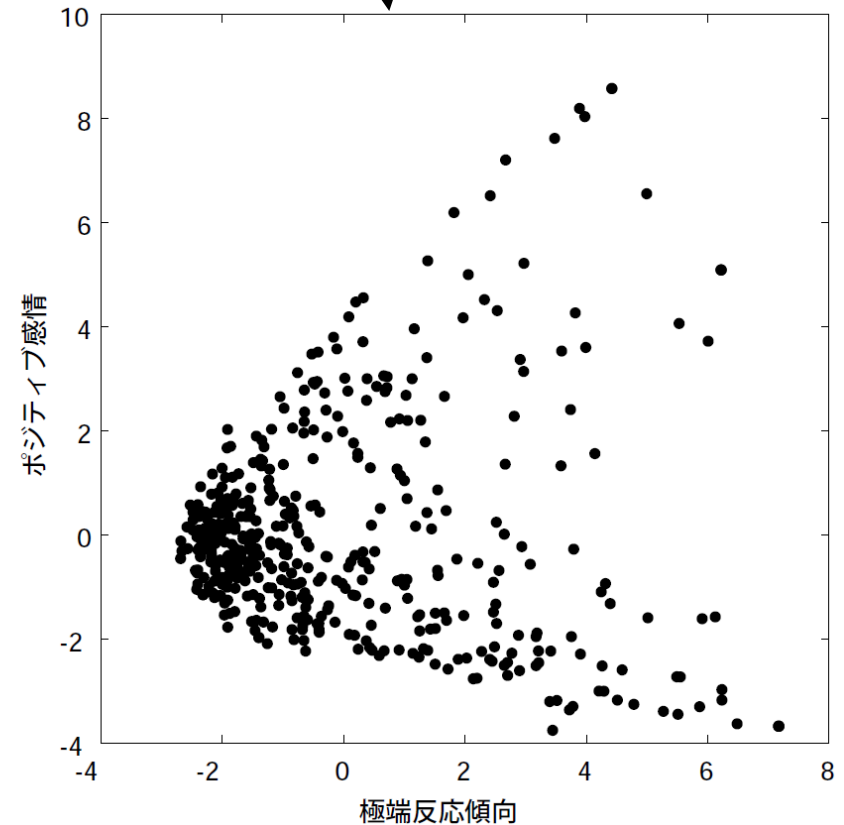
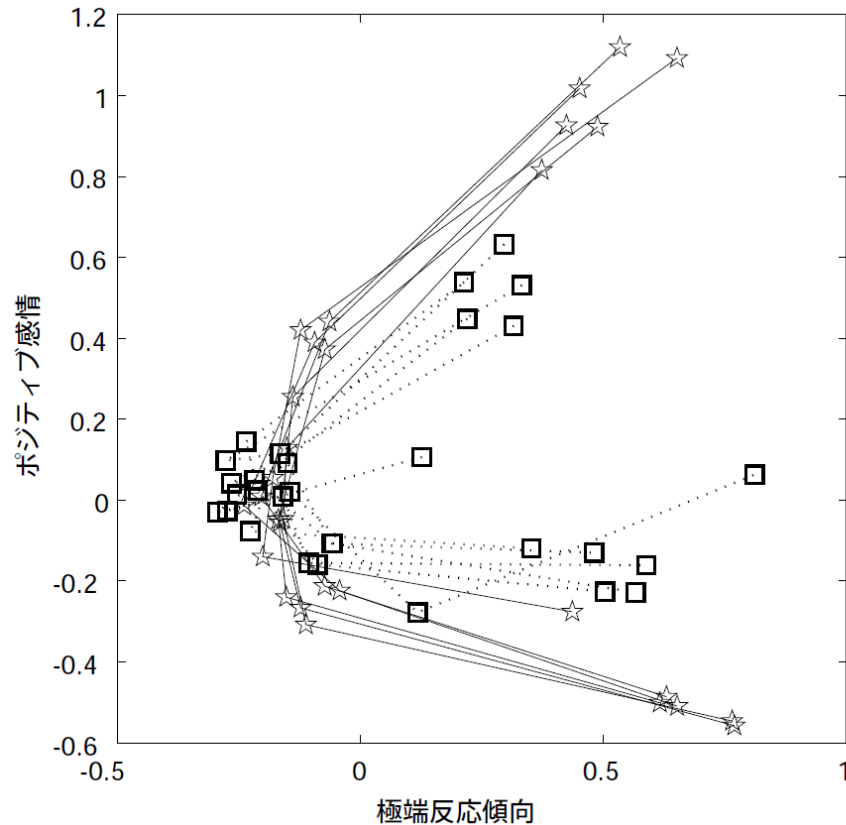
（1と2, 2と3の間隔が同じかということよりも, 1と5にばかり○をつける個人と3の周辺でほとんど動かない個人の差は？ 効用の個人間比較の不可能性の問題）

Q2. メトリックな探索的「因子分析」(exploratory “factor analysis”; EFA)によって, 多くに成果が得られている。しかし, ノンメトリックな多重対応分析 (multiple correspondence analysis; MCA)の結果とは著しく違う。なぜか？

Q3. MCAの結果を, EFA 風に負荷行列を回転して解釈することはできないのか？

MCAの結果

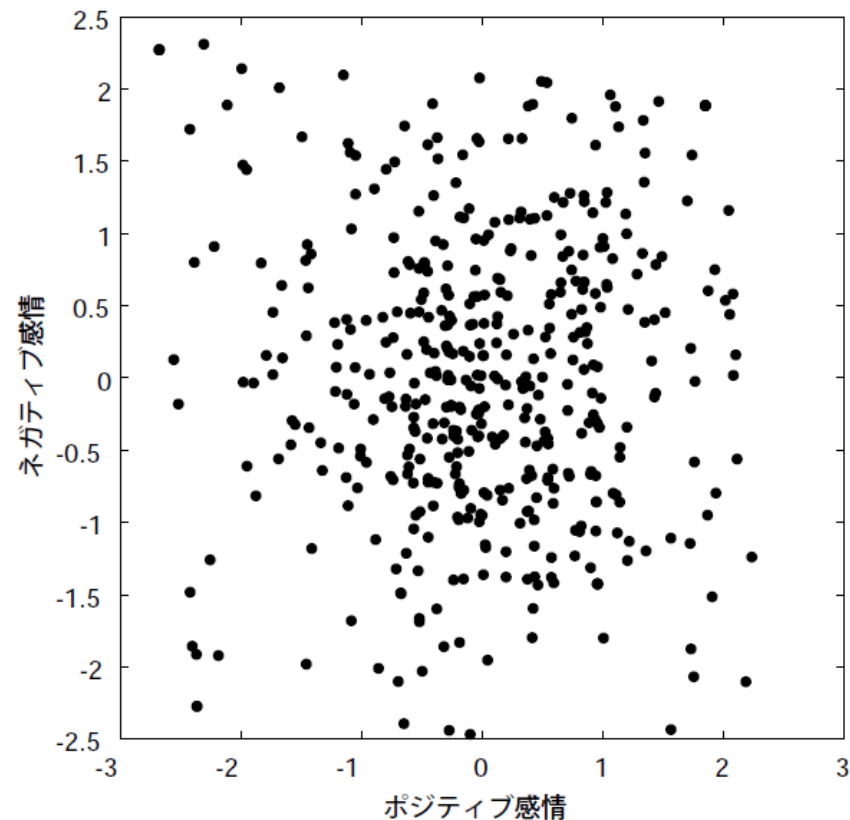
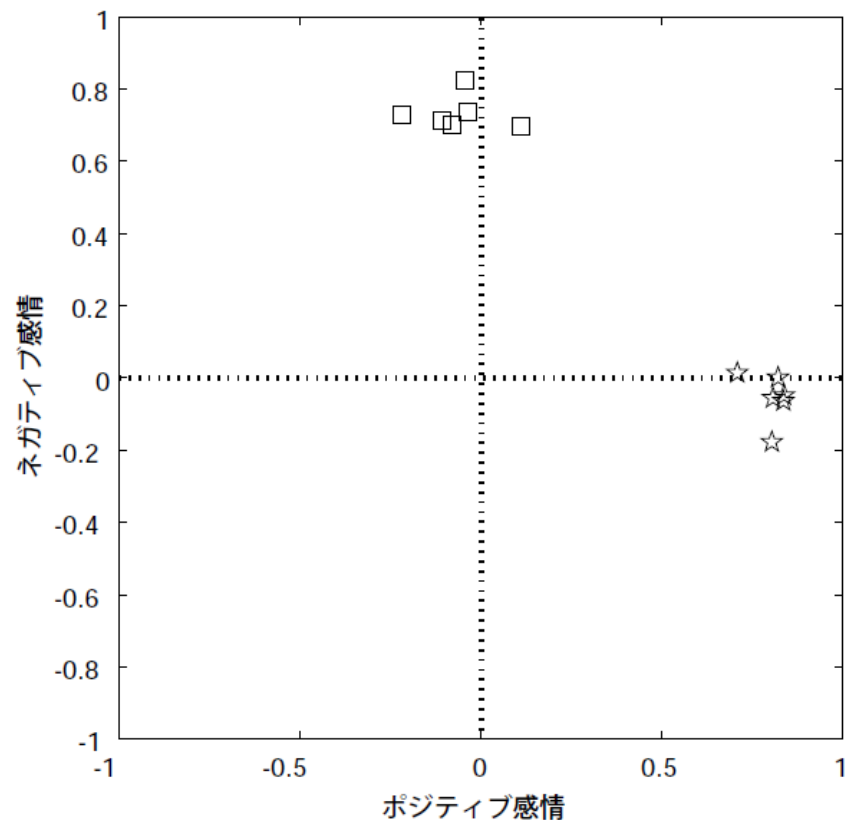
馬蹄現象 → 余剰次元？



☆ — ポジティブ項目, □ — ネガティブ項目

数量化重み (左) と数量化得点 (右)

PCA + Varimax rotation



☆ — ポジティブ項目, □ — ネガティブ項目

回転後の負荷行列(左)と主成分得点行列(右)

心理学者が EFA を好む理由

- 問題の焦点は、(社会ではなく)個人(差)にある。個人の特性のアセスメントが overt な目的となることも多い。
- 複数の次元で個人差をとらえたい。
- 分析は、この段階では(どちらかと言えば)まだ始まったばかり。得られた数量(あるいは、分析を通じて得られた尺度得点の定義)は、次の段階に持ち越される。
- 群間の差、他の変数との相関関係、これらを通じての尺度得点の妥当性検討、因果分析など。
- 空間表示よりは、個々の軸上での相対的位置に関心がある。
- そのことが、やや特殊な Likert 型項目への「偏愛」を生み出した。

本講演の概要

- 正規直交多項式主成分分析 (orthonormal polynomial principal component analysis; OPPCA) と呼ぶ方法を開発
- (1) MCAの解法の中核となる基準化Burt行列の固有値分解に、自明解の存在と、それにともなう重み行列の変数ごとの中心化という性質があることを確認する.
 - (2) MCAの重み行列を変数ごとに、カテゴリーの数量と数量化得点算出のための重みに分解する. カテゴリーの数量を用いて得られる相関行列は、基準化Burt行列と同じ固有値をもち、MCAはPCAの手続きに還元できることを示す. 実際、PCAによる主成分得点は数量化得点と一致する.
 - (3) その際、カテゴリーの数量は中心化され相互に直交している限り、全く任意であることが導かれる. 解を一意に定め、解釈を容易にするために、カテゴリーコードを数値として扱う正規直交多項式によって重みを定める.
 - (4) 重み行列を単純構造に向けて回転する.

記号の定義

- $\mathbf{X}$ --- n (個体) $\times$ p (変数) のデータ行列 (5段階の評定値)
- $\mathbf{G}_k$ --- 変数 k のダミー変数行列 ($n \times c (=5)$) の2値
- $\mathbf{V}_k$ --- 変数 k に対する重み (quantifying) 行列 ($c \times r$)
- $\mathbf{Y}_k = \mathbf{G}_k \mathbf{V}_k$ --- 変数 k の個別数量化 (quantified) 得点行列
- $\mathbf{F} = \mathbf{Y}_1 + \mathbf{Y}_2 + \cdots + \mathbf{Y}_p$ --- (全体) 数量化得点行列
- $\mathbf{B}_{kl} = \mathbf{G}_k' \mathbf{G}_l$ --- Burt 行列 (の部分行列)
- $\mathbf{D}_k = \mathbf{G}_k' \mathbf{G}_k$ --- 変数 k の単純集計を要素とする対角行列

$$\mathbf{G} = \begin{bmatrix} \mathbf{G}_1 & \mathbf{G}_2 & \cdots & \mathbf{G}_p \end{bmatrix}, \quad \mathbf{V} = \begin{bmatrix} \mathbf{V}_1 & \mathbf{V}_2 & \cdots & \mathbf{V}_p \end{bmatrix}', \quad \mathbf{F} = \mathbf{G}\mathbf{V}$$

$$\mathbf{B} = \begin{bmatrix} \mathbf{D}_1 & \mathbf{B}_{12} & \cdots & \mathbf{B}_{1p} \\ \mathbf{B}_{21} & \mathbf{D}_2 & \cdots & \mathbf{B}_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{B}_{p1} & \mathbf{B}_{p2} & \cdots & \mathbf{D}_p \end{bmatrix} = \mathbf{G}'\mathbf{G} \quad \mathbf{D} = \begin{bmatrix} \mathbf{D}_1 & \mathbf{O} & \cdots & \mathbf{O} \\ \mathbf{O} & \mathbf{D}_2 & \cdots & \mathbf{O} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{O} & \mathbf{O} & \cdots & \mathbf{D}_p \end{bmatrix} = \text{dg } \mathbf{G}'\mathbf{G}$$

データ行列 X とダミー変数行列 G

- カテゴリーコード(x)を, 順序のない2値変数(g)に。

$$g_{ijk} = \begin{cases} 1 & \text{if } x_{ik} = j \\ 0 & \text{if } x_{ik} \neq j \end{cases}$$

[illegible]

Burt 行列: $B = G'G$

	g_{11}	g_{21}	g_{31}	g_{41}	g_{51}	g_{12}	g_{22}	g_{32}	g_{42}	g_{52}	g_{13}	g_{23}	g_{33}	g_{43}	g_{53}	g_{15}	g_{25}	g_{35}	g_{45}	g_{55}
g_{11}	55	0	0	0	0	38	14	3	0	0	3	5	10	14	23	4	4	10	6	31
g_{21}	0	146	0	0	0	13	91	36	3	3	6	29	39	51	21	8	13	26	52	47
g_{31}	0	0	175	0	0	6	20	114	30	5	5	30	60	54	26	6	15	48	66	40
g_{41}	0	0	0	54	0	1	4	12	31	6	2	8	19	19	6	3	6	12	16	17
g_{51}	0	0	0	0	29	0	1	3	6	19	5	1	7	6	10	4	2	5	4	14
g_{12}						58	0	0	0	0	3	21	41	40	20	8	14	26	38	44
g_{22}		14	91	20	4	1	130	0	0	0	3	21	41	40	20	8	14	26	38	44
g_{32}		3	36	114	12	3	0	0	168	0	8	33	49	64	14	7	13	43	74	31
g_{42}		0	3	30	31	6	0	0	0	70	1	13	26	20	10	3	9	16	23	19
g_{52}		0	3	5	6	19	0	0	0	0	6	3	9	4	11	5	3	9	2	14
g_{13}											21	0	0	0	0	7	2	7	1	4
g_{23}											0	73	0	0	0	5	17	16	27	8
g_{33}											0	0	135	0	0	6	14	41	48	26
g_{43}											0	0	0	144	0	4	6	26	54	54
g_{53}											0	0	0	0	86	3	1	11	14	57
g_{15}																25	0	0	0	0
g_{25}																0	40	0	0	0
g_{35}																0	0	101	0	0
g_{45}																0	0	0	144	0
g_{55}																0	0	0	0	149

単純集計

クロス集計

MCAの制約条件と最適化基準

- 制約条件

$$\mathbf{V}_k' \mathbf{1} = \mathbf{0} \quad \mathbf{Y}_k \text{ の中心化}$$

$$n^{-1} \mathbf{V}' \mathbf{D} \mathbf{V} = \mathbf{I} \quad \mathbf{Y}_k \text{ の基準化}$$

- 最大化基準

$$\phi(\mathbf{V}) = \text{tr } n^{-1} \mathbf{V}' \mathbf{B} \mathbf{V} \quad \mathbf{F} \text{ の分散の和の最大化}$$

- この解は, $\mathbf{B}\mathbf{v} = \lambda\mathbf{D}\mathbf{v}$ を満たす (中心化条件は一旦無視)。
- 解は次の基準か Burt 行列の固有値分解に含まれる。

$$\mathbf{D}^{-1/2} \mathbf{B} \mathbf{D}^{-1/2} = \mathbf{\Lambda}$$

基準化 Burt 行列 $\mathbf{D}^{-1/2}\mathbf{B}\mathbf{D}^{-1/2}$

- この行列の最大固有値は p (変数の数), 対応する固有ベクトルは $\mathbf{s} = \mathbf{D}^{1/2}\mathbf{1}$ (反応数の平方根) である。
- このベクトルは, 他の固有ベクトルと, 変数ブロックごとに直交する。これによって, (結果的に) 中心化に関する条件も満たされる。

	g_{11}	g_{21}	g_{31}	g_{41}	g_{51}	g_{12}	g_{22}	g_{32}	g_{42}	g_{52}
g_{11}	1.000	0.000	0.000	0.000	0.000	0.673	0.166	0.031	0.000	0.000
g_{21}	0.000	1.000	0.000	0.000	0.000	0.141	0.661	0.230	0.030	0.043
g_{31}	0.000	0.000	1.000	0.000	0.000	0.060	0.133	0.665	0.271	0.066
g_{41}	0.000	0.000	0.000	1.000	0.000	0.018	0.048	0.126	0.504	0.142
g_{51}	0.000	0.000	0.000	0.000	1.000	0.000	0.016	0.043	0.133	0.614
g_{12}	0.673	0.141	0.060	0.018	0.000	1.000	0.000	0.000	0.000	0.000
g_{22}	0.166	0.661	0.133	0.048	0.016	0.000	1.000	0.000	0.000	0.000
g_{32}	0.031	0.230	0.665	0.126	0.043	0.000	0.000	1.000	0.000	0.000
g_{42}	0.000	0.030	0.271	0.504	0.133	0.000	0.000	0.000	1.000	0.000
g_{52}	0.000	0.043	0.066	0.142	0.614	0.000	0.000	0.000	0.000	1.000

第1残差行列 $S = D^{-1/2} B D^{-1/2} - p s s'$

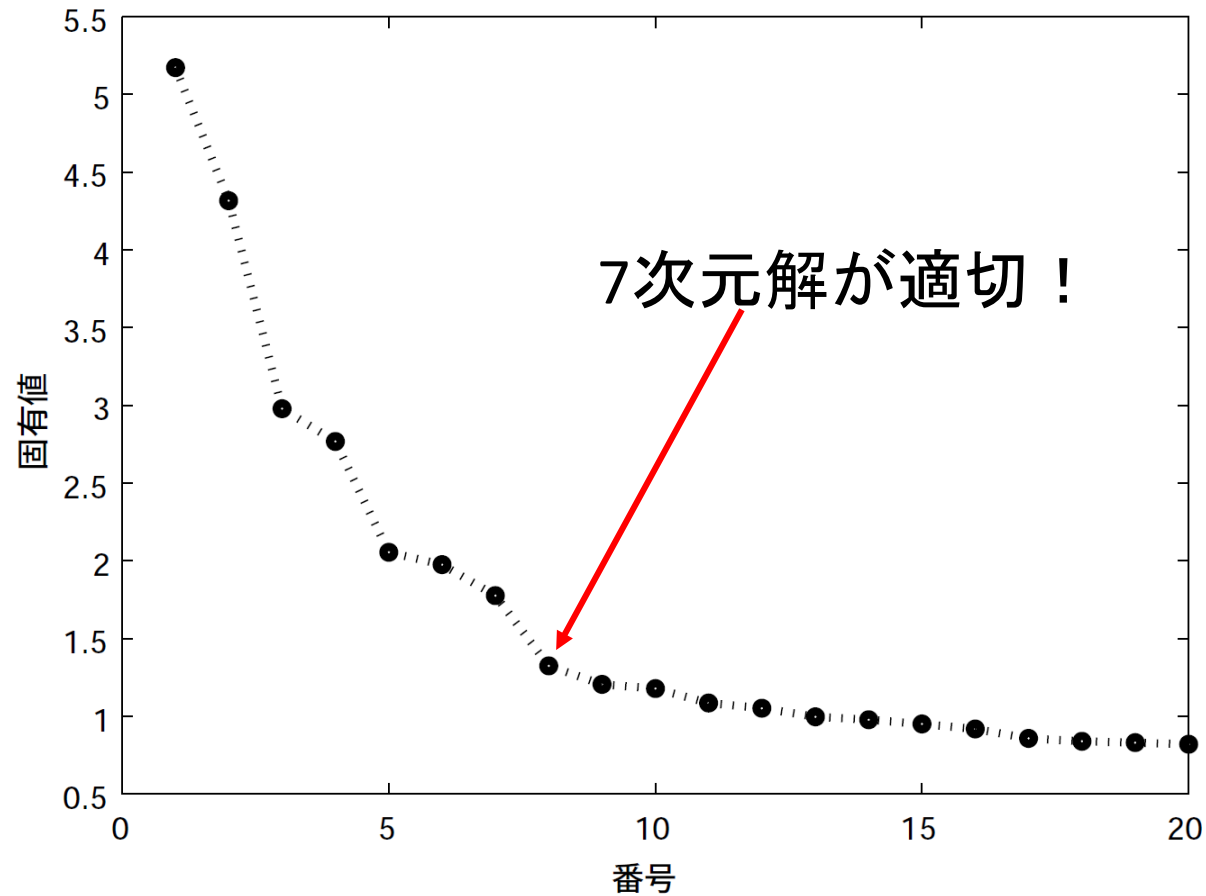
- この行列を $S = K \Lambda K'$ と固有値分解し、降順に r 番目までの固有値に対応する固有ベクトルからなる行列 $K^{(r)}$ を用いて $V = D^{-1/2} K^{(r)}$ によって、2つの制約条件を満たす数量化重み行列が得られる。

	g_{11}	g_{21}	g_{31}	g_{41}	g_{51}	g_{12}	g_{22}	g_{32}	g_{42}	g_{52}
g_{11}	0.880	-0.195	-0.214	-0.119	-0.087	0.550	-0.019	-0.178	-0.135	-0.093
g_{21}	-0.195	0.682	-0.348	-0.193	-0.142	-0.059	0.360	-0.111	-0.191	-0.108
g_{31}	-0.214	-0.348	0.619	-0.212	-0.155	-0.160	-0.196	0.291	0.030	-0.100
g_{41}	-0.119	-0.193	-0.212	0.882	-0.086	-0.104	-0.135	-0.082	0.370	0.050
g_{51}	-0.087	-0.142	-0.155	-0.086	0.937	-0.089	-0.117	-0.109	0.035	0.547
g_{12}	0.550	-0.059	-0.160	-0.104	-0.089	0.874	-0.189	-0.215	-0.139	-0.095
g_{22}	-0.019	0.360	-0.111	-0.191	-0.108	-0.189	0.717	-0.322	-0.208	-0.143
g_{32}	-0.178	-0.196	0.291	0.030	-0.100	-0.215	-0.322	0.616	-0.162	-0.105
g_{42}	-0.135	-0.191	0.030	0.370	0.050	-0.139	-0.208	-0.162	0.928	-0.105
g_{52}	-0.093	-0.108	-0.100	0.050	0.547	-0.095	-0.143	-0.162	-0.105	0.928

$$n^{-1/2} \chi_{jj'} = \frac{n_{jj'} - n_{j.} n_{.j'} / n}{\sqrt{n_{i.} n_{.j'}}$$

$$K'_k s_k = 0_{p(c-1)}$$

スクリー・プロット



それぞれの固有値に対応して、何が抽出されているのか？

個別主成分得点の正規直交化

- 変数ごとの個別主成分得点を考える。
- $\mathbf{F} = \mathbf{Y}_1 + \mathbf{Y}_2 + \cdots + \mathbf{Y}_p = \mathbf{G}_1 \mathbf{V}_1 + \mathbf{G}_2 \mathbf{V}_2 + \cdots + \mathbf{G}_p \mathbf{V}_p$
- 数量化の次元数 r がカテゴリ数マイナス1 ($c-1$) を上回っても, $c \times r$ の行列 $\mathbf{V}_k$ のランクは $c-1$ にとどまる。
- $\mathbf{V}_k$ の(結果的に $\mathbf{Y}_k$ の)1次独立な成分だけを取り出すことを考える。 $\mathbf{S}$ の固有ベクトルからなる行列 $\mathbf{K}$ の変数に対応するブロック $\mathbf{K}_k$ を(何らかの方法で)以下のように分解する。
- $\mathbf{K}_k = \mathbf{P}_k \mathbf{L}_k$ ただし, $\mathbf{P}_k' \mathbf{P}_k = \mathbf{I}_{c-1} \quad \rightarrow \text{次スライド}$
- これを用いて次のように定義する。

$$\mathbf{U}_k = \mathbf{D}_k^{-1/2} \mathbf{P}_k$$

- 数量化重み行列 $\mathbf{U}_k$ は, 中心化, 正規直交化されている。
- 変数ごとに正規直交な $n \times (c-1)$ のデータ行列 $\mathbf{Z}$ を定義する。

$$\mathbf{Z}_k = \mathbf{G}_k \mathbf{U}_k$$

K の分解と P, L の正規直交性

- K 全体としては、次のように分解される.

$$\mathbf{P} = \begin{bmatrix} \mathbf{P}_1 & \mathbf{O} & \dots & \mathbf{O} \\ \mathbf{O} & \mathbf{P}_2 & \dots & \mathbf{O} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{O} & \mathbf{O} & \dots & \mathbf{P}_p \end{bmatrix}, \quad \mathbf{L} = \begin{bmatrix} \mathbf{L}_1 \\ \mathbf{L}_2 \\ \vdots \\ \mathbf{L}_p \end{bmatrix}$$

- このとき, $\mathbf{K}'\mathbf{K} = \mathbf{I}_{p(c-1)}$, $\mathbf{P}'\mathbf{P} = \mathbf{I}_{p(c-1)}$ から、

$$\mathbf{L}'\mathbf{L} = \mathbf{I}_{p(c-1)}$$

相関行列 $\mathbf{R} = n^{-1}\mathbf{Z}'\mathbf{Z}$ のPCA

- $\mathbf{R}$ は $\mathbf{S}$ と全く同じ固有値をもつことが示される。すなわち,

$$\mathbf{L}'\mathbf{R}\mathbf{L} = \mathbf{L}'\mathbf{P}'\mathbf{S}\mathbf{P}\mathbf{L} = \mathbf{K}'\mathbf{S}\mathbf{K} = \mathbf{\Lambda}$$

- $\mathbf{W} = \mathbf{L}^{(r)}$ とすると, $\mathbf{W}'\mathbf{W} = \mathbf{I}_r$, これを制約条件として, $\ddot{\mathbf{F}} = \mathbf{Z}\mathbf{W}$ と定義し, この分散の和を最大化することを考える。

$$\text{tr } n^{-1}\ddot{\mathbf{F}}'\ddot{\mathbf{F}} = \text{tr } \mathbf{W}'\mathbf{R}\mathbf{W}$$

	z_{11}	z_{21}	z_{31}	z_{41}	z_{12}	z_{22}	z_{32}	z_{42}
z_{11}	1.000	0.000	0.000	0.000	0.733	0.005	-0.016	0.003
z_{21}	0.000	1.000	0.000	0.000	-0.043	0.583	0.039	-0.003
z_{31}	0.000	0.000	1.000	0.000	0.025	-0.026	0.472	-0.010
z_{41}	0.000	0.000	0.000	1.000	0.018	0.007	-0.060	0.339
z_{12}	0.733	0.005	0.025	0.018	1.000	0.000	0.000	0.000
z_{22}	0.000	0.583	0.026	0.007	0.000	1.000	0.000	0.000
z_{32}	0.000	0.039	0.472	0.060	0.000	0.000	1.000	0.000
z_{42}	0.003	-0.003	-0.010	0.339	0.000	0.000	0.000	1.000

ここに $\chi_{jj'}$ の情報はすべて含まれる。

カテゴリーに与えられる数量の任意性

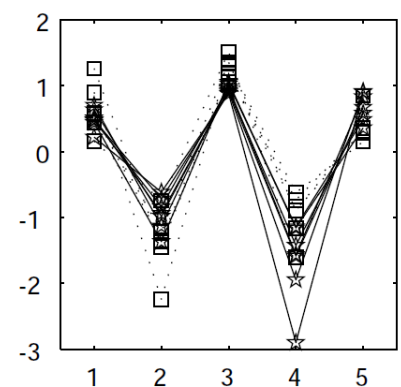
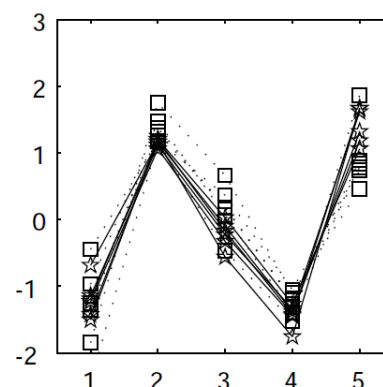
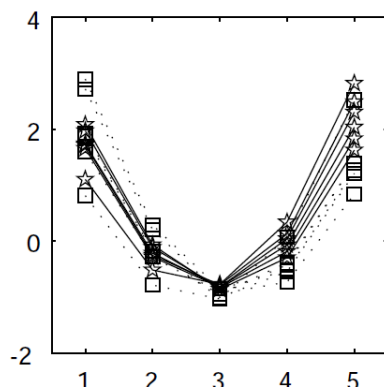
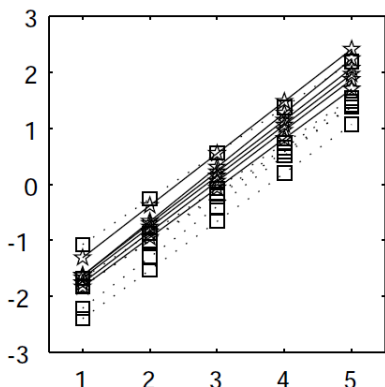
- こうして, MCAは個別数量化得点 $\mathbf{Y}_k$ を正規直交化した得点に PCA を適用することと等価であることがわかった。
- 相関行列の対角ブロックが単位行列であることを考えると, これは一般化正準相関分析 (GCA) とも見られる。これは, MCA, PCA, GCA をつなぐ missing link だった？
- しかし, $\mathbf{Z}_k$ は中心性と正規直交性を満たす限り, 全く任意であることがわかる。 $\tilde{\mathbf{T}}_1, \tilde{\mathbf{T}}_2, \dots, \tilde{\mathbf{T}}_p$ を任意の直交行列として,

$$\begin{bmatrix} \mathbf{U}_1 \tilde{\mathbf{T}}_1 & \mathbf{O} & \dots & \mathbf{O} \\ \mathbf{O} & \mathbf{U}_2 \tilde{\mathbf{T}}_2 & \dots & \mathbf{O} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{O} & \mathbf{O} & \dots & \mathbf{U}_p \tilde{\mathbf{T}}_p \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{T}}_1' \mathbf{W}_1 \\ \tilde{\mathbf{T}}_2' \mathbf{W}_2 \\ \vdots \\ \tilde{\mathbf{T}}_p' \mathbf{W}_p \end{bmatrix} = \mathbf{U} \mathbf{W} = \mathbf{V}$$

- 解釈を促進するような $\mathbf{U}_k$ の定義の仕方は何か？

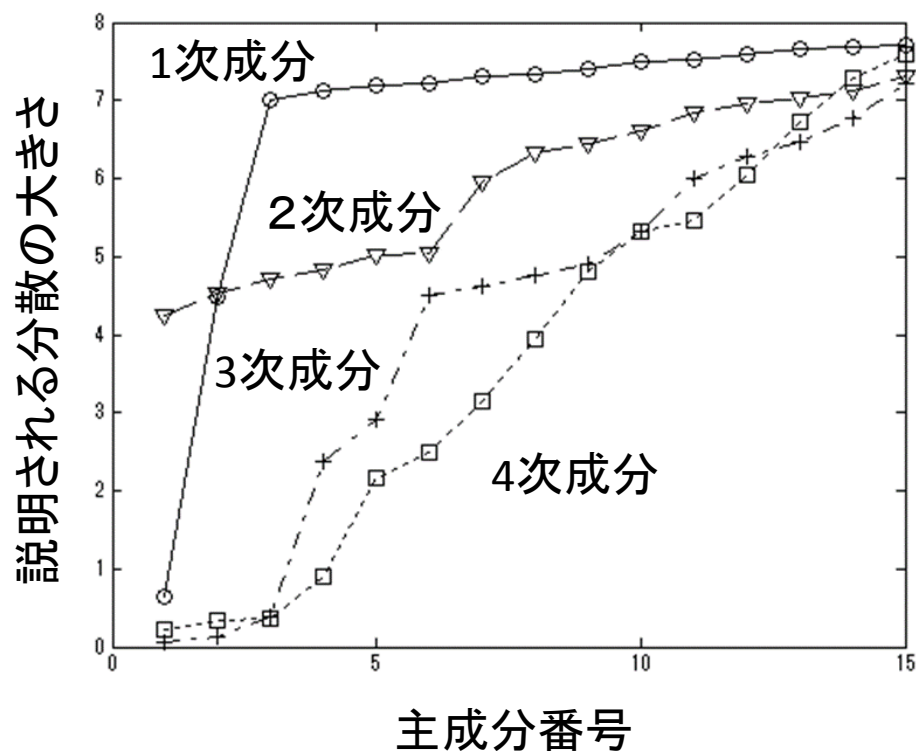
直交多項式による数量化

- 1次から $c - 1$ 次までの直交多項式によって数量を求める。
- 一般化された Gram-Schmidt の直交化法による。
- ノンメトリックな方法にメトリックな数量を持ち込むことへの反論はあろう。しかし, $c - 1$ 次までの多項式は, MCA の解に完全に埋め込まれる。
- 直交多項式という性質が明確な数量を導入することにより, すべての変数にわたって, MCA が抽出した成分の素性を記述できる. つまり直交多項式は MCA を測る gauge となり得る。



直交多項式によるgauging(例)

- 少なくとも、大きい固有値に関する限り、特定の次数の変量を集中的に説明するものとなっている。たとえば、最大固有値は2次成分(馬蹄が横倒しになる理由)、第2, 第3が1次成分に、主として関わっている。 $r = 3$ としていたら、PCAと比較可能な結果が得られていたであろう。



重み行列の直交回転:

空間中心の見方から軸単位の見方へ

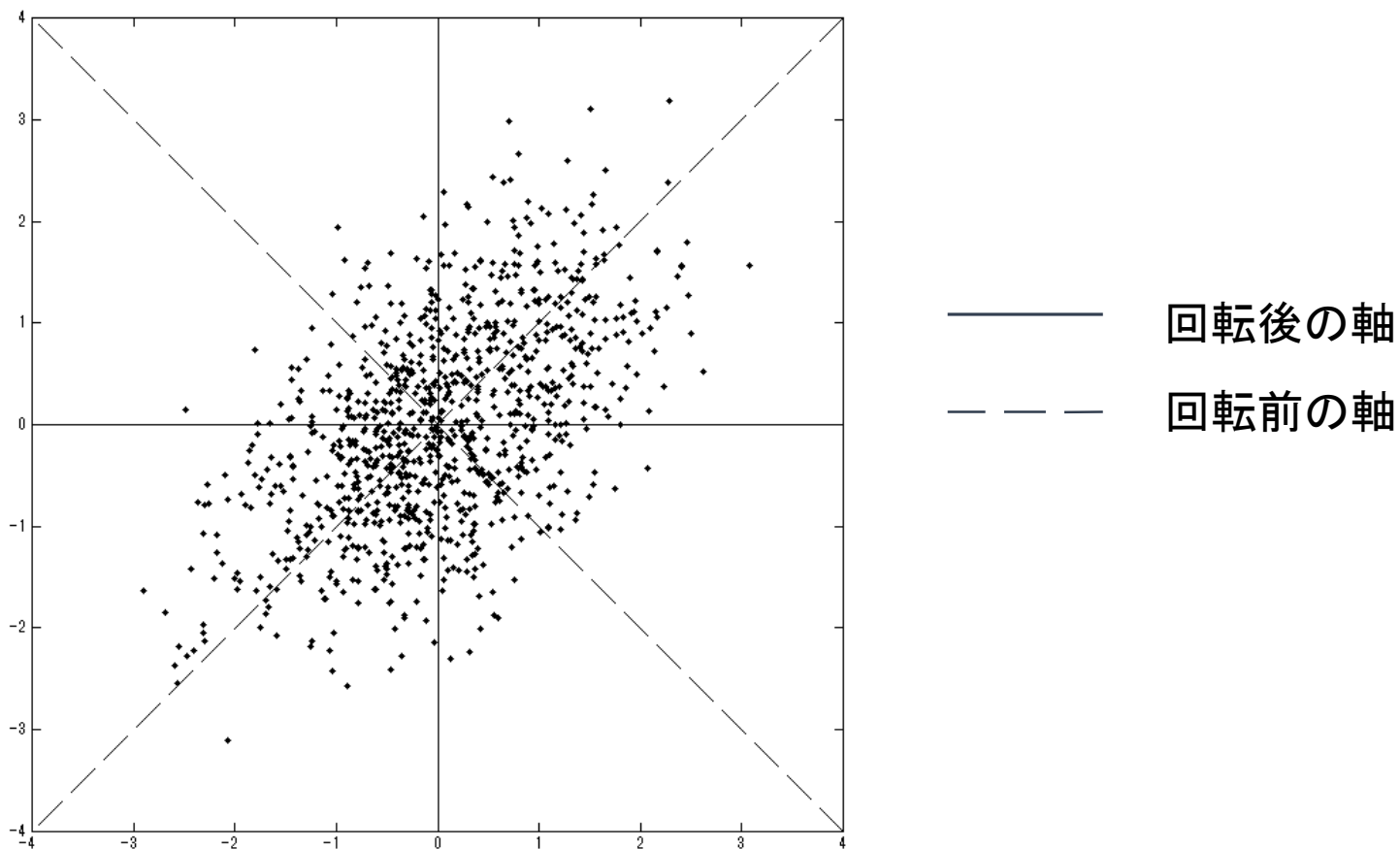
- 最大化基準の値も制約条件も, 重み行列の直交回転について不変である。T を任意の直交行列として,

$$\mathbf{T}'\mathbf{W}'\mathbf{W}\mathbf{T} = \mathbf{T}'\mathbf{T} = \mathbf{I}_r \quad \text{tr } \mathbf{T}'\mathbf{W}'\mathbf{R}\mathbf{W}\mathbf{T} = \text{tr } \mathbf{W}'\mathbf{R}\mathbf{W}\mathbf{T}\mathbf{T}' = \text{tr } \mathbf{W}'\mathbf{R}\mathbf{W}$$

- (負荷行列ではなく) 重み行列を直交回転する (Harris & Kaiser による独立クラスター回転)。
- すべての orthomax 基準は quartimax 基準に帰着
- 次で計算されるパターン行列は列ごとに直交, かつ列ごとの平方和が対応する主成分 (全体数量得点) によって説明される分散の大きさと一致 (合成変量間相関と「因子寄与」がともに得られる)。

$$\mathbf{A}_k = n^{-1}\mathbf{Z}'_k\mathbf{F}\mathbf{T}\left(n^{-1}\mathbf{T}'\mathbf{F}'\mathbf{F}\mathbf{T}\right)^{-1}\left(\text{dg } n^{-1}\mathbf{T}'\mathbf{F}'\mathbf{F}\mathbf{T}\right)^{1/2} = \mathbf{W}\mathbf{T}\left(\mathbf{T}'\mathbf{\Lambda}^{(r)}\mathbf{T}\right)^{1/2}$$

重みの回転で主成分間に相関が生じる理由



OPPCAの手続き(フロー)

- 1) 欠損値をリストワイズに除去した上で，入力 $\mathbf{X}$ をダミー変数行列 $\mathbf{G}$ に変換する．
- 2) Burt 行列 $\mathbf{B} = \mathbf{G}'\mathbf{G}$ とその対角要素によって，カテゴリーへの反応数 $\mathbf{d}$ を求める．
- 3) 直交多項式による重み行列 $\mathbf{U}$ を求める．
- 4) $\mathbf{R} = n^{-1}\mathbf{UBU}$ によって相関行列を求める．
- 5) $\mathbf{R}$ の全固有値と対応する固有ベクトルの行列 ($\mathbf{\Lambda}, \mathbf{K}$) を求める．
- 6) PCA 重み行列 $\mathbf{W} = \mathbf{K}^{(r)}$ を Quartimax 回転して $\mathbf{WT}$ とする．
- 7) 負荷行列，主成分間相関行列，重相関係数の 2 乗などを算出する．
- 8) $\mathbf{F} = \mathbf{GUWT}$ によって主成分得点を求め，ヒストグラム，散布図などを描く．

主成分パターン行列(上半分)

項目		I	II	III	IV	V	VI	VII	R^2	
1.	うれしい気持ち	1次	0.83	0.04	-0.01	0.01	-0.06	0.01	-0.03	0.70
		2次	0.01	0.63	-0.03	0.01	0.14	-0.04	0.01	0.50
		3次	-0.04	-0.04	-0.68	-0.05	0.08	0.08	0.04	0.47
		4次	0.00	-0.26	-0.06	0.07	0.15	0.02	0.53	0.32
2.	満足な気持ち	1次	0.80	-0.02	-0.02	-0.11	-0.02	-0.01	-0.02	0.68
		2次	0.00	0.63	0.03	0.01	0.16	-0.04	-0.01	0.51
		3次	-0.06	-0.02	-0.66	-0.09	0.12	0.03	0.00	0.45
		4次	-0.03	-0.29	-0.07	0.01	0.18	0.06	0.50	0.31
3.	いらいらした気持ち	1次	0.01	0.07	0.03	0.72	0.01	0.02	0.02	0.54
		2次	-0.01	0.04	-0.09	-0.05	0.54	0.04	0.06	0.33
		3次	-0.06	0.01	0.12	-0.07	-0.01	-0.55	0.00	0.32
		4次	0.03	0.34	-0.04	-0.08	-0.24	0.03	0.31	0.22
4.	楽しい気持ち	1次	0.84	0.01	0.03	-0.01	0.03	0.02	-0.02	0.71
		2次	-0.03	0.76	-0.04	0.02	0.03	-0.02	-0.04	0.59
		3次	0.04	0.02	-0.69	0.05	-0.10	0.03	0.07	0.48
		4次	0.03	-0.17	0.07	0.06	0.18	-0.04	0.54	0.35
5.	心配な気持ち	1次	-0.02	0.03	-0.01	0.70	-0.02	0.04	-0.02	0.50
		2次	0.00	0.09	0.04	-0.01	0.60	0.09	-0.06	0.42
		3次	0.00	0.06	0.01	0.02	0.03	-0.44	-0.01	0.20
		4次	-0.03	0.36	-0.03	0.00	-0.27	0.19	0.33	0.27
6.	面白い気持ち	1次	0.84	-0.07	0.04	0.06	0.11	0.05	-0.06	0.69
		2次	-0.02	0.73	0.03	0.01	0.03	-0.07	-0.07	0.55
		3次	0.05	0.15	-0.63	0.05	-0.20	0.01	-0.01	0.45
		4次	0.06	0.02	0.06	0.00	0.04	-0.07	0.45	0.24

ポジティブ項目とネガティブ項目がそれぞれ1次, 2次, 3次主成分をもっている。

主成分パターン行列(下半分)

		1次	2次	3次	4次	5次	6次	7次	8次
7. 悲しい気持ち	1次	0.15	-0.05	-0.09	0.71	-0.04	-0.05	0.10	0.51
	2次	0.05	0.12	0.11	0.10	0.54	-0.05	0.08	0.42
	3次	0.10	-0.04	-0.10	-0.05	0.07	-0.52	-0.04	0.31
	4次	0.05	0.29	0.00	-0.06	-0.30	0.00	0.34	0.22
8. 充実した気持ち	1次	0.80	0.03	-0.01	-0.02	0.04	-0.03	0.05	0.66
	2次	-0.01	0.76	0.01	0.02	0.03	0.03	-0.13	0.59
	3次	0.04	-0.06	-0.61	0.08	0.02	-0.07	-0.08	0.40
	4次	-0.05	-0.08	0.02	0.04	0.06	-0.17	0.50	0.29
9. 憂うつな気持ち	1次	-0.16	0.04	0.01	0.72	-0.01	-0.01	-0.01	0.58
	2次	0.02	0.06	0.03	-0.04	0.64	0.04	0.07	0.47
	3次	0.00	0.11	0.08	-0.02	-0.06	-0.59	-0.08	0.36
	4次	-0.11	0.25	-0.15	-0.13	-0.23	0.01	0.41	0.25
10. 安心した気持ち	1次	0.70	-0.01	-0.03	0.07	-0.11	-0.04	0.12	0.53
	2次	0.07	0.60	-0.05	0.03	0.20	0.03	0.00	0.52
	3次	-0.02	-0.02	-0.59	-0.03	0.06	-0.12	-0.06	0.39
	4次	-0.04	-0.16	-0.01	-0.05	0.25	-0.02	0.38	0.21
11. 満たされない気持ち	1次	-0.05	0.02	0.05	0.71	-0.01	0.03	-0.05	0.52
	2次	-0.06	0.18	-0.03	0.02	0.57	0.02	0.04	0.47
	3次	0.00	-0.08	0.00	0.05	0.00	-0.60	0.10	0.39
	4次	0.02	0.24	0.14	-0.04	-0.22	0.02	0.29	0.17
12. 腹立たしい気持ち	1次	0.03	-0.04	0.00	0.83	0.01	-0.01	0.00	0.68
	2次	0.03	0.04	-0.03	-0.05	0.74	-0.02	0.07	0.58
	3次	-0.05	0.04	-0.08	0.04	-0.05	-0.63	0.05	0.41
	4次	0.03	0.33	0.13	-0.02	-0.27	0.08	0.40	0.31
主成分分散		3.73	4.01	3.34	2.66	2.28	2.03	2.99	21.04

4次成分は、すべての変数に共通、やや独自の働きもしている。

主成分間相関行列

- ほぼ独立した7つの軸が抽出される。
- 2次成分のみ, やや高い相関がある。内容によらない極端反応の影響と見られる。

主成分間相関係数

I	1.00	0.05	-0.07	-0.15	-0.06	0.02	0.04
II	0.05	1.00	0.02	0.12	0.44	0.04	0.13
III	-0.07	0.02	1.00	0.04	0.04	0.04	0.13
IV	-0.15	0.12	0.04	1.00	0.09	-0.07	0.04
V	-0.06	0.44	0.04	0.09	1.00	0.02	0.13
VI	0.02	0.04	0.04	-0.07	0.02	1.00	0.01
VII	0.04	0.13	0.13	0.04	0.13	0.01	1.00

単純集計,メトリックPCAとの比較など

- 1次変量に対応する主成分だけを取り出せば,メトリックPCAとの差はほとんどない。

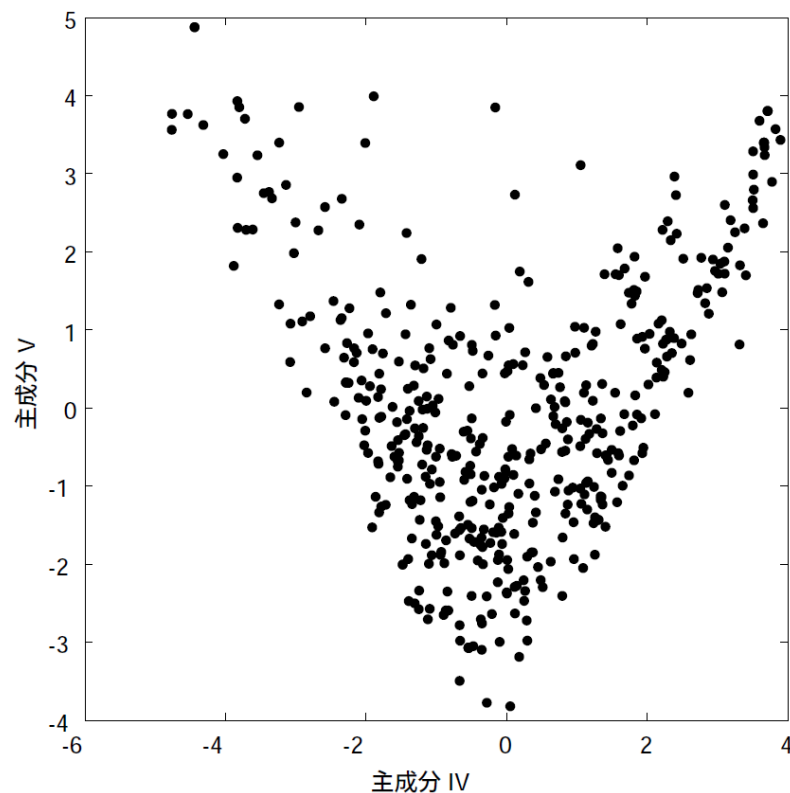
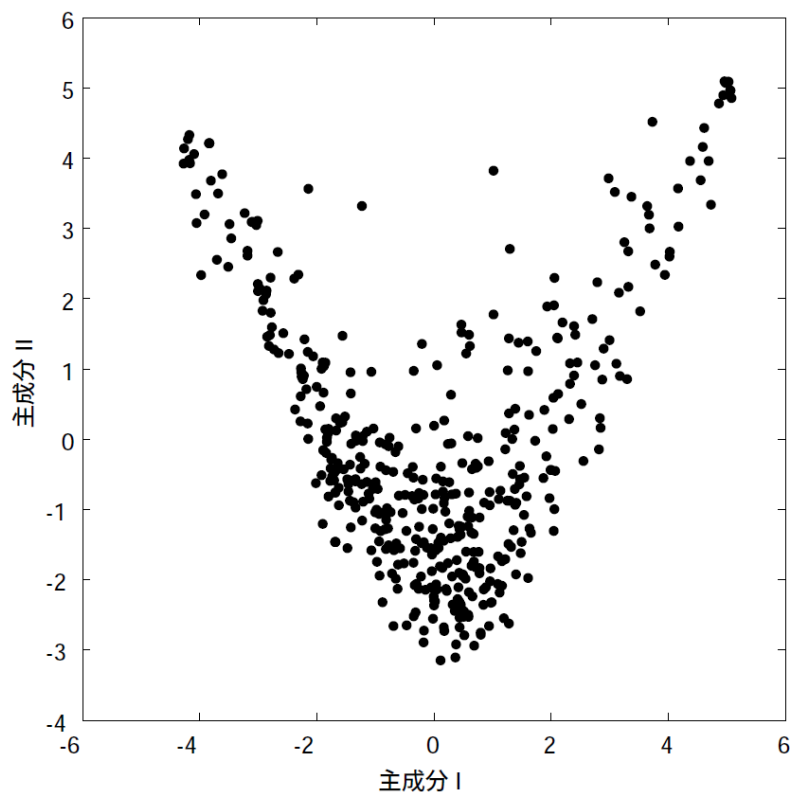
	頻度 (d_k')					PCA (Varimax)			OPPCAの1次主成分	
	1	2	3	4	5	I	II	R^2	I	IV
うれしい気持ち	55	146	175	54	29	0.84	-0.05	0.71	0.83	0.01
満足な気持ち	58	130	168	70	33	0.80	-0.18	0.68	0.80	-0.11
いらいらした気持ち	21	73	135	144	86	-0.04	0.74	0.55	0.01	0.72
楽しい気持ち	44	113	176	78	48	0.84	-0.06	0.70	0.84	-0.01
心配な気持ち	25	40	101	144	149	-0.08	0.70	0.50	-0.02	0.70
面白い気持ち	40	99	164	98	58	0.82	0.00	0.68	0.84	0.06
悲しい気持ち	151	127	94	59	28	0.11	0.69	0.50	0.15	0.71
充実した気持ち	54	116	179	69	41	0.81	-0.06	0.66	0.80	-0.02
憂うつな気持ち	45	87	117	118	92	-0.22	0.73	0.58	-0.16	0.72
安心した気持ち	103	146	154	31	25	0.71	0.01	0.50	0.70	0.07
満たされない気持ち	47	80	135	121	76	-0.11	0.71	0.52	-0.05	0.71
腹立たしい気持ち	61	78	153	92	75	-0.04	0.82	0.68	0.03	0.83
主成分分散						3.96	3.28	7.24		
主成分間相関係数										
PCA (Varimax)										
I						1.00	0.00		1.00	-0.08
II							1.00		-0.07	0.99
OPPCA										
I									1.00	-0.15
II										1.00

集団の構造だけを考える限り、
伝統的な扱いで、ほぼ問題が
ない。しかし・・・。

集団の構造だけを考える限り,伝統的な扱いで, ほぼ問題がない。しかし・・・。

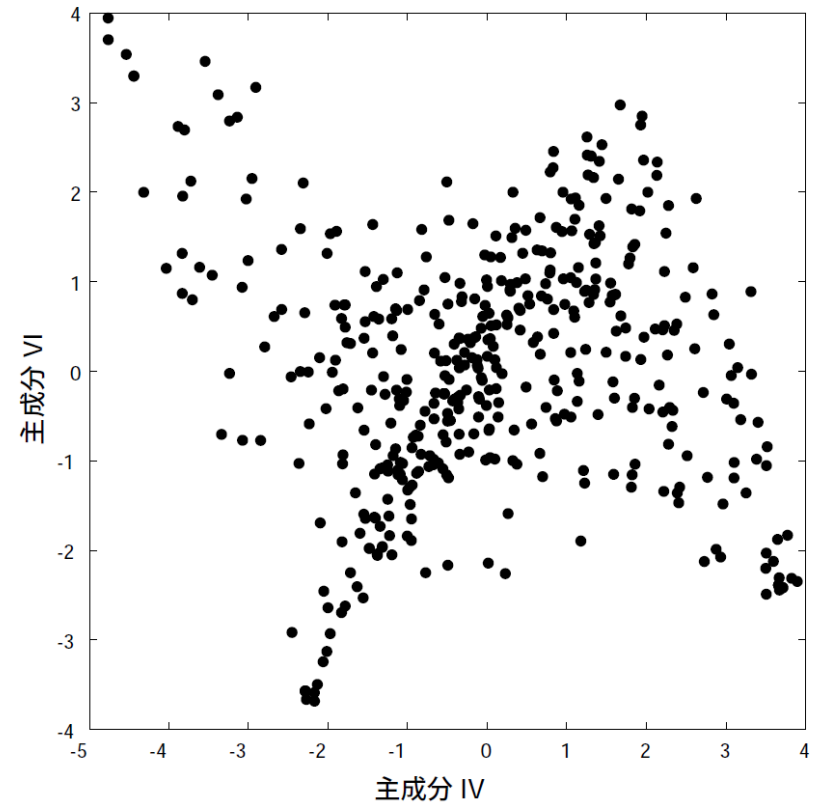
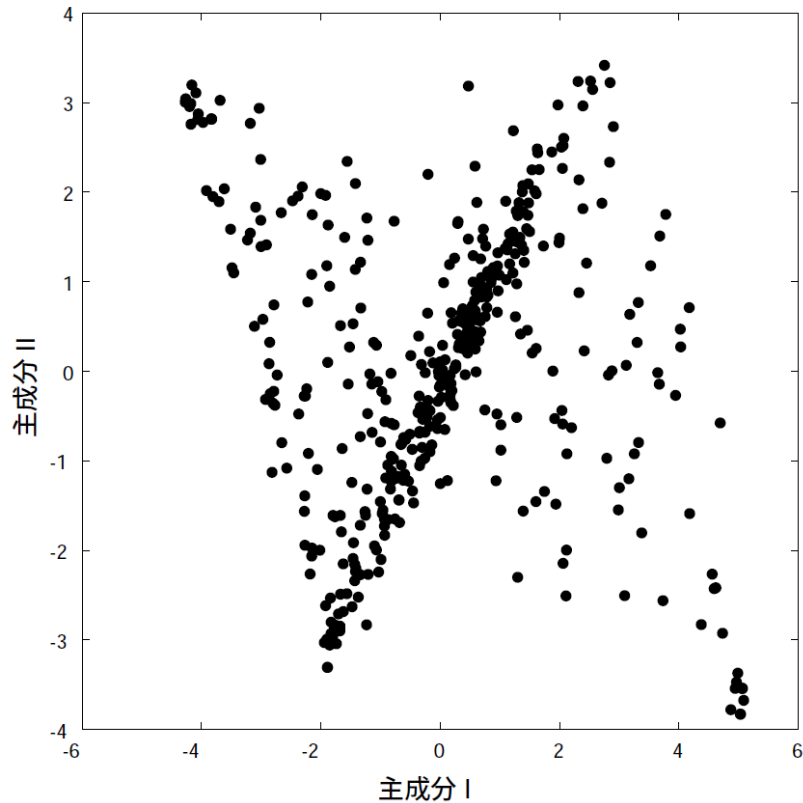
1次成分と2次成分

- 主成分Ⅱと主成分Ⅴの間には約 0.44 の相関がある。測定内容と独立な反応傾向と見られる。
- 1次成分間の相関にいくらか影響するケースも。



1次成分と3次成分

- N字の屈曲点あたりのケースは、やはり気になる。
- Screening, assessment 目的の使用では要注意。

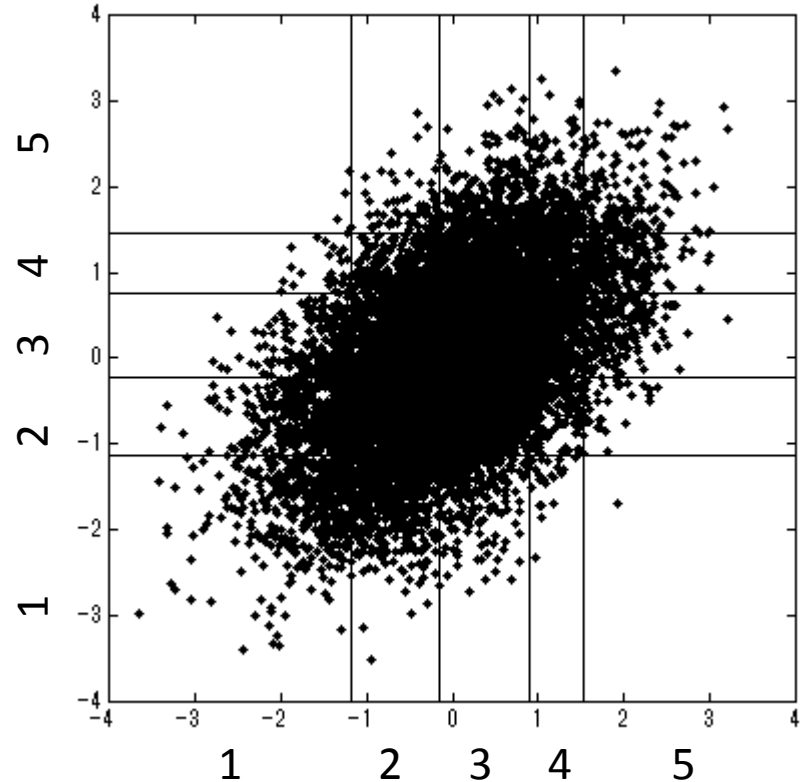


OPPCAの特徴

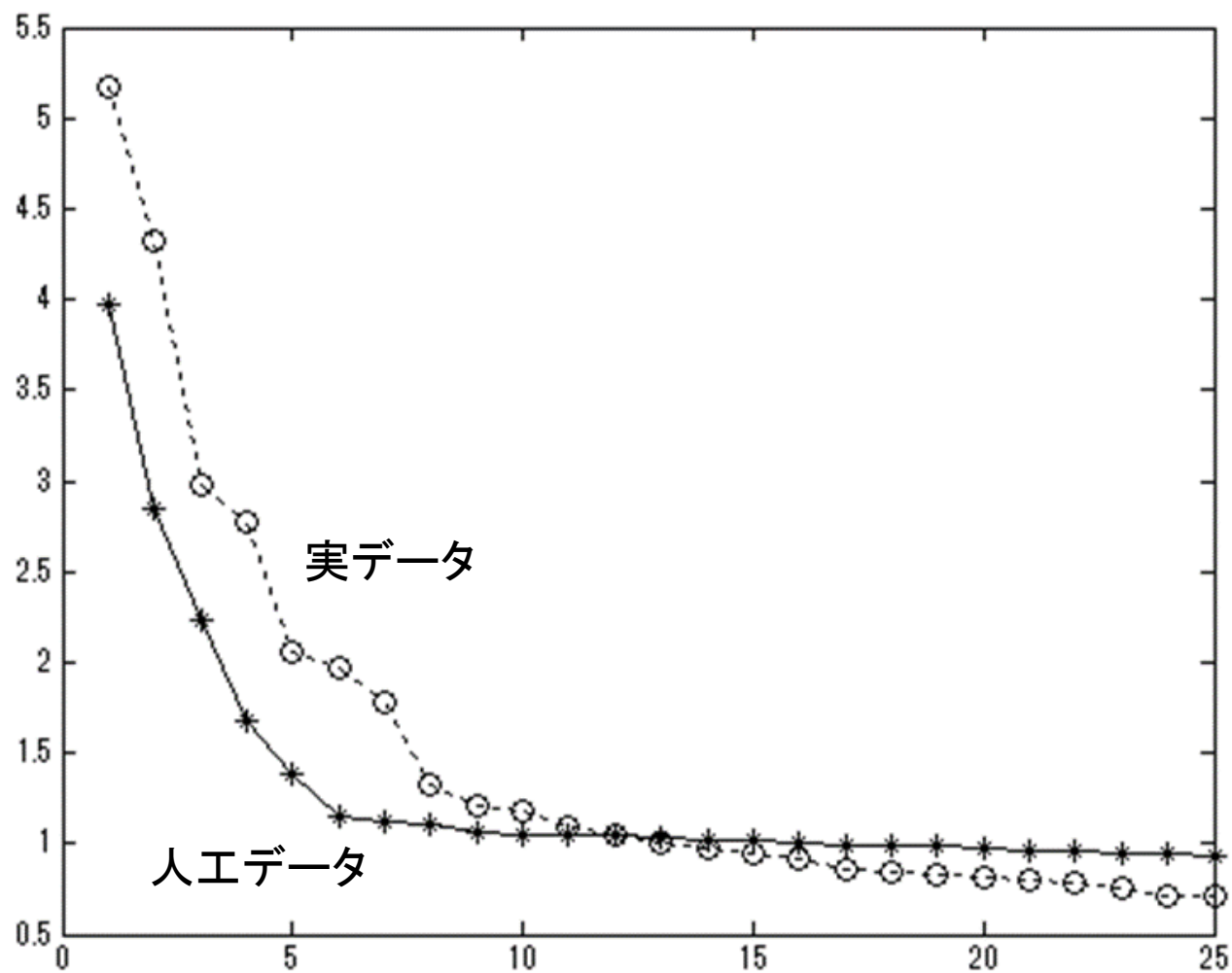
- あくまでも MCA に許容される変換の範囲
- アルゴリズムの頑健性. 固有値計算と Quartimax 回転以外の反復計算がない.
- 初等的な線型代数と三角関数の範囲で理解可能である.
- MCAについて, 理論的, 経験的に蓄積された成果が利用できる.

シミュレーション

- 問い: やはり2次以上の成分は余剰次元ではないのか?
- 実データと同一の相関行列をもつ正規乱数を生成
- 実データと同一の度数分布をもつカテゴリカル変数に変換
- OPPCAを適用



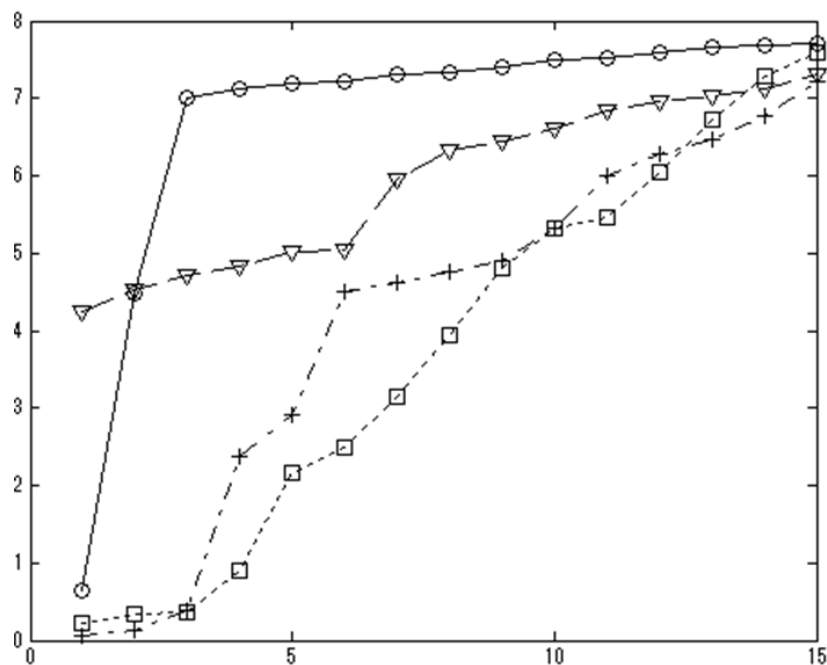
スクリーンプロット



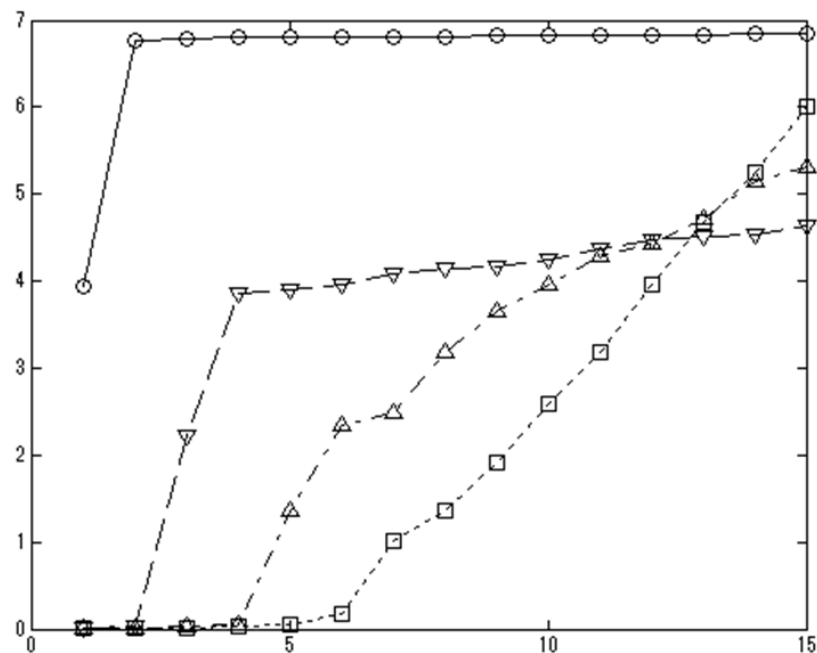
Gauging

多変量正規分布からは、1次、2次、3次と低次成分から順に出現する。

他方、実データの2次、3次の成分には、真の個人差も含まれると考えられる。

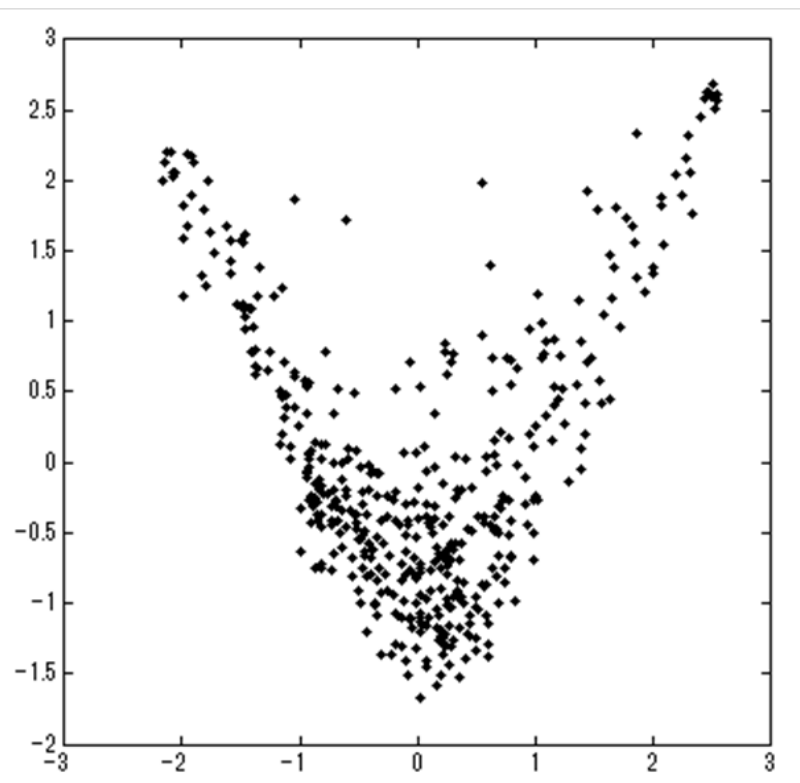


実データ

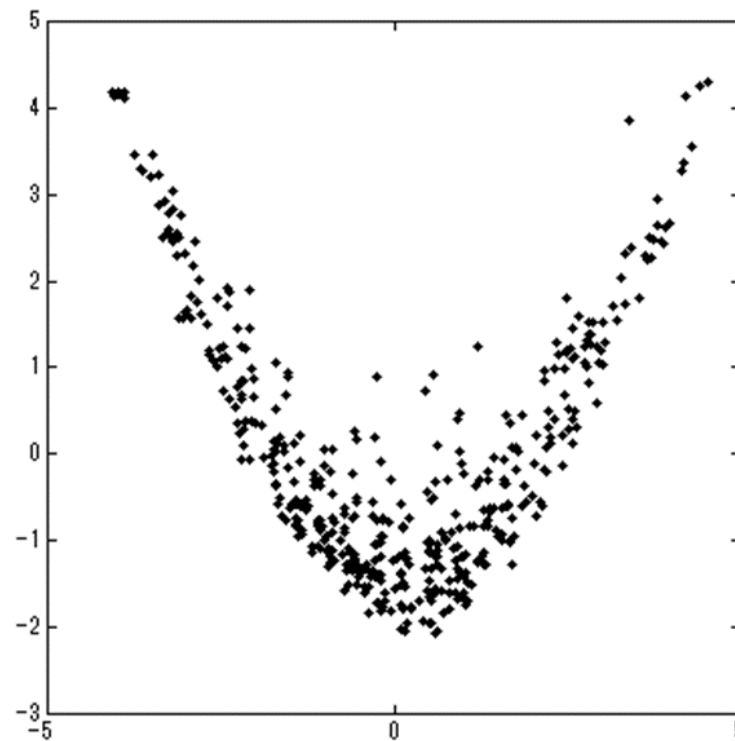


多変量正規分布から作成した人工データ

1次-2次主成分得点プロット



実データ



シミュレーション

ほとんど同じに見えるが、2次成分間に相関がないことと、多次元正規分布の方が、極端な値が生じていることが違う。

問いへの答え(1)

Q1. 順序尺度に過ぎない Likert 型項目を間隔尺度として扱っていいのか？(1と2, 2と3の間隔が同じかということよりも, 1と5にばかり○をつける個人と3の周辺でほとんど動かない個人の差は？ 効用の個人間比較の不可能性の問題)

A1. 相関研究に用いる限り, 等間隔としての扱っても大きな問題を引き起こすことはなさそうである。しかし, assessment の道具として考えるなら, 控えめながら一貫した反応をしている回答者のかかえている「問題」を見逃す危険性はある。

問いへの答え(2)

Q2. メトリックな探索的「因子分析」(exploratory “factor analysis”; EFA)によって、多くに成果が得られている。しかし、ノンメトリックな多重対応分析 (multiple correspondence analysis; MCA)の結果とは著しく違う。なぜか？

A2. 実際には MCA の解がメトリックな探索的「因子分析」の解を含んでいる場合が多いようである。ただ、3次元以上のMCAの解を検討することがあまり行われなかったために、そのことに気づかれなかったのだと思われる。

問いへの答え(3)

Q3. MCAの結果を, EFA 風 to 負荷行列を回転して解釈することはできないのか？

A3. ここで述べたような若干の工夫をすることによって可能である。これによって, カテゴリーコードに対して非線型な成分を分離して, 通常のEFAの結果を取り出すことができる。また, 通常の MCA の解の graphical な表示では困難であった4次元以上の解の様相についても知ることができる。

結びに替えて: Likert 尺度の特異性

- Likert 型項目の集合である Likert 尺度は、順序のついたカテゴリカル変数としては、非常に特異なものである。
- 冗長性の高さ: 同じ特性に関わる質問が、wording を変えて繰り返される。
- カテゴリー判断の主観性: カテゴリーの境界は、個人の判断に任される。
- 事実に関わる問い、年齢段階(10歳刻み)や年収(近似的対数尺度?)などとは異なる。
- たとえば、社会学者にとって妥当性とは、回答者が本当のことを答えているかどうか、ということにほぼ限定されているようだ。
- 心理学者にとっては、抽象概念としての特性(構成概念)の測定が目標となり、冗長性は必須であるとともに、この種の分析以降の扱いが問題になる。