*Bulletin of the Faculty of Education,
Nagoya University (Educational Psychology)
1985, Vol. 32, 75 – 93.*

多集合データのための探索的因子分析*

村　上　　隆

1. 序—問題の所在

図1のような項目からなる調査を考える。** 一応，この全体を分析の対象とするとして，これらの項目は4つの部分からなっている。現在研究は探索的な段階にあり，この調査から得られる結果について何ら明確な仮説はないものとする。例えば，各部分に幾つの，どのような内容の因子が潜在しており，それらの間にどのような関係があるかは，わかっていないものとする。

このとき，とりあえずの分析として，ここに含まれる計64（20＋10＋27＋7）の項目全体を一挙に因子分析することには，多くの人はためらいを感ずるのではなかろうか。そうであるとすれば，その理由はどこにあるのであろうか。

それは，恐らく，これらの4つの問（項目のグループ）が，それぞれ独立した内容領域にかかわるものであり，相互に関連はあるとしても，各グループごとに，因子が抽出されるべきであると考えられるせいであろう。領域間の関係は，それにもとづいて改めて（因子得点間相関等により）検討される。

このようなデータを多集合データ（multiset data）と呼ぼう。“集合”という語は，図1の質問から得られたデータで言えば，問，すなわち項目のグループを指す（表1参照)。

大部分の質問紙調査は，この多集合データの形態をとるものと考えられる。図1にあげたようなもの以外にも例えば，

1）中学生の母親イメージと母親の養育態度に対する認知，
2）高校生の職業興味と性格，
3）何らかの対象に対する態度における認知的側面，感情的側面，行動的側面，

といった例がただちに思い浮ぶ。

このような多集合データの探索的な分析方法は，従来次の2つの方法のうちのどちらかであったと思われる。

a）各集合ごとに個別の因子分析（または主成分分析）を適用し，因子得点（主成分得点），または何らかの合成得点によって集合間の相関関係を検討する。
b）データ全体を，正準相関分析し，集合間の相関を最大化するような合成得点を求める。

これらの方法は，ともに問題を含んでいると考えられる。まず，aの方法では，専ら各集合内の内部相関に重点がおかれ，集合間の関係が（因子抽出の段階では）無視される結果，実際には存在するかもしれない集合間の興味ある関係が見逃される可能性がある。

bのやり方では，まず，通常の正準相関分析は，集合の数が2に限定されてしまうことが問題になる。ただしこの点では，集合の数が3以上の場合に対する一般化が幾つか提唱されてはいる（例えば，Horst　1965や，Kettenring　1971等参照)。しかし，それ以上に，正準相関分析の結果は，多くの場合解釈が困難である。更にこの方法から得られる合成変量は，集合間相関のみによって求められる結果，因子分析の場合とは逆に内的整合性に著しく欠けることになる。(このことが，結果の解釈が困難になる1つの理由であろう。)

古典的テスト理論の観点からすれば，aは測定の信頼性のみを重視するやり方であり，bの方法は，測定の（基準関連）妥当性にウエイトのかかりすぎる方法である。この問題を回避するための1つの方法は，最初に少し触れたように，多集合データに含まれる全項目間の相関行

＊　本研究の数値計算は名古屋大学大型計算機センターFACOM　M—382によった。

＊＊　これらは，昭和58年度に，経済企画庁の委託調査として愛知県物価対策課（現消費生活課）および愛知学院大学林英夫助教授（現関西大学教授）と筆者によって行われた「消費者の購買行動と物価情報に対するニーズに関する調査」の一部である。データの使用を許可していただいた林教授，並びに関係各位に深く感謝する。調査の全容は，同名の報告書（愛知県　1983）に詳しい。

問1　以下の項目はどれも，主婦のかたがたを対象とした日常の買い物についてのアンケート結果の一部です。あなたご自身はこれらの結果にどの程度似ているとお感じになりますか。**（各項目について１つづつ）**

	よく似ている	かなり似ている	どちらともいえない	あまり似ていない	ぜんぜん似ていない
1）日常の買い物には，予算の腹づもりをして出かける。	1	2	3	4	5
2）近所の人や知人らから，商品やお店など値打ちな買い物について話を聞く。	1	2	3	4	5
3）日用品や食料品などは，特売や安い時にまとめ買いをする。	1	2	3	4	5
4）加工食品などの製造年月日や内容の表示などを確かめてから買う。	1	2	3	4	5
5）日用品や食料品などは　単位価格（たとえば100 *g* 当りいくら）を見て，もっとも経済的なものを買う。	1	2	3	4	5
6）高額のファッション品などを選ぶ時は，値段は二の次にして好みに合わせて買う。	1	2	3	4	5
7）スーパーの店頭で予定になかったものを買ってしまう。	1	2	3	4	5
8）生鮮食料品は出盛りで，一番安い旬（しゅん）になってから買う。	1	2	3	4	5
9）安くて良い商品を売っている店を探す。	1	2	3	4	5
10）近所の人と共同で買い物をして分けあう。	1	2	3	4	5
11）新聞広告や折込み・チラシ広告を見て買い物をする。	1	2	3	4	5
12）自分にとって本当に必要な物を必要な時に必要なだけ買う。	1	2	3	4	5
13）家で使う主な品目の買い物計画リストを作り，なくなりかけたものを追加して買う。	1	2	3	4	5
14）たくさんの人たちが行列しているような売場で買う。	1	2	3	4	5
15）肉や食肉加工品を買う場合「はかり売り」よりも「パック売り」のものを買う。	1	2	3	4	5
16）買い物をしてレシートをもらっても，すぐすててしまう。	1	2	3	4	5
17）新聞記事や「暮しの手帖」などですすめている商品を選ぶ。	1	2	3	4	5
18）実物を直接に見ないと納得できないので，テレビショッピングや通信販売は利用しない。	1	2	3	4	5
19）日常の買い物以外の大きな買い物は，家族と相談のうえどれを買うか決める。	1	2	3	4	5
20）生鮮野菜や鮮魚などは，値段が高くても品質のよいものを買う。	1	2	3	4	5

問3　わたしたちが，よりよい暮しをしていくためには，いろいろな情報を通じて知識を得ることが必要だと思われます。次にあげたような情報は，現在，あなたご自身にとって，十分な状態にあるとお考えになりますか。それとも不十分でしょうか。それぞれで，あなたのお感じにもっとも近いものをお選びください。**（各項目について１つづつ）**

	十分であり関心がある	不十分であるが関心はある	関心はない
1）食品の安全性についての記事や報道	1	2	3
2）商品を長持ちさせたり保存したりする知識の提供	1	2	3
3）生鮮食料品の今日，明日の市場価格のお知らせ	1	2	3
4）教育費の負担状況	1	2	3
5）石油製品（灯油，ガソリン，プロパンガス）の価格の動き	1	2	3
6）国や県・市などで実施する各種の調査の結果	1	2	3
7）生鮮食料品などの基礎的な商品知識や流通の仕組みのやさしい解説	1	2	3
8）公共料金の改定のお知らせ	1	2	3
9）省エネルギーのための知識	1	2	3
10）生鮮食料品の価格が高い時，その代わりになる品目の価格などのお知らせ	1	2	3
11）県や市町村が主催する食品などの安売りデー，青空市場，朝市のお知らせ	1	2	3
12）賃金やボーナスなどの見通し	1	2	3
13）よりよい暮しをするために役立つアイデア，ヒントや生活の知恵	1	2	3
14）生鮮食料品の値動きの向う１ケ月程度の見通し	1	2	3
15）年金制度についてのやさしい解説	1	2	3
16）商品テストの結果	1	2	3
17）商品価格の安い商店街，ショッピングセンター，市場などの解説	1	2	3
18）サービス料金の動向の解説	1	2	3

問2　あなたのふだんの生活の仕方についておたずねします。以下に2つの相反する文章が対になってあげてありますが，それぞれであなたはどちらの方に近いとお感じになりますか。**（各項目について1つづつ）**

A	Aの方	どちらかといえばAの方	どちらともいえない	どちらかといえばBの方	Bの方	B
1）将来のため・貯蓄や保険にお金をまわすようにしている。	1	2	3	4	5	現在の生活を，より豊かにするためにお金を使うようにしている。
2）欲しいものは借金をしてでも買う。	1	2	3	4	5	欲しいものは貯蓄してから買う。
3）近所の人が持っているものを買うことが多い。	1	2	3	4	5	近所の人が持っていないものを買うことが多い。
4）現金買いをすることが多い。	1	2	3	4	5	月賦を利用することが多い。
5）値切って買うことが多い。	1	2	3	4	5	言い値で買うことが多い。
6）日常の買い物は節約して耐久消費財のようなものを買う時どんとお金を使う。	1	2	3	4	5	耐久消費財のようなものを買う時は節約して，日常の買い物にどんとお金を使う。
7）いつも同じ銘柄のものを買うようにしている。	1	2	3	4	5	銘柄を気にせず買い物をしている。
8）衣服はコーディネイト（全体の組合せと調和）できるよう単品をいくつか購入する。	1	2	3	4	5	衣服は単品よりもワンピース，ツーピース，スーツなど上下揃いのものを買う。
9）買い物をする時は，いつも決まった店でする。	1	2	3	4	5	買い物をする時は，いろいろな店でする。
10）定期的に買うものの値段が変わるとすぐ気がつく。	1	2	3	4	5	定期的に買うものの値段が変わってもすぐには気がつかない。

19）一般の消費者の方々の物価，消費生活に対する意見，お考えの紹介	1	2	3
20）冠婚葬祭の知識や世間のとおり相場	1	2	3
21）国の経済政策や国際経済の情勢についてのやさしい解説	1	2	3
22）旬（しゅん）のものを使った料理方法	1	2	3
23）景気や消費生活の見通し	1	2	3
24）バーゲンや特売の情報	1	2	3
25）老後の生活設計のための助言	1	2	3
26）県や市が開催する物価教室，消費生活講座などの案内	1	2	3
27）新製品の紹介	1	2	3

問4　次にあげたような手段によって得られる商品やサービスについての知識や情報はどの程度信頼できるとお考えですか。**（各項目について1つづつ）**

	非常に信頼できる	かなり信頼できる	やや信頼できる	どちらとも言えない	やや信頼できない	あまり信頼できない	ぜんぜん信頼できない
1）新聞・雑誌・テレビ・ラジオなどの広告	1	2	3	4	5	6	7
2）広告以外の新聞・雑誌の記事やテレビ・ラジオ番組	1	2	3	4	5	6	7
3）新聞の折込み・チラシ広告	1	2	3	4	5	6	7
4）消費者雑誌（「月刊消費者」や「暮しの手帖」など）の記事	1	2	3	4	5	6	7
5）近所の人や知人らの話	1	2	3	4	5	6	7
6）店の人の話	1	2	3	4	5	6	7
7）県・市町村などによる広報紙の記事	1	2	3	4	5	6	7

図1　多集合データを生み出す質問紙（例）

表１　多集会データ

個人番号	問　1	問　2	問　3	問　4
1）	24311312334445533111	2532322323	2232223223222222333322223232	2232443
2）	44112331141252341113	1531515121	2223222112122331222222221222	2121231
3）	22212331113323252223	3333434322	1112111211211111111111111111	5232341
4）	33211332131313341313	3333233332	1122222211121211121121111121	4452443
5）	22222232241333341133	2432323332	2222222222222222222222221212	3233233
6）	12113112451554155132	4431412112	1221213113321323333223221333	5433333
7）	22213342312324232112	3531511511	2222232112132232332332333332	2232433
8）	12111421151335521112	1531214151	1112111111121121121121111222	1112322
9）	11111333121133353211	2432442222	2222233222232223232232222223	4232432
10)	11113422121233351123	1531333331	212222112122	
11)	22111333111345353223	2531332331	2212311112　・	
12)	23112331242454222113	2551422121	123211　・	
13)	13112232222243242212	2542222522	22　・	
	・		・	
	・		・	
	・		・	

（以　下　略）

$$
\begin{bmatrix} \boldsymbol{R}_{11} & \boldsymbol{R}_{12} & \boldsymbol{R}_{12} \\ \boldsymbol{R}_{21} & \boldsymbol{R}_{22} & \boldsymbol{R}_{23} \\ \boldsymbol{R}_{31} & \boldsymbol{R}_{32} & \boldsymbol{R}_{33} \end{bmatrix} = \frac{1}{N} \begin{bmatrix} \boldsymbol{Z}_1^{\mathrm{T}} \\ \boldsymbol{Z}_2^{\mathrm{T}} \\ \boldsymbol{Z}_3^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} \boldsymbol{Z}_1 & \boldsymbol{Z}_2 & \boldsymbol{Z}_3 \end{bmatrix}
$$

図２　多集合データ $\boldsymbol{Z}$ から算出される partitioned correlation matrix $\boldsymbol{R}$

$$
\begin{bmatrix} \boldsymbol{R}_{11} & \boldsymbol{R}_{12} & \boldsymbol{R}_{13} \\ \boldsymbol{R}_{21} & \boldsymbol{R}_{22} & \boldsymbol{R}_{23} \\ \boldsymbol{R}_{31} & \boldsymbol{R}_{32} & \boldsymbol{R}_{33} \end{bmatrix} = \begin{bmatrix} \boldsymbol{A}_1 \\ \boldsymbol{A}_2 \\ \boldsymbol{A}_3 \end{bmatrix} \begin{bmatrix} \boldsymbol{A}_1^{\mathrm{T}} & \boldsymbol{A}_2^{\mathrm{T}} & \boldsymbol{A}_3^{\mathrm{T}} \end{bmatrix} + \begin{bmatrix} \boldsymbol{D}_1 & \boldsymbol{O} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{D}_2 & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} & \boldsymbol{D}_3 \end{bmatrix}
$$

図３　$\boldsymbol{R}$ 全体の因子分析の基本方程式

列（図2）を因子分析すること（図3）であろう。しかし，これもまた実際にやってみると解釈に困難をきたすことが多い。一般に集合間の項目間相関は，集合内の項目間相関より低い場合が多いから，単に集合ごとに別の因子があらわれてくるにすぎない場合が多い。また，この方法ではパラメータの数が多くなりがちで無駄が多い。

それ以上に問題になるのは，このやり方が多集合データの本来の理念に反していると思えることであろう。集合間では，いかに相互相関の高い項目でも別のものを測っているという前提が漠然とあり，これをいわば同一の次元で扱うことには抵抗があるのである。

我々の直面している事態は，Cronbach（1970）の言う bandwidth－fidelity dilemma に相当するものであり，2つの目標を同時に追求することは本質的に不可能である。どちらにより大きなウエイトを置くかによって，多くの中間的な立場が考えられる。

しかしながら，すべての項目間の相関行列を最大限に説明する，という基準はそれなりに納得のいくものではないかと思われる。

ここでは，今述べたような基準にもとづく因子分析の一方法である，多集合因子分析（multiset factor analysis）を提案する。* この方法は，端的に言えば，モデルとしては，ブロックごとの因子分析を考え，解法としては，項目間相関行列全体へのあてはめをはかるものである。（集合ごとの個別因子分析では，図2の対角部分の行列，$\boldsymbol{R}_{11}$，$\boldsymbol{R}_{22}$，$\boldsymbol{R}_{33}$ の部分のみにあてはめられる。）

ここで扱われるモデルは，より一般的なモデルである Jöreskog（1973）の共分散構造分析（covariance structure analysis）の特殊な場合にすぎないとも言える。しかし，特殊な場合のみを目指した結果，アルゴリズム的には，固有値計算の反復に帰することができ，演算速度，記憶容量の面では大幅な節約が可能になっている。すなわち，項目数と因子数（全体としてパラメータの数）に関する制約は著しく緩和されている。一方，共分散構造分析にあるような，モデルのあてはまりに関する検定統計量はこの方法には備わっていない。したがって，この方法は最初に述べたような研究の最も初期の段階，項目分析と尺度構成の段階において有効に利用される可能性のある探索的方法であるといえよう。

* Horst（1965）は，正準相関分析の一般化を Multiple set factor analysis と呼んでおり，このモデルには重集合因子分析という訳語があてられている。

2. モデル

各集合 k（$=1, \ldots, p$）は，それぞれ n_k 個の項目を含むものとする。各集合は少くとも3つの項目を含む，すなわち $n_k \geqq 3$ とする。これは，モデルの識別性（identifiability）を考慮してのことである（Anderson & Rubin, 1956）。各項目は，標準得点化して考えるものとする。したがって，個体の数を N とするとき $N \times n_k$ のデータ行列 $\boldsymbol{Z}_k$ による

$$\boldsymbol{R}_{kk'} = \frac{1}{N}\boldsymbol{Z}_k^{\mathrm{T}}\boldsymbol{Z}_{k'} \qquad k, k' = 1, \ldots, p \tag{1}$$

は，2つの集合 k と k' の間の項目間相関行列である。（正方行列とは限らない。）$\boldsymbol{Z}_k^{\mathrm{T}}$ は $\boldsymbol{Z}_k$ の転置を示す（以下同じ）。

$$\boldsymbol{Z} = [\boldsymbol{Z}_1, \boldsymbol{Z}_2, \ldots, \boldsymbol{Z}_p] \tag{2}$$

は，データ全体をあらわすが，これを用いた，

$$\boldsymbol{R} = \frac{1}{N}\boldsymbol{Z}^{\mathrm{T}}\boldsymbol{Z} \qquad k = 1, \ldots, p \tag{3}$$

は，$\sum_{k=1}^{p} n_k$ 次の正方行列で，いわゆる partitioned correlation matrix となる（図2）。$N > \sum_{k=1}^{p} n_k$ で $\boldsymbol{R}$ は特異でないと仮定する。

各集合ごとに通常の因子モデル

$$\boldsymbol{Z}_k = \boldsymbol{F}_k\boldsymbol{A}_k^{\mathrm{T}} + \boldsymbol{E}_k \tag{4}$$

を考える。因子数は集合 k ごとに異なってもよいものとし，q_k とかく。$\boldsymbol{F}_k$ は $N \times q_k$ の因子得点行列で，

$$\frac{1}{N}\boldsymbol{F}_k^{\mathrm{T}}\boldsymbol{F}_k = \boldsymbol{I}_{q_k} \qquad k = 1, \ldots, p \tag{5}$$

を満たすものとする。ここで $\boldsymbol{I}_{q_k}$ は，q_k 次の単位行列である。すなわち，各因子得点は分散1で，各集合内で直交するものとする。$\boldsymbol{A}_k$ は $n_k \times q_k$ の因子負荷行列，$\boldsymbol{E}_k$ は独自成分と誤差の和を要素とする $N \times n_k$ 行列で，

$$\frac{1}{N}\boldsymbol{F}_k^{\mathrm{T}}\boldsymbol{E}_{k'} = \boldsymbol{O} \qquad k, k' = 1, \ldots, p \tag{6}$$

$$\frac{1}{N}\boldsymbol{E}_k^{\mathrm{T}}\boldsymbol{E}_{k'} = \boldsymbol{D}_{kk'} \qquad k, k' = 1, \ldots, p \tag{7}$$

を満足するものとする。ここで $\boldsymbol{O}$ はゼロ行列，$\boldsymbol{D}_{kk'}$ は $k = k'$ のとき対角行列，$k \neq k'$ ならゼロ行列となるような行列である。(6)，(7)は通常の因子モデルにおける仮定の自然な拡張と言ってよいであろう。すなわち，(6)は因子得点と独自性が無相関であるという仮定が，集合間でも成立すると考えるものであり，(7)は異なる項目の独自性が相互に無相関であることを述べているにすぎない。

さて，(4)を(1)に代入し，(5)〜(7)を用いて整理すると，

$$\begin{bmatrix} R_{11} & R_{12} & R_{13} \\ R_{21} & R_{22} & R_{23} \\ R_{31} & R_{32} & R_{33} \end{bmatrix} = \begin{bmatrix} A_1 & 0 & 0 \\ 0 & A_2 & 0 \\ 0 & 0 & A_3 \end{bmatrix} \begin{bmatrix} I & S_{12} & S_{13} \\ S_{21} & I & S_{23} \\ S_{31} & S_{32} & I \end{bmatrix} \begin{bmatrix} A_1^T & 0 & 0 \\ 0 & A_1^T & 0 \\ 0 & 0 & A_3^T \end{bmatrix} + \begin{bmatrix} D_{11} & 0 & 0 \\ 0 & D_{22} & 0 \\ 0 & 0 & D_{33} \end{bmatrix}$$

図 4　多集合因子分析の基本方程式

$$R_{kk'} = A_k S_{kk'} A_{k'}^T + D_{kk'} \tag{8}$$

という基本方程式を得る。ここで

$$S_{kk'} = \frac{1}{N} F_k^T F_{k'} \tag{9}$$

であり，これは，因子得点間相関行列である。$k = k'$ のときは，(5)より，

$$S_{kk} = I_{q_k} \tag{10}$$

となる。

基本方程式を図示すると，図4のようになる。これを図3と比較されたい。因子負荷行列の多くの部分は強制的に0とされ，かわりに因子得点間相関行列が付け加えられる。この結果，多くの場合パラメータの数は減少することが見てとれるであろう。また，因子負荷は，その集合内のみに関わることから解釈は容易になり，因子得点間相関によって集合間の関係も見やすくなる。

次節で述べるアルゴリズムは，この基本方程式を最小2乗法で解くものである。これによって，$R_{kk'}(k \neq k')$ の部分（異なる集合に属する項目間相関）を，因子負荷 A_k，$A_{k'}$ に反映させることができる。もう1つのメリットとして，この方法によれば推定された因子得点によることなく，因子得点間の相関を求めることができる。

3.　アルゴリズム

ここで述べるアルゴリズムは，(8)の右辺を最小2乗法の意味で左辺の相関行列に最も近づけるものであり，共通性反復推定を伴う主因子法と同じ考え方である。

モデル自体は，通常の共通性推定を伴う因子分析の個別的な適用とほとんど同一であるから，もし，データに標本誤差が含まれなければ，（これは実際上標本の大きさが無限大であれば，ということであるが，）個別の因子分析と全く同じ結果を与えるはずである。

最適解の算出　最小化すべき基準は，

$$S^2 = \mathrm{tr} \sum_{k=1}^{p} \sum_{k'=1}^{p} (R_{kk'} - D_{kk'} - A_k S_{kk'} A_{k'}^T)(R_{kk'} - D_{kk'} - A_k S_{kk'} A_{k'}^T)^T \tag{11}$$

である。この基準は，換言すれば仮定(7)を最大限満足させようとする基準，あるいは残差行列の要素の2乗和を最小にする基準（MINRES基準，Harman & Jones, 1966）と言いかえてもよい。これを(10)の制約条件を満たしながら解くことは容易ではない。しかし，U_k を任意の q_k 次の非特異行列とするとき，

$$A_k^* = A_k U_k^{-1} \tag{12}$$

$$S_{kk'}^* = U_k S_{kk'} U_{k'} \tag{13}$$

のように変換したとき，

$$A_k^* S_{kk'}^* A_{k'}^{*T} = A_k S_{kk'} A_{k'}^T \tag{14}$$

だから，A_k，$S_{kk'}$ は，非特異変換に関して不定性をもつことがわかる。したがって，制約条件を(10)のかわりに

$$A_k^{*T} A_k^* = I_{q_k} \tag{15}$$

に置きかえて，A_k^*，S_{kk}^* を求めても同じ S^2 の最小値に到達するはずである。そうすると，最小化基準は，

$$S^2 = \mathrm{tr} \sum_{k=1}^{p} \sum_{k'=1}^{p} (R_{kk'} - D_{kk'} - A_k^* S_{kk'}^* A_{k'}^{*T})^T (R_{kk'} - D_{kk'} - A_k^* S_{kk'}^* A_{k'}^*) \tag{16}$$

となる。これを，A_k^*, $S_{kk'}^*$, D_{kk} によって偏微分し，そのすべての項を0とおいて整理することによって，

$$\{\sum_{k'=1}^{p} (R_{kk'} - D_{kk'}) A_{k'}^* A_{k'}^{*T} (R_{k'k} - D_{k'k})\} A_k^* = A_k^* (\sum_{k'=1}^{p} S_{kk'}^* S_{k'k}^*) \tag{17}$$

$$S_{kk'}^{*} = A_k^{*T}(R_{kk'} - D_{kk'})A_{k'}^{*} \tag{18}$$

$$D_{kk} = \mathrm{diag}(R_{kk} - A_k^{*}S_{kk}^{*}A_k^{*T}) \tag{19}$$

を得る。ここで，(17)の右辺の $\sum_{k'=1}^{p} S_{kk'}^{*} S_{k'k}^{*}$ は，正定符号行列だから，正数を対角要素としてもつ対角行列 L_k と，直交行列 T_k^{*} により，

$$\sum_{k'=1}^{p} S_{kk'}^{*} S_{k'k}^{*} = T_k^{*}L_kT_k^{*T} \tag{20}$$

のように対角化できる。したがって，(17)は，

$$\{\sum_{k'=1}^{p}(R_{kk'} - D_{kk'})A_{k'}^{*}A_{k'}^{*T}(R_{k'k} - D_{k'k})\}A_k^{*}T_k^{*}$$
$$= A_k^{*}T_k^{*}L_k \tag{21}$$

とかきかえることができる。すなわち，A_k^{*} は，(21)の左辺の｛　｝内の行列の固有ベクトル，またはその直交回転である。$A_k^{*}T_k^{*}T_k^{*T}A_k^{*T} = A_k^{*}A_k^{*T}$ であることも考えて，$A_k^{*}T_k$ を改めて A_k^{*} と呼ぶことにすれば，解は固有ベクトルとして一意的となり，

$$\{\sum_{k k'=1}^{p}(R_{kk'} - D_{kk'})A_{k'}^{*}A_{k'}^{*T}(R_{k'k} - D_{k'k})\}A_k^{*} = A_k^{*}L_k \tag{22}$$

に帰着する。なお，

$$S^2 = \mathrm{tr}\sum_{k=1}^{p}\sum_{k'=1}^{p}(R_{kk'} - D_{kk'})^{T}(R_{kk'} - D_{kk'}) - \mathrm{tr}\sum_{k=1}^{p}L_k \tag{23}$$

であることが(17)，(18)，(20)を用いて示せるから，S^2 を最小化するためには，固有値は常に代数的な大小順で q_k 番目までのものをとることになる。

そこで，アルゴリズムとしては，D_{kk} の，（実際には，

$$H_{kk} = I - D_{kk} \tag{24}$$

によって定義される共通性の）適当な推定値から出発して，（例えば R^{-1} による Guttman の SMC など）（$R_{kk} - D_{kk}$）の固有ベクトルによって A_k^{*} の初期値とし，これを(22)の｛　｝内の行列のところに代入して，第1次近似の A_k^{*} を求めたあと，以下(18)，(19)，(22)をこの順に収束するまで反復すればよい（図5参照）。

ここで $p = 1$ の場合について考えてみると，(17)は，

$$(R_{11} - D_1)A_1^{*}S_{11}^{*} = A_1^{*}S_{11}^{*2}$$

という形となる。この両辺に右から S_{11}^{*-1} をかけ，更に，(20)が，

$$S_{11}^{*2} = T_1^{*}L_1T_1^{*T}$$

すなわち，

$$S_{11}^{*} = T_1^{*}L_1^{\frac{1}{2}}T_1^{*T} \tag{25}$$

となることから，結局

$$(R_{11} - D_{11})A_1^{*}T_1^{*} = A_1^{*}T_1^{*}L_1^{\frac{1}{2}} \tag{26}$$

に帰着する。$A_1^{*}T_1^{*}$ は $R_{11} - D_{11}$ の固有ベクトル，$L_1^{\frac{1}{2}}$ は固有値を対角要素とする対角行列であり，これから，(19)は，

$$D_{11} = \mathrm{diag}\,\{R_{11} - A_1^{*}T_1^{*}L_1^{\frac{1}{2}}(A_1^{*}T_1^{*})^{T}\} \tag{27}$$

となる。(25)と(26)を交互に反復するのが，共通性の反復推定をともなう主因子法であることは容易に見てとれるであろう。この意味で，先に述べたアルゴリズムは，主因子法の自然な拡張であると考えることができる。

解の変換　さて，A_k^{*} と $S_{kk'}^{*}$ を，本来の A_k，$S_{kk'}$ に変換することを考えよう。$k = k'$ の場合に対応する S_{kk}^{*} は，通常は正定符号行列となるから，正数を対角要素としてもつ対角行列 Λ_k と，直交行列 X_k が存在して，

$$S_{kk}^{*} = X_k\Lambda_kX_k^{T} \tag{28}$$

と分解することができる。したがって，

$$S_{kk}^{*} = S_{kk}^{*\frac{1}{2}}S_{kk}^{*\frac{1}{2}}$$

とすることのできる対称行列 $S_k^{*\frac{1}{2}}$ を，

$$S_{kk}^{*\frac{1}{2}} = X_k\Lambda_k^{\frac{1}{2}}X_k^{T} \tag{29}$$

によって定義することができる。そこで，(12)，(13)の U_k としては，この $S_k^{*\frac{1}{2}}$ をとればよい。すなわち，

$$A_k = A_k^{*}S_{kk}^{*\frac{1}{2}} \tag{30}$$

$$S_{kk'} = S_{kk}^{*-\frac{1}{2}}S_{kk'}^{*}S_{k'k'}^{*-\frac{1}{2}} \tag{31}$$

なる変換によって，制約条件(10)は，

$$S_{kk} = S_{kk}^{*-\frac{1}{2}}S_{kk}^{*}S_{kk}^{*-\frac{1}{2}} = I_{q_k} \tag{32}$$

のようにして満足されることになる。

ここで再度 $p = 1$ の場合について考えてみるならば，(25)からわかるように，(28)において $\Lambda_1 = L_1^{\frac{1}{2}}$ でありかつ $X_1 = T_1^{*}$ だから，(29)は，

$$S_{11}^{*\frac{1}{2}} = T_1^{*}L_1^{\frac{1}{4}}T_1^{*T}$$

となる。したがって，(30)は，

$$A_1 = A_1^{*}T_1^{*}\Lambda^{*\frac{1}{4}}T_1^{*T}$$

すなわち，

$$\boldsymbol{A}_1\boldsymbol{T}_1^* = (\boldsymbol{A}_1^*\boldsymbol{T}_1^*)\Lambda^{*\frac{1}{4}} \tag{33}$$

となる。(26)により，$\boldsymbol{A}_1^*\boldsymbol{T}_1^*$ は $\boldsymbol{R}_{11}-\boldsymbol{D}_1$ の固有ベクトル $\Lambda^{*\frac{1}{4}}$ は固有値の平方根であるから，（直交回転以前の）因子負荷は，固有ベクトルに固有値の平方根を掛けたものである，という主因子法のアルゴリズムが再びあらわれた。(33)の左辺に直交回転 $\boldsymbol{T}_1^{*T}$ をほどこすことによって，最終的な因子負荷 $\boldsymbol{A}_1$ が得られる。）

なお，$p \neq 1$ の場合も含めて，$\boldsymbol{D}_{kk}=\boldsymbol{O}$ として，得られた固有ベクトル $\boldsymbol{A}_1^*$ をただちに(18)に適用して $\boldsymbol{S}_{kk'}^*$ を求め（ここでももちろん $\boldsymbol{D}_{kk}=\boldsymbol{O}$ として），(28)～(31)の適用へと進めば，これは集合ごとに個別に主成分分析（共通性をすべて1と仮定した主因子法）による構造ベクトル $\boldsymbol{A}_k$ と，主成分得点間相関行列 $\boldsymbol{S}_{kk'}$ を求めたことになる。

解の回転　$\boldsymbol{A}_k$，$\boldsymbol{S}_{kk'}$ にはなお，直交回転に関する自由度が残されている。（$p=1$ の場合である(33)には，これが explicit にあらわれている。$p \geqq 2$ でも，(21)でみられるように固有ベクトルを $\boldsymbol{A}_k^*\boldsymbol{T}_k^*$ としてそのまま扱っていけば同じ表現になる。）すなわち適当な直交行列 $\boldsymbol{T}_k$ により，

$$\boldsymbol{A}_k^{\dagger} = \boldsymbol{A}_k\boldsymbol{T}_k \tag{34}$$

$$\boldsymbol{S}_{kk'}^{\dagger} = \boldsymbol{T}_k^{\mathrm{T}}\boldsymbol{S}_{kk'}\boldsymbol{T}_{k'} \tag{35}$$

とすることができる。具体的には，各 $\boldsymbol{A}_k$ を個別に（例えば）バリマックス回転し，$\boldsymbol{S}_{kk}$ にも対応した回転を行なうことになる。

また，制約条件(10)を外すことを考えれば斜交回転の適用も当然考えられる。

更に，$p=2$ かつ $q_1=q_2$ の場合に限って言えば，$\boldsymbol{S}_{12}(=\boldsymbol{S}_{21}^{\mathrm{T}})$ を対角化するような回転も考えられる。これは，因子得点が異なる集合間でも直交するという制約を課したことに等しく，Corballis & Traub (1970) の縦断的因子分析（longitudinal factor analysis）と同等である。以上のアルゴリズムは，村上（1983）における“一般的モデル”（pp. 156～158）のそれを，集合ごとに変数が異なり，更に共通性推定を含む形に一般化したものである。

計算の流れ図を図5に示す。

1項目だけの集合　ここまで，各集合は最低3つの項目を含むものとした。項目数が2以下では解が一意に定まらないからである。しかし，対象者の“年齢”のように単独で意味をもつような項目も質問紙にはしばしば含まれるから，これと他の集合の因子得点との相関を求めることは，当然興味の対象となる。

この場合，$\boldsymbol{R}_{kk}$，$\boldsymbol{A}_k^*$，$\boldsymbol{S}_{kk}^*$，$\boldsymbol{D}_{kk}$ はすべて 1×1 の行列であり，

$$\boldsymbol{R}_{kk} = [1.0]$$

である。解は一意的には決まらないので，

$$\boldsymbol{A}_k^* = [1.0],\ \boldsymbol{S}_{kk}^* = [1.0],\ \boldsymbol{D}_{kk} = [0.0]$$

と定義する。すなわち単独の変数については，それ自身を因子得点とみなすことにする。（$\boldsymbol{Z}_k=\boldsymbol{F}_k$，ただし $\boldsymbol{Z}_k$ も $\boldsymbol{F}_k$ も N 次元ベクトルである。）

こうして，1項目だけの集合 k については，(22)，(18)は常に bypass されるが，(22)による他の集合の固有ベクトルの算出に定数 $\boldsymbol{A}_k^*$ は含められ，その $\boldsymbol{A}_k^*$ を用いて，(19)の計算が行なわれる。解の変換に関しては，

$$\boldsymbol{A}_k = \boldsymbol{A}_k^* = [1.0],\ \boldsymbol{S}_{kk} = \boldsymbol{S}_{kk}^* = \boldsymbol{S}_{kk}^{*\frac{1}{2}} = [1.0]$$

で，何もしなくてよい。回転は当然この集合については行わない（行えない。もっともこの点は因子数1の集合については常に同様である。）形式的には，

$$\boldsymbol{T}_k = [1.0]$$

とすればよい。このように，若干のアルゴリズムの改訂は必要となるものの，1項目からなる集合も，一応問題なく分析に含めることができる。

なお，その1項目は，“性別”，“有職か否か”といった2値変数であってもよい。その場合の因子得点間相関は群差を反映する点双列相関（point biserial correlation）とみなされる。

因子得点の推定　最後に，因子得点の推定法について触れておこう。このモデルは，再三繰り返すように研究の初期の段階における探索的な利用を主として考えており，また分析の対象となる集合間では因子得点間相関の推定値も得られるので，因子得点を推定する必要性はさほど感じられないであろう。しかしながら，デモグラフィックな変数との関係を検討する場合等に，因子得点の推定値を用いることはありうる。

ここでは，しばしば用いられる真の因子得点 $\boldsymbol{F}$ に対する最小2乗基準

$$S_k^2 = \mathrm{tr}\,\frac{1}{N}(\boldsymbol{F}_k-\hat{\boldsymbol{F}}_k)^{\mathrm{T}}(\boldsymbol{F}_k-\hat{\boldsymbol{F}}_k) \tag{36}$$

を用いることにしよう。ここで $\hat{\boldsymbol{F}}_k$ が因子得点の推定値であるが，この定義には少くとも2通りありうる。

まず第1は，因子得点をその集合のみのデータから推定するもので，$n_k \times q_k$ の重み行列を $\boldsymbol{W}_k$ とするとき，

$$\hat{\boldsymbol{F}}_k = \boldsymbol{Z}_k\boldsymbol{W}_k \tag{37}$$

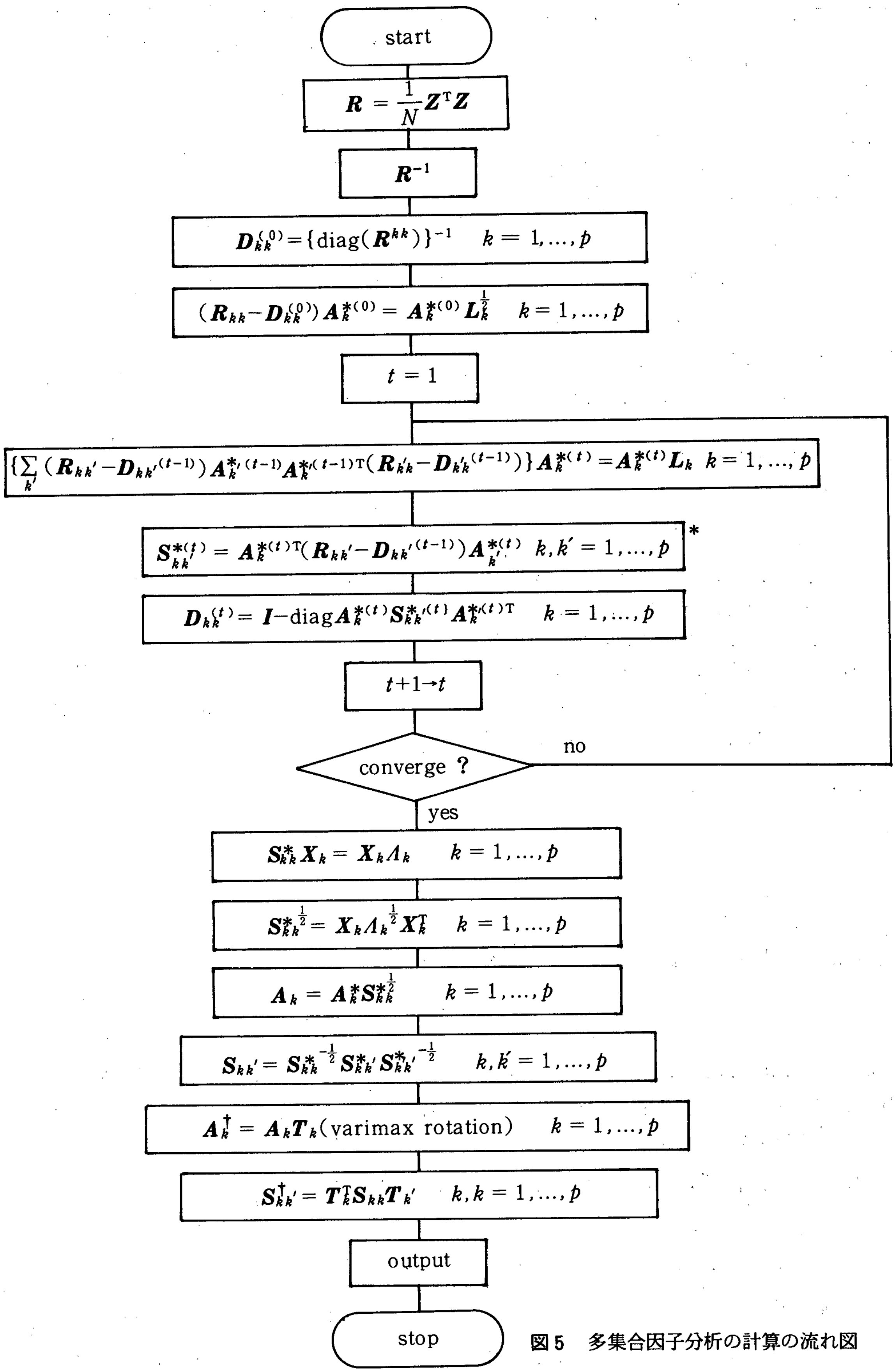

図 5　多集合因子分析の計算の流れ図

* 収束以前の段階では $k' = k$ の場合だけ計算し，収束後に $k' \neq k$ の場合の計算を行うようにすれば，時間は短縮される。

によって推定する。この場合，通常の推定式と全く同じ

$$\boldsymbol{W}_k = \boldsymbol{R}_{kk}^{-1}\boldsymbol{A}_k \tag{38}$$

が得られる。

第2に，因子得点を他の集合を含む全データから推定することが考えられる。k 番目の集合の因子得点を求めるための k' 番目の集合に対する重み行列を $\boldsymbol{W}_k^{(k')}$ とするとき，推定値を

$$\hat{\boldsymbol{F}}_k = \sum_{k'=1}^{p} \boldsymbol{Z}_{k'}\boldsymbol{W}_k^{(k')} \tag{39}$$

によって得るとする。(39)を(36)に代入し，$\boldsymbol{W}_k^{(k')}$ で偏微分して $\boldsymbol{O}$ とおくことにより，

$$\frac{1}{N}\boldsymbol{Z}_{k'}^{T}\boldsymbol{F}_k = \sum_{k''=1}^{p} \boldsymbol{R}_{k'k''}\boldsymbol{W}_k^{(k'')} \tag{40}$$

を得る。この左辺は，基本モデル(4)に，$\boldsymbol{F}_{k'}^{T}$ を左からかけることによって得られる，

$$\boldsymbol{A}_k\boldsymbol{S}_{kk'} = \frac{1}{N}\boldsymbol{Z}_k^{T}\boldsymbol{F}_{k'} \tag{41}$$

により，$\boldsymbol{A}_{k'}\boldsymbol{S}_{k'k}$ に等しい。((41)は $k'=k$ の場合を考えれば，“因子が直交であるならば，因子負荷は因子構造，すなわち粗点と因子得点との相関係数に等しい。” という個別因子分析の命題の一般化になっている。)

そこで(41)は，

$$\begin{bmatrix} \boldsymbol{A}_k\boldsymbol{S}_{k1} \\ \boldsymbol{A}_k\boldsymbol{S}_{k2} \\ \cdot \\ \cdot \\ \cdot \\ \boldsymbol{A}_k\boldsymbol{S}_{kp} \end{bmatrix} = \begin{bmatrix} \boldsymbol{R}_{11} & \boldsymbol{R}_{12} & \cdot\ \cdot\ \cdot & \boldsymbol{R}_{1p} \\ \boldsymbol{R}_{21} & \boldsymbol{R}_{22} & \cdot\ \cdot\ \cdot & \boldsymbol{R}_{2p} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \boldsymbol{R}_{p1} & \boldsymbol{R}_{p2} & \cdot\ \cdot\ \cdot & \boldsymbol{R}_{pp} \end{bmatrix} \begin{bmatrix} \boldsymbol{W}_k^{(1)} \\ \boldsymbol{W}_k^{(2)} \\ \cdot \\ \cdot \\ \cdot \\ \boldsymbol{W}_k^{(p)} \end{bmatrix} \tag{42}$$

とかける。右辺の左側の行列は，(3)で定義された $\boldsymbol{R}$ であり，この逆行列 $\boldsymbol{R}^{-1}$ を再び集合ごとに分割した部分行列を $\boldsymbol{R}^{kk'}$ とかくと，重み行列は，

$$\boldsymbol{W}_k^{(k')} = \sum_{k''=1}^{p} \boldsymbol{R}^{k'k''}\boldsymbol{A}_{k''}\boldsymbol{S}_{k''k} \qquad k' = 1, \ldots, p \tag{43}$$

によって求められることになる。

2種の推定方法の比較は，未だなされていない。

4. 適用例

ここでは図1に示した質問紙による調査データを分析してみる。

対象者は，愛知県内の消費者モニター（主婦）500名（悉皆）と，各モニターが2名づつ選んだ一般主婦1,000名であった。調査方法は，モニターについては郵送法，

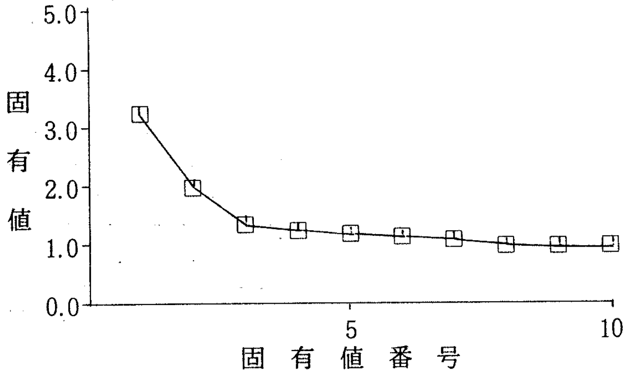

図6 問1の固有値の推移

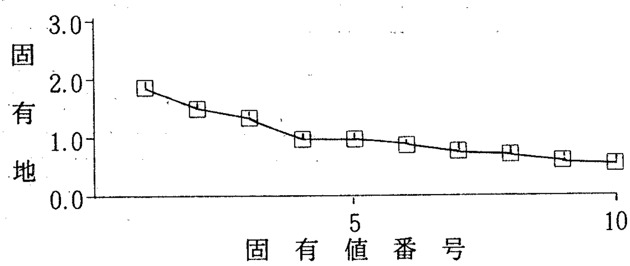

図7 問2の固有値の推移

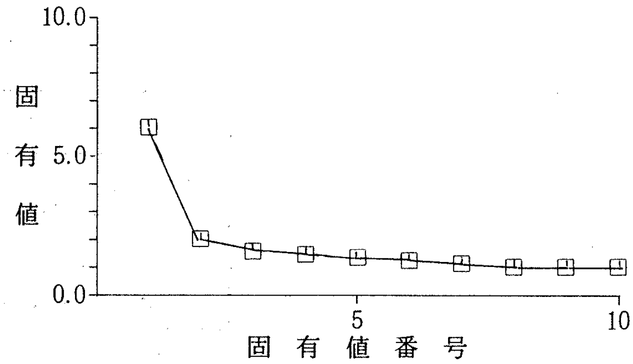

図8 問3の固有値の推移

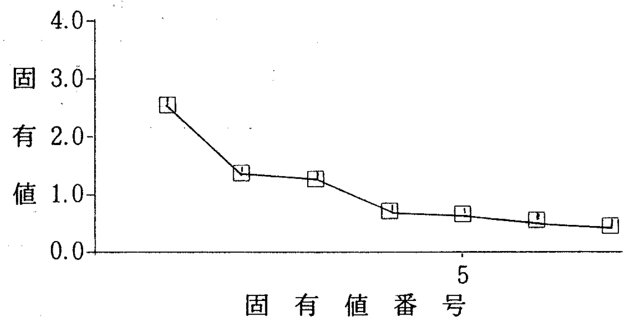

図9 問4の固有値の推移

一般主婦については，モニターによる留めおき法である。調査時期は1982年10月下旬から11月上旬であった。有効回収数は1342であったが，ここでは図1にあげた64項目のすべてに回答した1153名分のデータを分析の対象とする。図の問1～問4は，実際の調査では問6～問9にあたっており，分量的には調査全体の約2/5にあたる。

問1の20項目は，主婦の購買行動をとらえるための項目であり，5段階評定であるが，コーディングは図に示されたのとは逆に，値が大きい程似ている方向に逆転された。

問2の10項目は，購買態度，または意識をたずねる項目であり，強制選択の形をとっているが，これもコーディングの方向は逆転した。

問3は，情報に対するニーズと関心度を問う27項目からなる。この分析では，内容をニーズに限定するため，“関心はない”という選択肢は，“関心があり十分である”と同じ1とコーディングされている。結果的にこの問は1か2の2件法と同じことになっている。そして，ニーズが高いほど得点が高くなるものとみなされる。*

問4は，情報媒体に対する信頼度を調べるもので7項目からなる。これも，信頼度が高いほど値が高くなる方向にコーディングされる。

これら4つの問を，それぞれ項目の集合とみなして，前節のアルゴリズムを適用する。プログラムはFORTRAN77で書かれ，実数はすべて倍精度とした。

まず，個別の主成分分析によって得られた固有値を集合別に示したのが図6～9である。これらの図から，問1～問3の因子数（q_1，q_2，q_3）を2，問4の因子数（q_4）を3とした多集合因子分析によって求められた因子負荷行列を表2～5に，因子得点間相関行列を，表6に示した。すべての固有ベクトルの要素の変化が0.00001以下になるという停止基準により，収束までに45回の反復を要した。各表には比較のために，同じ因子数による主成分分析と，主成分得点間相関の値も示した。それぞれの因子負荷はバリマックス回転され，因子得点，主成分得点にも同じ回転が施されている。

まず，2つの方法による因子負荷は極めてよく似ていると言ってよいであろう。それぞれの問（集合）ごとに解釈してみる。

問1の第1因子は，“日用品や食料品などは……もっとも経済的なものを買う”（項目5），“安くて良い商品を売っている店をさがす”（項目9），“新聞広告や折込み・チラシ広告を見て買物をする”（項目11），“生鮮食料品は出盛りで，一番安い旬になってから買う”（項目8），“日用品や食料品などは，特売や安い時にまとめ買いをする”（項目3）等，できるだけ安くて良い商品を購入しようとする行動をあらわす項目の負荷が高く，“購買行動の経済性”と呼ぶことができよう。第2因子は，“自分にとって本当に必要なものを必要なときに必要なだけ買う”（項目12），“スーパーの店頭で予定になかったものを買ってしまう（否定）”（項目7），“…… 値段が高くても品質のよいものを買う”（項目20），といった項目が高く負荷し，更に，第1因子に正に負荷していた項目3，11が負の負荷をもち，“購買行動の計画性”と呼んでよいと思われる。

問2については，第1因子には，“定期的に買うものの値段が変わるとすぐ気がつく”（項目10），“将来のため貯蓄や保険にお金をまわす”（項目1），“欲しいものは貯蓄してから買う”（項目2）等，“経済的意識”とでも呼べる項目が高く負荷し，第2因子には，“いつも同じ銘柄のものを買う……”（項目7）と“買い物をするときは，いつも決まった店でする”（項目9）の2項目の負荷が高く，“購買行動の固定性”と解される。

問3は，情報へのニーズを問うものであったが，第1因子には，“生鮮食料品の値動きの向う1ケ月程度の見通し”（項目14），“一般消費者の方々の物価，消費生活に対する意見，お考えの紹介”（項目19），“サービス料金の動向の解説”（項目18），“国や県・市などで実施する各種の調査の結果”（項目6），“国の経済政策や国際経済の情勢についてのやさしい解説”（項目21）等，やや高度で一般的な，必ずしも即時的に役に立つわけではないような情報に関する項目の負荷が高い。この因子は“一般情報に対するニーズ”と呼んでおこう。第2因子は，“旬のものを使った料理法”（項目22），“よりよい暮しをするために役立つアイディア，ヒントや生活の知恵”（項目13），“食品を長持ちさせたり保存したりする知識の提供”（項目2），“省エネルギーのための知識”（項目9）といった，ただちに役立つような，やや特殊な情報に関する項目が高く負荷しており，“特殊情報に対するニーズ”と呼んでおく。

問4では，さまざまな情報媒体に対する信頼度が問われている。この部分の因子の解釈は極めて明解であり，第1因子（項目1～3の負荷が高い）は，“マスコミ情報への信頼度”，第2因子（項目4，7が高く負荷）は，“消費者雑誌・広報への信頼度”，第3因子（項目5，6

---

* このような2値データを因子分析することは，長らく問題視されてきたが，実際にはしばしば行なわれてきたし，近年では（少くとも探索的な方法としては）正当な方法と考えられるに至っている。この問題についてやや理論的な考察を加えたものに McDonald（1985，pp.198～202）がある。

表2　問1の因子負荷

	多集合因子分析			主成分分析		
	1	2	h^2	1	2	h^2
1　予算の腹づもり	**35**	20	16	34	**43**	30
2　近所の人の話	29	−17	11	**40**	−05	16
3　特売でまとめ買い	**44**	**−43**	38	**59**	**−41**	52
4　内容表示の確認	33	05	11	37	18	17
5　もっとも経済的なもの	**56**	−06	32	**62**	03	39
6　好みを重視	−05	14	02	−07	15	03
7　予定外のものを買ってしまう	−17	**−43**	21	−08	**−57**	33
8　生鮮食料品は旬に	**45**	−07	21	**53**	−03	28
9　安くて良い品の店	**51**	−19	30	**63**	−12	40
10　共同で買い物	22	−10	06	31	−03	10
11　新聞広告，折り込み広告	**48**	**−43**	42	**62**	−35	50
12　必要な物を必要なだけ	29	**44**	27	21	**62**	43
13　買物リストをつくる	**41**	22	21	**40**	38	31
14　行列ができるような売場	25	**−40**	22	39	**−40**	31
15　「はかり売り」より「パック売り」	−06	−24	06	00	−33	11
16　レシートはすぐ捨てる	−33	−06	11	−35	−17	15
17　新聞記事や「暮しの手帖」	**38**	−04	15	**40**	06	16
18　通信販売は利用しない	22	11	06	24	17	09
19　大きな買物は家族と相談	30	06	09	34	22	17
20　値段より品質	−01	**41**	17	−11	**51**	27
2　乗　和	231	135	365	314	205	519

（小数点省略，以下同じ）

表3　問2の因子負荷

	多集合因子分析			主成分分析		
	1	2	h^2	1	2	h^2
1　将来のため貯蓄	**40**	08	17	**62**	05	38
2　欲しいものは借金してでも	**−38**	−05	14	**−68**	24	52
3　人と同じものを買う	−01	−10	01	−05	−11	02
4　現金買いが多い	31	08	10	**62**	−11	39
5　値切ることが多い	**36**	−05	13	**40**	16	18
6　耐久消費財を重視	29	07	09	**44**	26	26
7　いつも同じ銘柄	07	**48**	24	04	**79**	63
8　衣服はコーディネイト	13	−03	02	17	08	03
9　いつもきまった店	−04	**62**	39	−06	**74**	56
10　値段がかわると気がつく	**59**	−03	35	**49**	13	26
2　乗　和	99	66	164	185	139	323

表4　問3の因子負荷

	多集合因子分析			主成分分析		
	1	2	h^2	1	2	h^2
1　食品の安全性	09	**43**	19	−05	**62**	38
2　食品の保存法	08	**52**	27	−03	**66**	43
3　今日，明日の市場価格	**45**	11	22	**52**	10	28
4　教育費の負担状況	22	25	11	15	38	16
5　石油製品の価格	19	**39**	19	16	**48**	25
6　国・県・市の調査結果	**51**	17	29	**55**	17	33
7　商品知識，流通のしくみ	**44**	20	24	**48**	21	28
8　公共料金の改定	16	**38**	17	17	**41**	19
9　省エネの知識	19	**52**	30	17	**57**	36
10　代替品目の知識	**39**	25	22	33	37	25
11　安売りデー・青空市場	28	21	13	29	25	15
12　賃金・ボーナスの見通し	31	26	17	31	33	20
13　暮しのアイデア・ヒント	18	**53**	31	12	**63**	41
14　向う1ケ月の値動き	**58**	06	34	**68**	01	46
15　年金制度の解説	31	33	20	32	37	24
16　商品テストの結果	**45**	32	30	**45**	36	34
17　安い商店街・市場	**41**	22	21	38	31	24
18　サービス料金の動向	**55**	02	31	**61**	01	38
19　一般消費者の意見	**56**	15	33	**63**	12	41
20　冠婚葬祭のとおり相場	29	29	17	31	32	20
21　経済政策や国際経済	**48**	12	24	**55**	09	31
22　旬のものの料理法	06	**53**	28	09	**51**	27
23　景気や消費生活の見通し	**47**	25	29	**54**	23	34
24　バーゲンや特売	17	**39**	18	22	37	19
25　老後の生活設計	**35**	34	23	36	37	27
26　物価教室・消費生活講座	**38**	22	19	**52**	13	28
27　新製品の紹介	33	28	19	39	26	22
2　乗　和	354	273	628	420	361	781

表5　問4の因子負荷

	多集合因子分析				主成分分析			
	1	2	3	h^2	1	2	3	h^2
1　マスコミ広告	**81**	12	06	68	**89**	09	04	81
2　広告以外の記事番組	**59**	**37**	00	49	**70**	**43**	−10	69
3　折込広告，チラシ	**55**	−02	**38**	44	**68**	−11	**42**	65
4　消費者雑誌	11	**68**	12	48	09	**84**	10	72
5　近所の人，知人の話	05	10	**72**	54	02	10	**83**	70
6　店の人の話	11	22	**52**	33	12	18	**78**	66
7　広報紙の記事	12	**66**	19	49	09	**81**	20	70
2　乗　和	135	110	99	345	179	161	153	494

表6　因子得点間相関行列

		多集合因子分析									主成分分析								
		1		2		3		4			1		2		3		4		
		1	2	1	2	1	2	1	2	3	1	2	1	2	1	2	1	2	3
1	1	100									100								
	2	00	100								00	100							
2	1	**82**	03	100							**46**	12	100						
	2	−09	**49**	00	100						−03	**21**	00	100					
3	1	**33**	−02	**27**	−09	100					**21**	02	12	00	100				
	2	**−30**	−12	**−25**	01	00	100				−14	−10	−11	−01	00	100			
4	1	05	−19	04	−06	−06	−10	100			08	−12	03	−02	−04	−09	100		
	2	**25**	−04	23	05	15	−09	00	100		**18**	00	13	02	07	−03	00	100	
	3	06	−13	04	−08	−03	−06	00	00	100	08	−05	03	−04	00	−05	00	00	100

が高く負荷）は，“ミニコミ・クチコミへの信頼度”と呼ぶことができよう。

さて，以上の各因子間の相関を表6でみてみよう。多集合因子分析と，個別の主成分分析ではほぼ同一の傾向が示されているものの，多集合因子分析では高い相関はより高くなり，いわばメリハリの効いた結果が得られている。これは，多集合因子分析では共通性の推定が行なわれる結果，因子得点からは誤差が取り除かれ，ちょうど稀薄化の修正を行った相関係数に相当するものが算出されているからである。（稀薄化の修正の場合と同様，稀には1を越える相関があらわれることがある。このような場合，集合間での項目内容の重なりが大きすぎることが多いようである。）

さて，多集合因子分析の方で数値を具体的に見てみると，まず問1と問2の各第1因子間の相関0.82が目立つ。これらは，ともに“経済性”にかかわる因子であり，ある程度当然の結果と言えよう。第2因子どうしも0.49の相関がある。計画的な購買行動は，どうしても固定的な行動につながらざるをえないのであろう。

注目すべきは，問1，問2の各々の第1因子と，問3の2つの因子との関係である。問1の第1因子と問3の第1因子とは0.33の正の相関，問3の第2因子とは−0.30という負の相関があらわれており，この間の関係は全く逆転している。問2の第1因子と問3の2つの因子の間でも同じ関係となっている。これは，経済的購買活動を行っている主婦は，平素の情報探索活動の結果，特殊情報についてはもはや“十分である”と感じており，彼女等の関心はむしろより高度な“一般的情報”に向けられている，と解されるかもしれない。なお，購買行動が，“経済的”であると，“消費者雑誌・広報”への信頼度が高い傾向にあること（問1の第1因子と問4の第2因子の相関0.25，問2の第1因子でも同様）も示されている。

ところで，以上のような傾向は，多少の稀薄化は伴うもののすべて個別の主成分分析からも読みとることのできるものであり，特に多集合因子分析でなければ得られないというものでもない。

次に，問1の因子数（q_1）を3に変えた場合の結果を示す。こちらは，収束までに72回の反復を要した。

表7は，問1の回転後の因子負荷を，やはり（回転された）主成分分析の結果とともに示したものである。問2と問3については，結果は表3，4とほとんど同一である（一部の負荷が0.01程度変化しているにすぎない。）ので，省略し，表8に問4の多集合因子分析の結果を示す。やや値の動きがあるが，先の解釈をくつがえすほどのものではない。表9は，因子得点間相関であり，再び主成分得点間相関とともに示した。

表7の2種の因子負荷は今度はかなり異なるものとなっている。多集合因子分析の第1因子は，項目5，8の負荷が低下し，3，11の負荷が目立つようになり，新たに“たくさんの人たちが行列しているような店で買う”（項目14），“……値段が高くても品質のよいものを買う（否定）”（項目20）がやや高い負荷をもつようになった。したがって，むしろ“安値指向の購買行動”といった名称がふさわしいように思われる。第2因子では，項目20の負荷が低下，項目12の負荷が上昇し，新たに，“……買物計画リストを作り，なくなりかけたものを追加して買う”（項目13），“日常の買い物には予算の腹づもりをして出かける”（項目1），及び第1因子で負荷の低下した項目5が加わった結果，より明確に“計画性”を示す因子となった。第3因子で目立つのは，“新聞記事や「暮

しの手帖」などですすめている商品を選ぶ”（項目17)という1項目だけである。

表9から，問1と他の問との相関が，因子の内容の変化に伴って著しく変化したことが見てとれる。第1因子は，問2の第1因子となお正の相関をもつ（0.38）ものの，同一の内容と言えるほどではなく，むしろ第2因子とかなり高い逆相関（－0.51）をもつことになった。すなわち，この因子は（安いものを求める）“購買行動の可変性”という意味あいが強いようである。また問4の“マスコミ情報への信頼度”ともはっきり正の相関(0.26)を示すようにもなった。あえて言えば，この因子得点の高い人は，やや“情報におどらされる”側面があるようである。一方，問3の2つの因子に対する相反する関係はほとんど消滅してしまった。逆に，問2の第1因子と

表7　問1の因子負荷（3因子）

	多集合因子分析				主成分分析			
	1	2	3	h^2	1	2	3	h^2
1　予算の腹づもり	00	**37**	16	16	15	27	**46**	31
2　近所の人の話	24	07	23	12	**52**	−05	16	30
3　特売でまとめ買い	**59**	11	19	40	**60**	26	−30	52
4　内容表示の確認	05	19	29	12	09	**47**	09	24
5　もっとも経済的なもの	32	**38**	28	33	39	**52**	04	42
6　好みを重視	−16	02	01	03	18	**−41**	35	32
7　予定外のものを買ってしまう	24	**−42**	02	24	22	−31	**−46**	36
8　生鮮食料品は旬に	29	30	21	21	**40**	36	02	29
9　安くて良い品の店	**40**	26	28	30	**56**	31	00	41
10　共同で買い物	11	02	27	09	**44**	−08	16	23
11　新聞広告，折り込み広告	**62**	14	20	44	**57**	34	−27	51
12　必要な物を必要なだけ	−18	**52**	03	31	04	16	**66**	46
13　買物リストをつくる	05	**47**	12	24	28	20	**47**	34
14　行列ができるような売場	**44**	−04	15	22	**60**	−10	−18	41
15　「はかり売り」より「パック売り」	24	−11	−15	09	19	−20	−24	13
16　レシートはすぐ捨てる	−11	−28	−15	11	03	**−62**	00	39
17　新聞記事や「暮しの手帖」	04	04	**60**	36	**42**	07	21	23
18　通信販売は利用しない	−03	16	21	07	03	33	11	12
19　大きな買物は家族と相談	03	18	27	11	10	39	17	19
20　値段より品質	**38**	18	05	18	−08	−19	**59**	39
2　乗　和	168	137	107	412	256	206	194	656

表8　問4の因子負荷（多集合因子分析で問1を3因子にした場合）

	1	2	3	h^2
1　マスコミ広告	**80**	13	05	66
2　広告以外の記事番組	**58**	**39**	−01	50
3　折込広告，チラシ	**56**	−03	**38**	46
4　消費者雑誌	09	**73**	12	56
5　近所の人，知人の話	04	11	**71**	53
6　店の人の話	12	20	**53**	33
7　広報紙の記事	13	**60**	20	42
2　乗　和	134	112	100	345

表9　因子得点間相関行列（問1，3因子）

		多集合因子分析										主成分分析									
		1			2		3		4			1			2		3		4		
		1	2	3	1	2	1	2	1	2	3	1	2	3	1	2	1	2	1	2	3
1	1	100										100									
	2	00	100									00	100								
	3	00	00	100								00	00	100							
2	1	**38**	**65**	**35**	100							**26**	**40**	12	100						
	2	**−51**	11	11	00	100						−06	−02	**22**	00	100					
3	1	08	11	**37**	**27**	−08	100					**18**	10	07	12	00	100				
	2	−06	**−30**	−13	**−25**	01	00	100				−11	−06	−14	−11	−01	00	100			
4	1	**26**	04	−17	04	−06	−06	−10	100			14	−03	−07	03	−02	−04	−09	100		
	2	−05	−12	**62**	23	05	14	−09	00	100		14	11	02	13	02	07	−03	00	100	
	3	12	−04	07	04	−08	−02	−06	00	00	100	12	−03	00	03	−04	00	−05	00	00	100

表10　問1の因子負荷（主因子法）

		1	2	3	h^2
1	予算の腹づもり	17	18	**38**	20
2	近所の人の話	33	06	05	11
3	特売でまとめ買い	**62**	09	−22	44
4	内容表示の確認	11	**53**	−03	30
5	もっとも経済的なもの	**45**	34	07	32
6	好みを重視	−04	−07	12	02
7	予定外のものを買ってしまう	07	−14	**−42**	20
8	生鮮食料品は旬に	**43**	14	09	21
9	安くて良い品の店	**55**	16	04	32
10	共同で買い物	24	05	05	06
11	新聞広告，折り込み広告	**60**	17	−20	43
12	必要な物を必要なだけ	09	04	**69**	49
13	買物リストをつくる	28	11	**41**	25
14	行列ができるような売場	**49**	−19	−09	28
15	「はかり売り」より「パック売り」	13	−24	−11	08
16	レシートはすぐ捨てる	−16	−33	−06	14
17	新聞記事や「暮しの手帖」	28	15	09	11
18	通信販売は利用しない	08	25	06	07
19	大きな買物は家族と相談	12	**38**	07	16
20	値段より品質	−21	14	28	14
	2　乗　和	215	100	120	436

高い相関（0.65）をもつようになったのは第2因子である。問2の第2因子には，“慎重性”と解される面もあり，このことが“計画性”との高い相関をもたらしたのであろう。そしてこの因子が，問3の第2因子“特殊情報へのニーズ”と逆相関（−0.30）を示すようになった。計画的な消費者にとって，即時的に有効な情報は，さほど意味をもたないのであろうか。

問1の第3因子は，まさに多集合因子分析のアルゴリズムが成立させたものである。これは，問4の第2因子“消費者雑誌・広報への信頼度”と0.62という高い相関

をもち，また問３の第１因子とも0.37の相関をもつ。いわば，この因子は，第１因子よりも高い意味での“経済性”の追求を示す因子である。これは問２の第２因子との高い相関とも矛盾しない。そしてどうやら，この因子を出現させたのは，主として問１の項目17と，問４の項目４にともに含まれる「暮しの手帖」であるように思われる。ただし負荷の２乗和からもわかるように，それだけでは説明しつくせない。

このように見てくると，問１と他の問との関係は，因子数を３にした場合の方がより納得のいくものになったと考えられる。図６の固有値の推移からすると，この部分で因子数を３とすることはややためらわれるのであるが，相関行列$\boldsymbol{R}$の全体を考えたときには，ありうるやり方であると思われる。ちなみに，S^2の値は$q_1 = 2$のとき，6.336，$q_1 = 3$では5.710であり，これを相関係数１個あたりの平均値の平方根（平均２乗誤差の平方根で，残差の絶対値の平均よりやや大きめの値となる）で除したものは，それぞれ約0.040，0.038である。この値は小さいようであるが，全体として相関の低い行列（共通性の推定値も含めた２乗平均の平方根がともに0.116程度）なので，いずれもそれほどの説明力ではない。

最後に，問１だけ単独に共通性推定を伴う主因子法を適用した結果を，表10に示す。表７の因子負荷が単なる共通性推定の効果だけであらわれたのでないことは，この第３因子が，表７のものと全く異なり，むしろ主成分分析の結果に近いことからわかる。

5. 結　語

多集合因子分析は，繰り返し述べてきたように，これによって最終的な結論を得るための方法ではない。前節でその一端を示したように，因子数を変えて heuristic にデータを検討するための方法である。

質問紙による尺度構成の研究は，ただ単に何本かの尺度ができればこと足りるのではなく，それを用いて別の領域の尺度との相関関係を検討することを目的としていることが多い。その“別の領域の尺度”の方も自分で構成するとすれば，予備調査の段階で，同一の質問紙に両方を相乗りさせることになるであろう。その際，まずはそれぞれの領域（本研究で言う集合）で項目分析を行ない尺度を固定してから，領域間の相関をとる，というのが通常の手順である。（１節のａの方法にあたる。）

これはもちろん手堅いやり方であり，それなりに妥当であるが，多集合因子分析によれば，分析のごく初期の段階で，２つの領域間の関係についての情報を得ることができる。この情報は，尺度構成のための項目選択にも反映するはずである。更に，集合内で相互に相関する項目が少いために，個別の因子分析によっては見出されなかったような尺度が生まれてくる可能性もある。

例えば適用例における表７の第３因子は，“経済的購買行動”に２つの異なる側面，すなわち単なる“安値指向”と，より高度ないわゆる“賢い消費者”といった面とがあることを示唆している。負荷そのものは余り高くないが，第３因子に最大の負荷を示す，項目４，10，18，19などの内容から，この因子が後者の側面をあらわしていることはうかがい知ることができる。そうすると，この側面を測る尺度を項目の wording を工夫するなどして新たに構成する，という方向が示唆されてくることになる。

このような，普通なら研究の少し先の段階で検討する事柄を先廻りして見ることができ，それによって従来のやり方では得られなかった情報が得られる可能性がある。

なお，適切な利用のために，結果の出力は工夫する必要がある。通常の因子分析では因子負荷の符号や，因子の並んでいる順番はさほど問題にならない（適当な符号反転や入れ替えは頭の中の操作として行なえる）が，集合間の因子得点間相関も含めて解釈するとなると，常に因子の方向と並び順は揃えておく必要がある。筆者は結果を一旦データセットに出力し，それを入力して符号の逆転，配列順の交換等を行ったうえで，英文論文浄書システム（富士通ＡＴＦ）の表作成のためのテキストを出力するプログラムを作成して利用している。（表２～10の原稿はこれによって作られた。）これは大いに解釈の助けになった。*

１節で述べた bandwidth—fidelity dilemma について言えば，これは結局のところ両者のバランスの問題であり，ここで用いた基準S^2は，このバランスのとり方についてややあいまいであることは指摘しておく必要があろう。すなわち，集合の数が多くなるほど，$\boldsymbol{R}$の全体の中に占める異なる集合に属する項目間相関の割合が増し，結果的には bandwidth の方が強調されることになる。この点は，しかし，通常の因子分析でも，因子はそこに含まれる他の項目の文脈の中で決まって来るものであることを考えれば，これは当然のこととも言えるであろう。またアルゴリズム的には，$k = k'$の場合と$k \neq k'$の場合でS^2への寄与を変えるように重みづけをするような修正を行うことは容易である。

最後に，ここで言う“集合”の概念について明確に定義することは難しい。表１でも，問１と問２は，単なる

---

*　このプログラムの作成には小池あき氏の協力を得た。記して謝辞とする。

form の違いにすぎないかもしれない。一方，これらと問3との差は，概念間の一種の hierarchy の差と言うこともできよう。そうすると，その間に単なる相関関係を超えて，因果関係を想定することもありうる。事実近年，因子分析を因果関係の探究の手段と見なすという観点があらわれている（McDonald, 1985)。しかしながら，それは，ここで述べた方法が対象としているよりはもう少し進んだ研究段階，少くとも事前にしっかりした理論モデルがたてられるような段階で，より確認的な方法にもとづいて検討されるべきことがらであろう。

一方，探索的段階でもある程度標本論的配慮は必要かもしれない。単なる標本誤差にすぎないものを因子と誤認する危険性は回避できた方がよい。更に，予備調査なしの一発勝負の調査を行わざるを得ない場合は現実にはかなりあるから，何らかの対策は必要であろう。それにしても，項目水準の分析では多変量正規分布を前提とすることは不可能である。この問題には本研究では全く触れることができなかったが， bootstrap 法（Efron, 1985）のような，シミュレーション的な方法をこの目的に応用することは検討に値すると考えられる。

文　　献

Anderson, T.W. & Rubin, H. 1956 Statistical inference in factor analysis. *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, Vol. V,* 111-150.

Corballis M.C. & Traub, R.E. 1970 Longitudinal factor analysis. *Psychometrika,* **35**, 79-98.

Cronbach, L.J. 1970 *Essentials of psychological Testing,* 3rd ed. Harper & Row.

Efron, B. 1985 The bootstrap method for assessing statistical accuracy. *Behaviormetrika,* **17**, 1-35.

Garman, H.H. & Jones, W.H. 1966 Factor analysis by minimizing residual (MINRES). *Psychometrika,* **31**, 351-368.

Horst, P. 1965 *Factor analysis of data matrices.* Holt, Rinehart & Winston.

Jöreskog, K.G. 1970 A general method for analysis of covariance structures. *Biometrika,* **57**, 239-251.

Kettenring, J.R. 1971 Canonical analysis of several sets of variables. *Biometrika,* **58**, 433-451.

McDonald, R.P. 1985 *Factor analysis and related methods.* Lawrence Erlbaum Associates.

村上　隆　1983　3相データにおける因子変化の記述のための諸方法（Ｉ）──因子分析型のモデル── 名古屋大学教育学部紀要──教育心理学科──，**30**, 145－176.

（1985年８月３日　受稿）

原　　　　　　　　著

ABSTRACT

EXPLORATORY FACTOR ANALYSIS OF MULTISET DATA

Takashi MURAKAMI

A new exploratory method, multiset factor analysis, which factor analyzes data consisting of several sets of variables related to different domains or levels was proposed. Basic equations of this method are written as

$$\boldsymbol{R}_{kk'} = \boldsymbol{A}_k \boldsymbol{S}_{kk'} \boldsymbol{A}_{k'}^{\mathrm{T}} + \boldsymbol{D}_{kk'} \qquad k, k' = 1, \ldots, p \qquad \cdots\cdots (1)$$

where $\boldsymbol{R}_{kk'}$ is n_k by n_k correlation matrix between variables of set k and k', $\boldsymbol{A}_k$ is n_k by q_k factor loading matrix for set k, $\boldsymbol{S}_{kk'}$ is q_k by $q_{k'}$ correlation matrix between factor scores satisfying

$$\boldsymbol{S}_{kk} = \boldsymbol{I} \qquad k = 1, \ldots, p \qquad \cdots\cdots (2)$$

and $\boldsymbol{D}_{kk'}$ is diagonal matrix including unique variances when $k = k'$, and zero matrix when $k \neq k'$

An alternating least squares algorithm is formulated. It assumes that

$$\boldsymbol{A}_k^{*\mathrm{T}} \boldsymbol{A}_k^* = \boldsymbol{I} \qquad k = 1, \ldots, p \qquad \cdots\cdots (3)$$

and

$$\sum_{k'=1}^{p} \boldsymbol{S}_{kk'}^{*} \boldsymbol{S}_{k'k}^{*} = \boldsymbol{L}_k \qquad k = 1, \ldots, p \qquad \cdots\cdots (4)$$

where $\boldsymbol{L}_k$ is diagonal and positive definite. Then following three equations

$$\sum_{k'=1}^{p} \{(\boldsymbol{R}_{kk'} - \boldsymbol{D}_{kk'}) \boldsymbol{A}_{k'}^{*} \boldsymbol{A}_{k'}^{*\mathrm{T}} (\boldsymbol{R}_{k'k} - \boldsymbol{D}_{k'k})\} \boldsymbol{A}_k^* = \boldsymbol{A}_k^* \boldsymbol{L}_k, \qquad \cdots\cdots (5)$$

$$\boldsymbol{S}_{kk'} = \boldsymbol{A}_k^{*\mathrm{T}} (\boldsymbol{R}_{kk'} - \boldsymbol{D}_{kk'}) \boldsymbol{A}_{k'}^{*} \qquad \cdots\cdots (6)$$

and

$$\boldsymbol{D}_{kk} = \mathrm{diag}(\boldsymbol{R}_{kk} - \boldsymbol{A}_k^* \boldsymbol{S}_{kk}^* \boldsymbol{A}_k^{*\mathrm{T}}) \qquad \cdots\cdots (7)$$

are alternated until convergence is reached. Final solutions satisfying (2) are obtained through the transformations

$$\boldsymbol{A}_k = \boldsymbol{A}_k^* \boldsymbol{S}_{kk}^{*\frac{1}{2}} \qquad \cdots\cdots (8)$$

and

$$\boldsymbol{S}_{kk'} = \boldsymbol{S}_{kk}^{*-\frac{1}{2}} \boldsymbol{S}_{kk'}^{*} \boldsymbol{S}_{k'k'}^{*-\frac{1}{2}} \qquad \cdots\cdots (9)$$

This algorithm is a natural extension of ordinary principal factor method.

A questionnaire data collected from housewives in Aichi prefecture was analyzed by this method. It consists of four sets, – buying behavior, attitude for buying, need for information and confidence for the various media. A factor which could not be found by separate factor analysis was derived in the first set and interesting relations between the factor and factors of other sets were revealed.