

# 構造方程式モデリングは，因子分析，分散分析，パス解析の すべてにとって代わるのか？<sup>1</sup>

狩野 裕<sup>2</sup>

Does structural equation modeling outperform traditional factor analysis,  
analysis of variance and path analysis?

Yutaka Kano

It is well-known that structural equation modeling (SEM) can represent a variety of traditional multivariate statistical models. This fact does not necessarily mean that SEM should be used for the traditional models. It is often said that a general model is more difficult to handle than a specific model developed for a given situation. In this paper, we shall clarify relative advantages between SEM and several traditional statistical models. Rather than comparison in mathematical properties, we shall discuss how and when SEM outperforms corresponding traditional models in practical situations. Special attention is paid to statistical analysis of a scale score, the sum of indicator variables determined by factor analysis.

In particular, we shall study relative advantages between (i) confirmatory factor analysis and exploratory factor analysis, (ii) multiple indicator analysis and correlational and regression analysis of scale scores, (iii) analysis of factor means and analysis of variance of scale scores, and (iv) path analysis and multiple regression analysis.

## 1 はじめに

構造方程式モデル (SEM; Structural Equation Modeling)<sup>3</sup>の端緒は，検証的因子分析を提案した Jöreskog (1969) と因子分析に因果分析の考え方を取り入れた Jöreskog (1973) と言ってよいだろう．そして，それらを実現するプログラム (後に LISREL とよばれることになる) の開発も重要な役割を果たした．その後，理論的にもプログラムにおいても大きな発展を遂げたことは周知の事実である<sup>4</sup>．

SEM は多くの既存のモデルを下位モデルとして含むと言われている (e.g., 豊田, 1992) . SEM が因子分析モデルや回帰モデルを含むのは当然として，たとえば，多くの分散分析のモデルも実現可能である．豊田 (2000, 3 章) では，SEM で 3 要因までの完全無作為要因計

---

<sup>1</sup>本論文は，2000 年 3 月に愛知学院大学で行われた日本行動計量学会主催のセミナーで講演した内容を基礎にして大幅に加筆したものである．

<sup>2</sup>大阪大学 大学院人間科学研究科 (Osaka University, Graduate School of Human Sciences) Suita, Osaka 565-0871, JAPAN

<sup>3</sup>共分散構造分析と同義語．近年は構造方程式モデルの方がよく用いられる．

<sup>4</sup>構造方程式モデリングの歴史については清水 (1994) が詳しい．

画や分割法計画，共分散分析などの分散分析モデルが SEM によってどのように実現されるかが詳説されている．しかし，「包含する」ということは直ちに「使うべきだ」ということを意味しない．すなわち，数学的に包含関係があったとしても，得られる情報が同じならば古典的な方法を用いてもよい．また，一般的なモデルは，目的を特化したモデルよりも使い勝手が悪いことがある．実際，普通の回帰分析や分散分析は SEM でも実行できるが，やはり既存のソフトウェアが使いやすいし，SEM では残差分析，多重共線性のチェックや多重比較ができないという感想を聞くことがある．そこで，本論文では，SEM と古典的多変量解析を比較対照することで，お互いの relative advantage を明確化する．数学的な関係にとどまらず，実際の状況で，SEM を用いることで分析者が実際に享受するメリット・デメリットを議論する．

探索的因子分析 (EFA) と検証的因子分析 (CFA) の違いはときに微妙である．両者とも似たような分析結果を出力する場合がある，そのようなときは CFA のメリットはあるのだろうか．また，CFA は EFA を拡張している，すなわち，CFA で分析できるが EFA では分析できないという状況が存在するのだろうか．

SEM はしばしば潜在変数間のパス解析と形容される．実際，測定モデルがあり，測定モデルによって規定された潜在変数間のパス解析という状況で適用されることが多い．しかし，SEM による潜在変数間のパス解析は，従来は「尺度化 + パス解析（重回帰分析の繰り返し）」というスタイルで分析してきた．では，SEM は従来の方法と比較してどのようなメリットがあるのだろうか．また，よく行われている「尺度化 + 分散分析」という方法についてはどうであろうか．

本論文では上記のような問題意識の下で，まず，2 節において EFA と CFA の比較をおこなう．両方の分析が可能な場合とそうでない場合を区別して議論する．続いて，3 節と 4 節では「尺度化 + 相関分析・回帰分析」並びに「尺度化 + 分散分析」と SEM による分析を比較検討する．5 節では重回帰分析の繰り返しという意味でのパス解析と対応する SEM の分析を比較する．6 節は本論文のまとめである．

SEM の発展によって可能になった分析として多母集団の同時分析や因子平均の分析がある．また，従来から存在するがやや標準的でない分析方法が SEM の枠組みで実行できることも指摘しておきたい．たとえば，三相データの分析モデルである PARAFAC (Bentler, Poon, & Lee, 1988)，行動遺伝学における ACE モデル，イプサティブデータの分析 (Chan, 2003)，潜在曲線モデル，潜在構造分析などの離散潜在変数のモデル，項目反応モデルなどである<sup>5</sup>．加えて，SEM で実行できる新しいモデル，たとえば，多変量二段抽出モデル，平均に特色をもたせる三相データの分析モデル (豊田, 2001) や因子分析と分散分析の統合モデル (豊田&真柳, 2001) などがある．Muthén (2002) は “Beyond SEM” というタイトルの論文で，SEM がカバーする広範なモデルを 4 つの枠組みに分けて議論している．SEM が第二世代の多変量解析とか多変量解析の高級言語と形容される理由は，このバラエティに富んだ下位モデル群にある．

本論文では，SEM のこのような特徴とは別に，対象をあえて歴史があり従来の枠組み

---

<sup>5</sup> 豊田 (2000) を参照のこと．

でそれなりに分析できる方法論にしぼった状況で、SEM の価値がどこにあるのかを議論していく。

## 2 探索的因子分析 versus 検証的因子分析

探索的因子分析 (EFA) と検証的因子分析 (CFA) を比較する際、まず大前提として CFA では因子に関する良質な仮説が必要ということがある。この意味で CFA はやや制約的である。一方、CFA は可能だが EFA は実行できないという状況がある。また、両者ともそれなりに分析できるという状況もある。本節では、両分析が可能である場合の比較を 2.1 節で、CFA のみで分析できるという状況を 2.2 節で論じる。

### 2.1 両分析が可能な場合

まずここでは、児玉・品川・茂木 (1978) による知能検査の標準化データ (WISC-R;6 歳児) を因子分析してみる。南風原 (1992) は、このデータを 2 因子と 3 因子の因子分析モデルで解析し、それらの解釈や共通性、信頼性、特殊性などについて論じている。また、豊田 (1992) は WISC-R のデータ (12 歳児) の高次因子分析結果を報告している。

WISC-R は 12 個の下位検査から成っており、それらは 6 個ずつの言語性検査と動作性検査に大別される。われわれはまず、言語性因子と動作性因子を想定した 2 因子モデルによって分析を行った。探索的因子分析 (EFA) と検証的因子分析 (CFA) の結果を表 1 と表 2 に示す。

EFA においては 3 種類の斜交回転 (Harris-Kaiser 法, Promax 法, Procrustes 法) を行ってみた。Procrustes 法は、分析者がもっている因子に関する仮説をターゲットに回転する方法であり、ここでは、本知能検査が想定している構造 — X01 ~ X06 は言語性因子、X07 ~ X12 は動作性因子 — をターゲット行列に指定した。

3 種類の斜交回転に対して因子負荷量の推定値は比較的安定している。しかし、細かくみていくと、「X06 数唱」の因子負荷の大小が回転方法によって変化し、この検査を言語性因子に入れるべきか動作性因子に入れるべきかがはっきりしない。また、因子相関の推定値が回転方法によってずいぶん異なることがわかる。このように、探索的因子分析の欠点の一つは因子回転による解の不安定性にある。

CFA においては、言語性因子と動作性因子が想定された構造から始めて、EFA の結果や LM 検定<sup>6</sup>・Wald 検定を参考にしながら因子の仮説を修正し、表 2 に報告した 2 つのモデルを最終モデルとした。モデル 1 では言語性因子から「X10: 組合せ」へのパスを引いていないのに対してモデル 2 はそのパスは引いている。ここでは、どちらのモデルがより適切であるかは議論しない。そして、主にモデル 1 について議論していく。

表 1 と表 2 の結果を比べると、基本的に同じような分析結果で解釈も大きく違わないように見える。では、CFA を実行する価値はどこにあるのだろうか。

<sup>6</sup>Lagrange Multiplier 検定・Rao 検定、Score 検定ともよばれる。

表 1: WISC-R (6 歳児) 標準化データの探索的因子分析 ( $n = 100$ )  
(最尤法 (ML) による分析)

	Harris-Kaiser 回転		Promax 回転		Procrustes 回転	
	言語性 因子	動作性 因子	言語性 因子	動作性 因子	言語性 因子	動作性 因子
X01 知識	0.58	0.02	0.58	0.04	0.59	0.01
X02 類似	0.52	0.23	0.53	0.24	0.58	0.21
X03 算数	0.13	0.41	0.15	0.40	0.22	0.37
X04 単語	0.78	-0.09	0.77	-0.07	0.76	-0.09
X05 理解	0.77	-0.10	0.76	-0.07	0.75	-0.10
X06 数唱	0.22	0.29	0.24	0.28	0.29	0.26
X07 絵画完成	0.08	0.44	0.11	0.43	0.18	0.41
X08 絵画配列	0.10	0.31	0.12	0.30	0.18	0.28
X09 積木模様	-0.11	0.77	-0.06	0.75	0.07	0.71
X10 組合せ	0.21	0.40	0.23	0.39	0.30	0.36
X11 符号	-0.09	0.44	-0.07	0.43	0.01	0.40
X12 迷路	-0.10	0.54	-0.07	0.52	0.02	0.49
因子相関	1	0.44	1	0.37	1	0.24
	0.44	1	0.37	1	0.24	1

カイ 2 乗値=51.15; df=43;  $p$  値=0.18; AIC=-31.13

因子パターンで空欄となっているのは因子負荷量が 0 に固定されていることを表す。

CFA の一つの価値は因子回転の問題が生じないということである。表 1 でのようにいくつかの回転解の中でどの結果を採用すべきかという問題に悩まなくて済むのである。CFA の具体的な成果として、「X06: 数唱」の行方について、CFA による統計的な観点からは動作性因子により関連が深いという結論が得られている。それは以下の 3 つの統計的証拠による: (i) 動作性因子への因子負荷量が  $\hat{\lambda}_{6,2} = 0.43$  としっかりとした値として推定されていること, (ii)  $H_0: \lambda_{6,2} = 0$  の Wald 検定が有意であり、そのカイ 2 乗値が 15.04 とかなり大きな値になっていること, (iii) 「X06: 数唱」の言語性因子への因子負荷量  $\lambda_{6,1}$  の LM 検定, i.e.  $H_0: \lambda_{6,1} = 0$  が非有意で、そのカイ 2 乗値 2.30 は 15.04 と比較してもずいぶん小さな値になっていることである<sup>7</sup>。もちろん、パスを引くべきかどうかは、統計的な情報のみから判断されることはないが、客観的な情報が得られることは大きなメリットであろう。

実は、先に述べた「CFA においては因子回転の問題が生じない」という言及は、正確には誤りである。CFA では、いくつかの因子負荷量を 0 とおくという因子回転の方法を採っている、というのが正確な表現である。EFA における因子回転の方法の多くは、因子負荷量の 2 乗がよりばらつくように (計算機によって) 回転を定めており<sup>8</sup>、それは経験的に解釈しやすい解を報告すると言われている。しかし、これはあくまでも経験則であって、実質科学的にはその根拠は薄弱であると言わざるを得ない。というのは、「因子負荷量の 2 乗が

<sup>7</sup>両者のパスを入れたモデルにおいて行った Wald 検定においても、 $\lambda_{6,1}$  は非有意で  $\lambda_{6,2}$  は有意となっている。

<sup>8</sup>原点からの距離の 2 乗がばらつくように回転することと同義。因子回転の詳細は芝 (1979) や柳井・繁樹・前川・市川 (1990) などを参照のこと。

表 2: WISC-R(6 歳児) 標準化データの検証的因子分析 ( $n = 100$ )  
(最尤法 (ML) で分析)

	モデル 1: CFA( $\lambda_{10,1}$ なし)				モデル 2: CFA( $\lambda_{10,1}$ あり)			
	因子パターン		Wald or LM Test		因子パターン		Wald or LM Test	
	言語性 因子	動作性 因子	言語性 因子	動作性 因子	言語性 因子	動作性 因子	言語性 因子	動作性 因子
X01 知識	0.60		33.11	0.67	0.59		32.56	0.77
X02 類似	0.47	0.31	16.81	7.26	0.49	0.30	19.76	7.13
X03 算数		0.51	0.19	21.61		0.51	0.44	21.54
X04 単語	0.74		54.62	0.07	0.74		55.17	0.06
X05 理解	0.74		53.22	0.17	0.74		54.34	0.22
X06 数唱		0.43	2.30	15.04		0.42	2.93	14.43
X07 絵画完成		0.50	0.05	21.08		0.51	0.19	21.41
X08 絵画配列		0.36	0.79	9.96		0.36	0.89	10.09
X09 積木模様		0.67	4.41	40.96		0.68	2.34	42.08
X10 組合せ		0.50	3.41	21.16	0.22	0.38	3.58	9.79
X11 符号		0.36	0.69	10.31		0.37	0.47	10.80
X12 迷路		0.45	1.24	16.83		0.46	0.50	16.99
因子相関	1	0.44			1	0.38		
	0.44	1			0.38	1		
	カイ 2 乗値=66.15; df=52 $p$ 値=0.09; AIC=-37.85				カイ 2 乗値=62.78; df=51 $p$ 値=0.13; AIC=-39.23			

よりばらつく」という数学的性質と実際のデータ発生機構との間をつなぐ十分な理論がないからである。2つの異なった仮定からは異なった結論が導かれることがある。2つの異なった因子回転の方法からは異なった推定結果が得られることがある。では、この異なった推定結果はデータ発生過程に関するどのような仮定の違いによって生み出されたのであろうか。この仮定が理解できないとどちらの結果がよりふさわしいのか判断できない。

一方、CFAにおいて因子負荷量を0とおくことは、対応する因子からの影響はないという理解しやすい仮定を意味している。そして、この仮定はLM検定やモデルの全体的評価を行う適合度によって検証可能である。

二つ目は、CFAではパスを引く価値と引かない価値を統計的に判断できることである。EFAでいえば、解釈する因子負荷量と無視する因子負荷量の区別が統計的に判断できるかという問題になる。これは、因子回転のところでも述べたが、因子負荷量=0の検定が可能かどうかに関連している。CFAにおいては、引いたパスに対して引く価値があったかどうかを検討するためのWald検定を、そして、引かなかったパスに対して引かなくてもよかったかどうかを検討するためのLM検定を実行することができる。表2の右側にはそれらの検定統計量の値が出力されている。これらの値を、自由度1の $\chi^2$ -分布の上側5%点(=3.841)と比較しパスの有無を決めるのも一案であるが、この値にこだわる必要はない。より重要な観点は、引かなかったパスのカイ2乗値のすべてが、引いたパスのカイ2乗値のすべてより小さいことであろう。パスを引く価値と引かない価値をカイ2乗値で評価するとき、逆転現象が生じていないということが肝要である。LM検定やWald検定はモデル修正のための情

報だけでなく、モデルの局所評価も与えている。

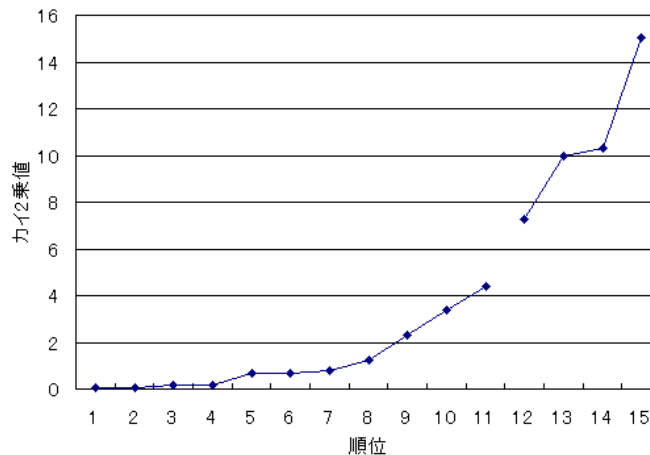


図 1: LM 検定と Wald 検定 (小さい方から 15 個のみの表示)

図 1 では、表 2 のモデル 1 において、推定した因子負荷量に対する Wald 検定統計量と推定しなかった因子負荷量に対する LM 検定統計量のカイ 2 乗値を小さい順に並べ図示している。11 番目と 12 番目の間で折れ線が途切れており、これは、11 番目以前が 0 と固定された因子負荷量の LM 検定の値を、12 番目以降が推定された因子負荷量の Wald 検定の結果を示している。パスを引くか引かないかの統計的判断は、折れ線をどこで切るかに対応する。

EFA においても因子負荷量等の有意性検定は可能である。<sup>9</sup> 実際、SAS 第 8 版ではそのオプションが追加されている。しかし、因子回転による不安定性の上で行われる検定にどの程度信頼がおけるか疑問である。回転方法を変えると検定結果も変わるといことが容易に想像される。

この節での主張は EFA を否定するものではない。EFA と CFA の両者が実行可能でよく似た結果が得られたとき、CFA の価値がどこにあるかを明らかにしたものである。CFA は回転の自由度を殺すだけ十分な因子に関する仮説が必要である。良質な仮説を得ることができないとき EFA が登場する。また、同じ結果が得られるのなら、仮定の弱い EFA の結果の方が望ましいという主張も説得力をもつ。

本節での議論を一言でまとめるならば、より強い仮定が必要な CFA からはより信頼のおける統計的証拠が得られるというまっとうな結論であり、結局は、より強い仮定というデメリットとより強い結論というメリットのどちらを重視するかということになる。

## 2.2 探索的因子分析が不可能な場合

EFA は不可能であるが CFA は可能であるという状況はたくさんある。以下にいくつかの場合を示す。

<sup>9</sup>M. W. Browne による CEFA というプログラムがフリーで配られている。URL は次のとおり：  
<http://quantrm2.psy.ohio-state.edu/browne/>

- (i) 2変数にしか影響しない因子がある場合
- (ii) 誤差共分散が存在する場合
- (iii) 因子が多すぎる場合
- (iv) 因子負荷量や独自性に関する様々な仮説の統計的吟味をする場合
- (v) 複数個の集団の因子分析結果を推測統計的に比較したい場合（多母集団の同時分析）
- (vi) 因子平均に興味がある場合
- (vii) 因子モデルの変形・拡張

以下の節では主に (i)–(iii) を取り上げる。(iv) の例として、同族テストにおいて因子負荷量が等しい、独自性が等しいなどの仮説を統計的に検定できることがある。詳細は豊田 (1992, 10.4.1 節) を参照されたい。(v) と (vi) の詳細は標準的な SEM のテキストを参照されたい (e.g., Bollen, 1989; 豊田, 1998; 狩野, 1997)。なお、(vi) は分散分析との比較でふれる。(vii) に含まれるものとしては、二値データの因子分析、イプサティブデータの因子分析、特殊因子と誤差因子の分離、背景変数のある因子分析モデル (MIMIC モデル) などを挙げることができよう<sup>10</sup>。

### 2.2.1 2変数にしか影響しない因子がある場合

2変数にしか影響しない因子がある場合、すなわち、因子負荷行列のある列において大きな値の因子負荷量が2個しか存在しないとき、識別性の問題<sup>11</sup>が生じ解は不安定になり信用できない。しばしば共通性が1を越える不適解が報告される。このことを実例で示そう<sup>12</sup>。

前節で紹介した WISC-R の分析において、南風原 (1992) は、2因子解に対して『ここでの第2因子のように、異質なものが混じって純粋さが損なわれている因子は、因子数が少なすぎるときに生じやすい』ということを指摘し、3因子を仮定した因子分析も行っている。その他の先行研究においても2因子と3因子の両モデルによる分析結果が報告されている。

ところが、2.1 節で分析した6歳児のデータを3因子でEFAを行うと、反復計算の途中で共通性が1を越え不適解となってしまう<sup>13</sup>。不適解の発生原因としては (i) 標本変動 (ii) 識別性 (iii) 因子モデル不適合などが考えられている (e.g., Van Driel, 1978; 佐藤, 1987; 狩野, 1998a)。不適解に遭遇したとき、分析者がまず行うことは不適解の原因を特定することである。そのための一つの方法は、共通性が1を越えても反復を止めずに推定値を計算し続けてみることである<sup>14</sup>。SASで分析すると、反復は11回で収束するものの、X03の共通性は大きく1を越え111.120と推定された。このように共通性が非常識的に大きな値として推定されたり、また反復が収束しないときは、その不適解は悪性で、比較的罪の軽い「(i) 標本変動」ではないことを示している。

<sup>10</sup> 豊田 (2000) などを参照のこと。

<sup>11</sup> データから共通性を一意に決定できないことをいう。

<sup>12</sup> 吉田・狩野・原田 (in press) も参照されたい。

<sup>13</sup> 最尤法，最小二乗法，反復主因子法などで不適解となる。

<sup>14</sup> SASではPROC FACTORにultraというオプションを付加する。

不適解の原因を追求するためにさらなる解析を行う．表 3 に非反復主因子法で求めた初期解（未回転）と VARIMAX 直交解を報告している．初期解の第 3 因子 (F3) において，大

表 3: 主因子解

	初期解			VARIMAX 直交解		
	F1	F2	F3	F1	F2	F3
X01 知識	0.51	-0.29	0.00	0.55	0.08	0.20
X02 類似	0.65	-0.14	-0.04	0.55	0.28	0.25
X03 算数	0.49	0.12	<u>0.46</u>	0.14	0.14	<u>0.65</u>
X04 単語	0.57	-0.41	-0.13	0.71	0.08	0.09
X05 理解	0.56	-0.40	-0.12	0.69	0.08	0.10
X06 数唱	0.45	-0.01	<u>0.37</u>	0.22	0.07	<u>0.53</u>
X07 絵画完成	0.45	0.21	0.04	0.16	0.38	0.29
X08 絵画配列	0.35	0.15	-0.24	0.19	0.40	0.01
X09 積木模様	0.53	0.40	0.04	0.08	0.55	<u>0.36</u>
X10 組合せ	0.50	0.10	-0.11	0.30	0.39	0.16
X11 符号	0.29	0.33	-0.17	0.01	0.47	0.05
X12 迷路	0.37	0.40	-0.20	0.03	0.57	0.07

きな値と言えるのは「X3: 0.46」と「X6: 0.37」の2個であろう．一方，VARIMAX 回転解では「X9: 0.36」も結構大きく3個の因子負荷が大きな値をとっている．このように『2変数にしか影響しない因子がある』という基準は因子回転に依存する．より正確に定義を述べるならば『適当な回転を行ったときに2変数にしか影響しない因子がある』となり，このとき推定に問題が生じる．逆に述べるならば『どんな回転を行っても各因子が3変数以上に影響する』ならば問題は起こらないと考えてよい<sup>15</sup>．

以上の考察により，WISC-R(6歳児)のデータの構造は第3因子がX03とX06にのみ大きな影響をもつ因子であるがゆえに，識別性の問題によって不適解が生じていると考えられる．この事実から，2因子の因子分析モデルでは，X03とX06の相関が十分に説明できていないことが示唆される．実際，2因子解での残差（標本相関行列と2因子モデルに基づく相関行列との差）を見てみると，(3,6)成分の値は0.24830で最大である<sup>16</sup>．

以上の考察により以下のことが結論される．WISC-R(6歳児)のデータは，2因子のモデルでは「X03: 算数」と「X06: 数唱」の相関 $r_{36}$ が十分に説明できず，3因子解における第3因子は，実質的にこの相関 $r_{36}$ のみを説明するため大きな因子負荷量が2つしかなく，識別性の問題が生じ不適解となる．

EFAで可能な分析はここまでである．一方，CFAはこのような構造のデータを見事に分析してみせる．すなわち，言語性因子を構成するのはX01, X02, X04, X05の4変数，動作性因子を構成するのは当初の想定どおりX07-X12の6変数，そして，X03とX06の背後に3番目の因子を想定した．分析結果を図2の(1)に与える．なお，本モデルは豊田(1992,199頁)のモデルと同値関係にある．南風原(1992)では，第三の因子の解釈として「転導性から

<sup>15</sup> 数学的には必要条件にすぎないが，ここではこれ以上議論しない．詳細は Anderson & Rubin (1956) を参照のこと．

<sup>16</sup> SASで残差行列を出力するには，オプション residuals を付加する．



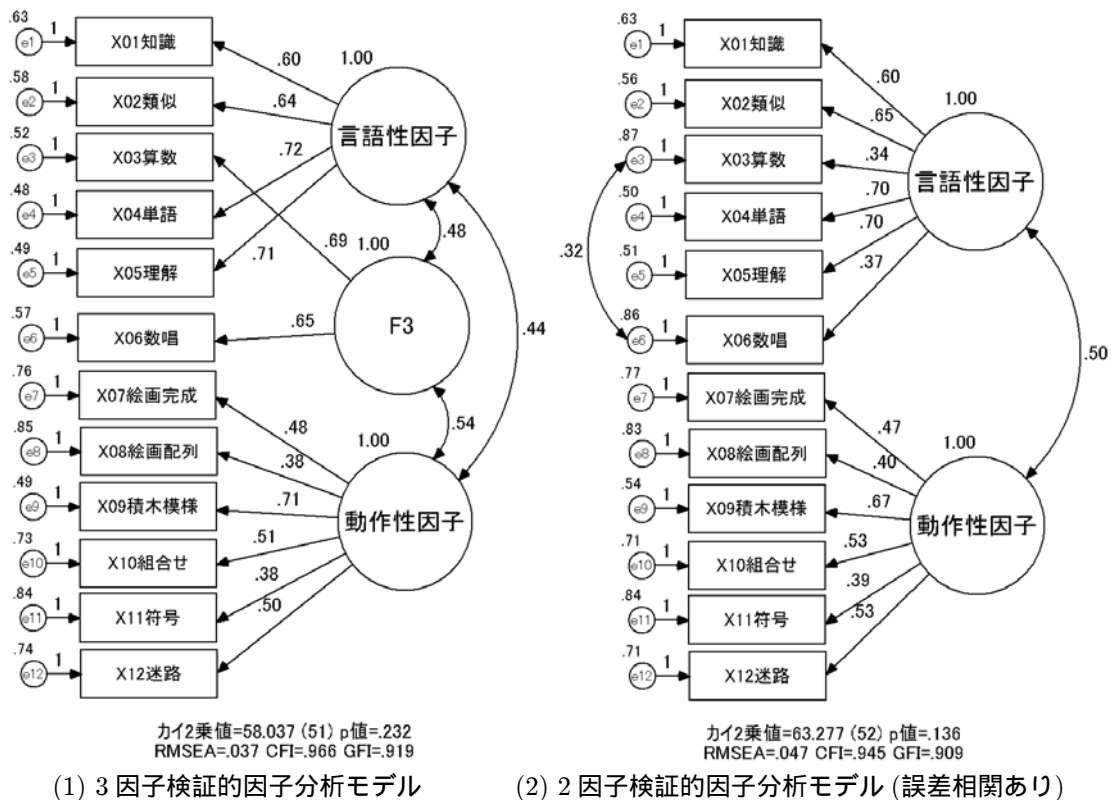


図 2: 検証的因子分析のモデルと推定結果

の解放」「短期記憶」「数量的な能力」を紹介している。

図 2 の (2) では別のモデルを紹介している。X01-X06 を言語性因子の下位検査，X07-X12 を動作性因子の下位検査であることを所与としたモデルである。しかし，今までの分析から，X03 と X06 の背後にはこれらの因子以外の要素が存在するのは明らかであるから，その影響を誤差共分散で拾っている<sup>17</sup>。このように，CFA では，因子として取り上げるほどではないと考える共変動を誤差共分散としてモデリングすることができる<sup>18</sup>。これも CFA のメリットである。誤差共分散を正確に推定することは，次節で述べる尺度の信頼性の評価にも重要である。

変数 X03 と X06 の因子 F3 への因子負荷量を固定して，主要 2 因子を EFA によって抽出するというモデルも可能である<sup>19</sup>。

### 2.2.2 尺度と信頼性：誤差共分散の扱い

因子分析を実行して因子ごとに観測変数（項目）を分類する。一つの因子に関わる観測変数群は意味的なまとまりをもつので，そこから因子の命名が行なわれる。観測変数群を，

<sup>17</sup> 誤差共分散を外すとモデルは適合しない。

<sup>18</sup> 分析の道具として EFA しか持ち合わせていなかった時代は，誤差共分散の影響は無視されていたのではないと思われる。

<sup>19</sup> 分析には CFA と EFA の因子回転を用いる。詳細は Kano(1997) を参照されたい。

その意味を測る「ものさし」と見なし尺度といい、観測変数（に対する回答）の合計を尺度得点という。事前の項目検討から始まって、因子分析を経由し尺度得点を求める一連の作業を尺度化という。尺度化を行うと変数の数が減少するのでその後の分析がしやすくなる。次元縮小のために変数を分析から外すと、せっかく採ったデータが無駄になり情報を適切に活用できない。尺度化はすべてのデータ（情報）を使っているという精神的満足感も与えてくれる。

尺度の信頼性とは、尺度  $X$  を真の得点  $x$  と誤差  $e$  とに分解したとき、尺度の分散に占める真の得点の分散の割合をいい、

$$\frac{V(x)}{V(X)} = 1 - \frac{V(e)}{V(X)} \quad (2.1)$$

で定義される。1 因子の因子分析モデルは

$$X_i = \mu_i + \lambda_i f + u_i \quad (i = 1, \dots, p) \quad (2.2)$$

と表され、通常、 $V(f) = 1$  として  $f, u_1, \dots, u_p$  は互いに無相関という仮定がおかれている。因子分析モデルの下では、尺度  $X = \sum_{i=1}^p X_i$  における真の得点は  $\mu_i + \sum_{i=1}^p \lambda_i f$  と考えられることが多く、このとき、信頼性は

$$\rho = \frac{(\sum \lambda_i)^2}{(\sum \lambda_i)^2 + \sum V(u_i)} \quad (2.3)$$

で見積もることができる。

信頼性の測度として最もしばしば用いられているものに 係数がある (Cronbach, 1951)。

係数は観測変数の分散・共分散の簡単な関数であるため、係数の算出に因子分析を必要としないという計算上のメリットがある<sup>20</sup>。1 因子の因子分析モデルの仮定が正しいとき、係数は (2.3) による  $\rho$  の下限を与え、等号は  $\lambda_1 = \dots = \lambda_p$  のときに成立する<sup>21</sup>。

言語性因子を測定する尺度項目として X01-X06 を考える。これらの言語性検査の CFA による分析結果を図 3 に示す。言語性因子の尺度 (X01-X06) の 係数の推定値は 0.747 である。しかしこの値は、言語性因子とは無縁と考えられる X03 と X06 の共変動によって水増しされており不正確である。では、(2.3) による  $\rho$  はどうか。これも機能しない。というのは、誤差共分散をモデリングしない EFA では、X03 と X06 の共変動を (無理やり) 説明するため因子負荷量が過大推定されるので、それが  $\rho$  の値を押し上げるからである。実際、 $\hat{\rho} = 0.746$  であり 係数とほとんど変わらない。

図 3 のモデルに基づいて誤差共分散を正確に評価し、それを  $\rho$  の分母に反映する公式

$$\rho' = \frac{(\sum \lambda_i)^2}{(\sum \lambda_i)^2 + \sum V(u_i) + \sum_{i \neq j} \text{Cov}(u_i, u_j)} \quad (2.4)$$

を用いて信頼性を求めると  $\hat{\rho}' = 0.703$  となる。誤差共分散が存在する場合、係数や EFA に基づく信頼性係数が不正確であることが見てとれるだろう。

<sup>20</sup>しかし、昨今の計算機の発達はこのメリットを失わせている。

<sup>21</sup>この条件は essential  $\tau$ -equivalence といわれる。2.2 節で述べたように、このような因子負荷量に関する仮定を検定するのにも CFA が有効である。

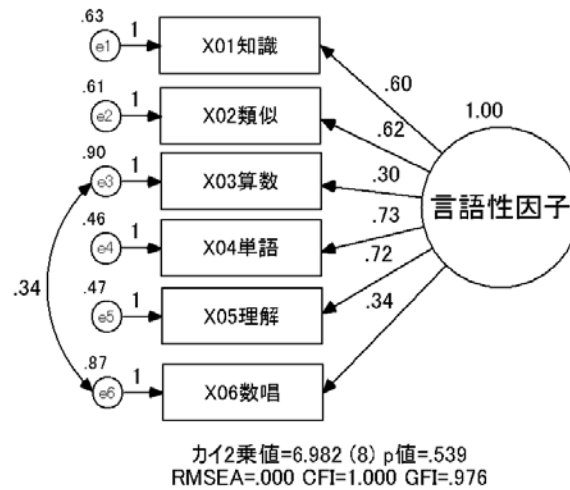


図 3: 言語性検査の検証的因子分析

EFA を実行した際、誤差共分散の存在によってモデルの適合度が悪くなる。実際、図 3 のモデルで共分散を外すと  $\chi^2 = 22.111(df=9)$ ,  $p$  値=0.009 となる。したがって、無批判な  $\alpha$  や  $\rho$  の適用を躊躇するかもしれないが、EFA の枠組みでは何もできない。誤差共分散をモデリングできる CFA が不可欠である。Kano & Azuma (2003) は CFA を基礎におき、誤差共分散をモデリングして信頼性を (2.4) によって評価するプログラムを開発している。

ところで、因子分析モデルを正確に書き下すと

$$X_i = \sum_{k=1}^m \lambda_{ik} f_k + s_i + e_i \quad (i = 1, \dots, p)$$

である。ここで  $s_i$  は特殊因子であり  $e_i$  は誤差変数である<sup>22</sup>。信頼性分析では、 $s_i$  の影響は真の得点の変動へ含めるべきであるが、一般には  $s_i$  と  $e_i$  が区別できないので、 $s_i$  の影響は誤差とみなされることが多い。上記の分析でもそのようにしている。南風原 (1992) はこの点について詳しい議論を行っており、下位検査の信頼性を折半法で求め、信頼性から共通性を減じることで特殊性  $V(s_i)$  を導出している。SEM は、独自性の分離に関して、CFA モデルを発展させた高次因子分析モデルを用いてより洗練された方法論を提供するが、紙幅の関係でここでは省略する。興味のある読者は、豊田 (1998) や狩野 (1998b) を参照されたい。

### 2.2.3 因子が多すぎる場合 (Ledermann の限界)

EFA で抽出できる因子の数には制限がある。具体的には、観測変数の数を  $p$  とすると

$$\text{因子数} \leq (2p + 1 - \sqrt{8p + 1}) / 2 \quad (2.5)$$

である。これを Ledermann の限界という (Ledermann, 1937)。いくつかの  $p$  について抽出可能な因子数の上限を表 4 に示してある (e.g., 丘本, 1986, 21 頁) :

<sup>22</sup>(2.2) は、 $u_i = s_i + e_i$  としたものである。

表 4: 抽出可能な因子数の上限

観測変数の数 $p$	2	3	4	5	6	7	8	9	10	...
因子数の上限	0	1	1	2	3	3	4	5	6	...

Ledermann の限界を超えた因子数に対しても，因子負荷量に関する何らかの仮説があれば，検証の方法によって因子分析を行うことができることがある．以下に，シンプレックスモデル（図 4），円環モデル（図 5），潜在曲線モデル（図 6）の例を紹介する<sup>23</sup>．

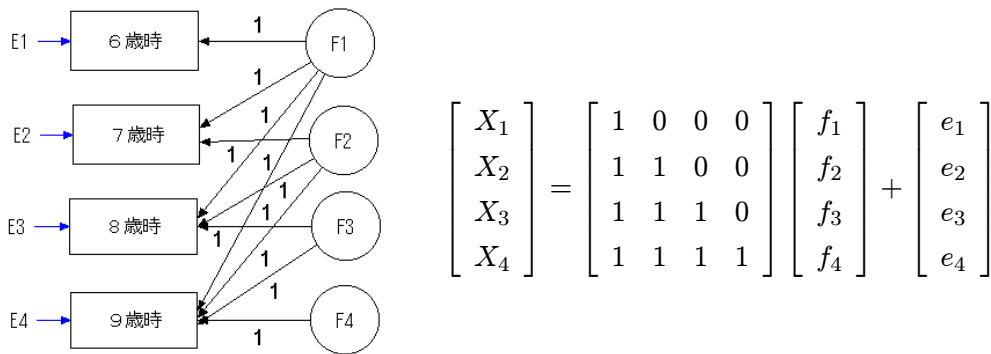


図 4: シンプレックスモデル

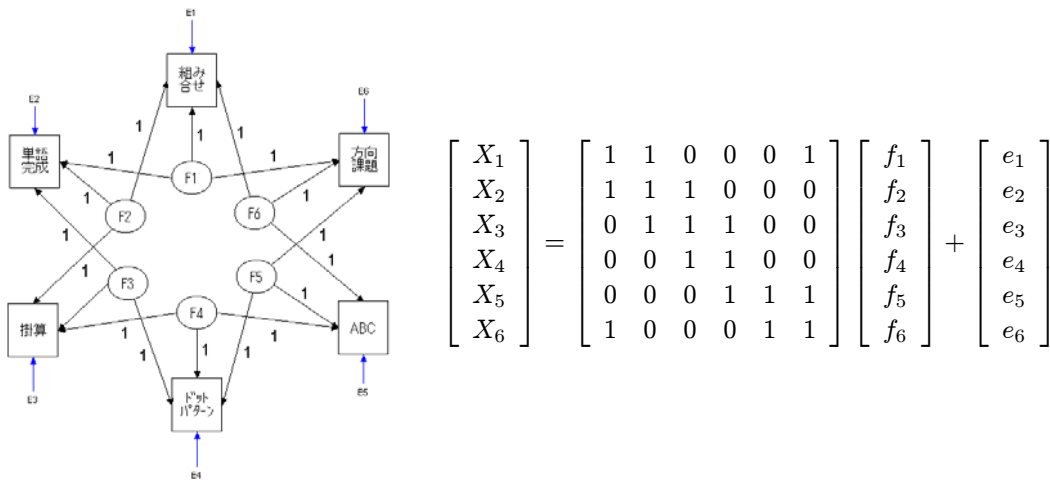


図 5: 円環モデル

<sup>23</sup> シンプレックスモデルと円環モデルは豊田 (1992) からの引用である．

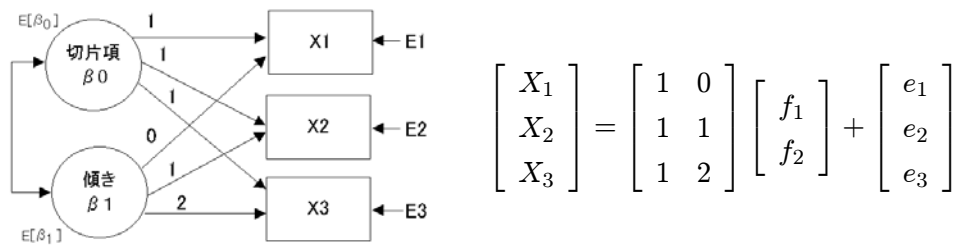


図 6: 潜在曲線モデル

これらのモデルでは因子の数が Ledermann の限界を超えており，探索的因子分析を実行することができない．しかし，因子負荷行列はいずれも固定パラメータのみから構成されており，この性質を使うとモデルは同定可能である．そして，それを実現するのが検証的因子分析であるといえる．

### 3 「尺度化 + 相関分析・回帰分析」と SEM

2.2.2 節でふれたように，尺度構成は因子分析の重要な応用である．因子に関連が深い観測項目の合計得点を尺度得点とし，尺度得点間で相関分析や回帰分析が行われてきた．社会科学におけるこのような伝統的研究手法に問題点があるのだろうか．もしあるとすれば SEM はそれをどのように解決するのだろうか．本節では，これらの論点に関して相関分析と回帰分析のそれぞれについて議論を深めていく．

#### 3.1 相関分析

2 節で分析した WISC-R のデータを再び取り上げよう．言語性因子と動作性因子との間の因子間相関に興味があるものとする．比較すべき手法は CFA によって因子間相関を推定する方法 (図 2) と尺度化 (下位検査の合計得点) を行ってから相関を計算する方法である．また，言語性因子の指標 (尺度) として 6 項目すべてを用いる場合 (図 2(2)) と X03 と X06 を除いた 4 項目を用いる場合 (図 2(1)) を区別して比較する．分析結果を尺度の信頼性係数と合わせて表 5 に示してある．尺度化による相関係数の推定値は CFA と比較してずいぶん小さいことがわかる．この現象を「相関係数の希薄化」または「誤差がコントロールされていない」という．

なぜこのような小さな値に推定されるのであろうか．この問題は尺度の信頼性と大いに関連している．二つの尺度の構造を  $X = x + e$ ,  $Y = y + u$  とし， $\rho_X$  と  $\rho_Y$  をそれぞれ尺度  $X$  と  $Y$  の信頼性係数とすると，尺度間相関  $\text{Cor}(X, Y)$  と真の得点間の相関  $\text{Cor}(x, y)$  との関係を決めることができる：

$$\frac{\text{Cov}(X, Y)}{\sqrt{V(X)V(Y)}} = \frac{\text{Cov}(x, y)}{\sqrt{(V(x) + V(e))(V(y) + V(u))}}$$

表 5: 因子間相関の比較と信頼性

言語性尺度の 項目数	因子間相関		尺度の信頼性		
	検証的因子分析	尺度間相関	尺度名	$\alpha$ 係数	$\rho$ or $\rho'$
4	0.444	0.354	言語性尺度	0.764	0.766
			動作性尺度	0.661	0.668
6	0.495	0.405	言語性尺度	0.747	0.703
			動作性尺度	0.661	0.668

$$\begin{aligned}
 &= \frac{\text{Cov}(x, y)}{\sqrt{V(x)V(y)}} \sqrt{\frac{V(x)}{V(x) + V(e)} \cdot \frac{V(y)}{V(y) + V(u)}} \\
 &= \frac{\text{Cov}(x, y)}{\sqrt{V(x)V(y)}} \sqrt{\rho_X \rho_Y} \quad (3.1)
 \end{aligned}$$

(3.1) から、尺度間相関(合計得点による相関)は、真の得点間の相関を信頼性係数の幾何平均倍したものであることがわかる。したがって、信頼性が低い尺度では希薄化は大きな問題となる。希薄化は、回帰分析を実行したときの偏回帰係数においても生じる(次節参照)。尺度化をする際「信頼性が重要である」と教えられる理由はここにある。

希薄化というバイアスは、内的整合性のある尺度項目を多く採り尺度の信頼性を高めることによって緩和される。CFA のメリットは、理論的には二個以上尺度項目があれば誤差をコントロールすることができ、相関係数や偏回帰係数をバイアスなく推定できるところにある。つまり、十分多くの尺度項目を構成できず信頼性が低くとも、CFA でモデル化を行えば、誤差をコントロールして希薄化されない正しい推定値を求めることができるのである。

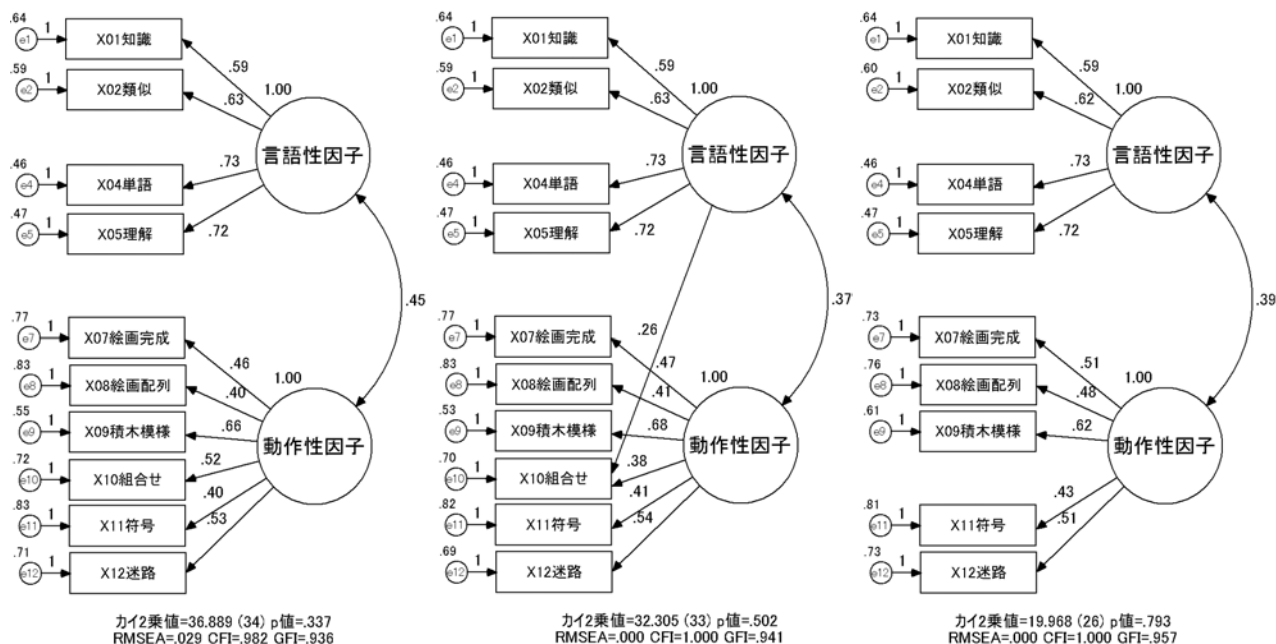
さて、CFA による因子間相関と合計得点による尺度間相関の関係 (3.1) を WISC-R のデータで確かめてみると以下ようになる：

$$\begin{aligned}
 0.354 &\approx 0.318 = 0.444 \times \sqrt{0.766 \times 0.668} \\
 0.405 &\approx 0.340 = 0.495 \times \sqrt{0.703 \times 0.668}
 \end{aligned}$$

残念ながら、理論と現実とでは乖離があるようで、両辺はお世辞にも近いとはいえない。

これらの乖離が生じる理由として、データとモデルとの間に距離があることが考えられる。すなわち、図 2 のモデルから導かれた相関行列がデータの相関行列を十分に近似していないのである。しかし、カイ 2 乗値を始めとする種々の適合度指標はよい適合を示していたのではないかという反論がありえるが、サンプルサイズ ( $n = 100$ ) が十分大きくないこと、よく検討してみると、GFI の値 (0.919, 0.909) は十分大きいとはいえない。そこで、表 3 の VARIMAX 解や LM 検定を参考にして、言語性因子から「X10: 組合せ」へパスを引いてみた。現在興味のない F3 を削除し、このパスを引く前と引いたモデルの両分析結果を図 7(1)(2) に示す。

このパスを引くべきかどうか、すなわち、データ(相関行列)が真実を表しているのか、単純構造のモデルが信じるべきものであるかの最終決定は心理学者に委ねることになる。こ



(1) 単純構造の場合 (2) 単純構造でない場合 (3) 単純構造を崩す項目を削除

図 7: 単純構造を崩す項目の扱い

ここでは、パスを引くことによる言語性因子と動作性因子の相関の違いに注目する。すなわち、パスを引かないモデルでは  $\widehat{Cor}(\text{言語}, \text{動作}) = 0.45$  であるのに対して、パスを引いたモデルでは  $\widehat{Cor}(\text{言語}, \text{動作}) = 0.37$  となって大きな食い違いが生じている。もしこのパスが不要であるならば相関係数の推定値間にこのような差異が生じないはずであるから、統計的にはパスを引くモデルが正しいということになる。そして、この例は、必要なパスを無視すると推定値にバイアスが生じることを示している。

さて、図 7(2) のモデルが正しいとしたとき、尺度化はどのようにして行えばよいのだろうか。2 つ以上の因子に大きな因子負荷量をもつ観測変数は、因子分析の段階で分析から外すことがある。そうする第一の理由は、そのような観測変数は因子分析のモデルに適合しておらず、分析から外すことによってモデルの適合度がぐっと改善されることが経験的に多いからである。もう一つの理由は尺度化する際そのような観測変数の扱いに困るからである。もし、「X10: 組合せ」を両方の尺度に入れ相関係数を求めたならば、同じ項目が入っている尺度間の相関係数を求めることになり違和感を感じるだろう。しかし、観測変数がモデルに合っていないときは分析から落すのは仕方ないが、尺度化に困るからといってモデルに合っている変数を削除してしまうのはいかにももったいない気がするのは筆者だけであろうか。いずれにしても、とりあえずは「単純構造を崩す項目はどの尺度にも含めない」という常套手段に従うことにしよう。

図 7(3) は X10 をモデルから外したときの CFA の結果である。興味の対象である因子相関は 0.39 と推定されており、図 7(2) と大きな変化はない。このように、CFA の大きな特徴は、モデルが適合しているならば、変数の削除によって分析結果が大きな変化を受けない

ということである。逆に言えば、単純構造を乱す変数が存在してもそれを適切にモデリングすれば、分析に問題は生じないということになる。

他方、尺度化してから相関係数を計算するという方法はどのようなになるか。まず、既に指摘したように単純構造を乱す変数は削除せざるを得ない。そして、変数を削除すると、一般に信頼性が低下する。実際、X10 を削除することによって、動作性尺度の信頼性係数  $\rho$  の推定値は 0.668 から 0.640 に低下する。そして信頼性の低下が希薄化をさらに進めることになる。尺度間相関は 0.294 となり、図 7(3) による推定値 0.394 より大幅に下回るのである。

蛇足であるが、尺度間相関と真の得点間の相関の間に成立する公式 (3.1) を確かめておこう。

$$0.294 \approx 0.276 = 0.394\sqrt{0.766 \times 0.640}$$

今度は、近似的に (3.1) が成立しているといっておいてよいだろう。

ここまでは、尺度の信頼性という観点から CFA の価値を示してきたわけであるが、実は、尺度において最も重要な性質は (信頼性ではなく) 内容的妥当性である。内容的妥当性は尺度項目が本来測定したいものをきちんと測定しているかという基準であり、言い換えると、構成概念の意味する様々な側面をもらさずに測定しているということを要求している。その定義から信頼性とは整合的でないこともあるが、尺度構成において重要なのは、内容的妥当性を保った下で如何にして信頼性を向上させるかということであろうか。内容的妥当性を強調すればするほど信頼性を高めることが難しくなり、この事実は、CFA の重要性をさらに強調することになる。

さて、本節では (観測変数の単純合計である) 尺度得点によって構成概念間の相関を推定する方法を取り上げたが、他の方法として、観測項目の重み付けられた合計得点である因子得点を用いる方法や EFA の斜交解によって因子間相関を推定する方法がある。これらの方法においては単純構造を乱す変数を削除する必要がなく希薄化の影響も受けない。また、因子得点は、尺度得点 (単純合計) と比べて予測の精度が高いことがわかっている。ではこれらの方法にはどのような問題があるのだろうか。まず両者に係わる問題として、これは 2.1 節でも述べたが、因子回転による結果の不定性がある (具体例として表 1 の因子間相関をみよ)。この問題に加えて、因子得点を用いる方法では、因子得点の推定方法が多種多様存在し、それらによって異なった結果が得られることがある。したがって、CFA が可能な場合は CFA を適用する方がよい。因子得点の問題点については 4 節で再びふれる。

### 3.2 回帰分析

ここでは、尺度化後における回帰分析について考える。あまり現実的ではないが、前節の例を用いて、図 7(3) において「言語性因子」から「動作性因子」への回帰モデルを考えると、相関分析とまったく同じ結論を得ることができる。すなわち、SEM では標準化回帰係数が 0.394 と推定されるが、尺度得点間で単回帰分析を行うと、標準化回帰係数は 0.294 となる。相関分析と同様に「尺度化 + 回帰分析」も希薄化の影響を受けるのである。

重回帰分析を用いるときは状況はより深刻になる。というのは、信頼性が異なると説明



変数の統計的な比較ができなくなるからある．説明変数の比較は偏回帰係数の検定や標準化偏回帰係数の大きさに基づいて判断されることが多い．しかし，ある説明変数の偏回帰係数が有意でなく標準化偏回帰係数の推定値が小さな値であったとしても，それは単に信頼性が低いというのが理由であって，その説明変数の真の得点は有意に結果変数に影響しており，他の説明変数と同等（またはそれ以上）に重要かもしれないのである．

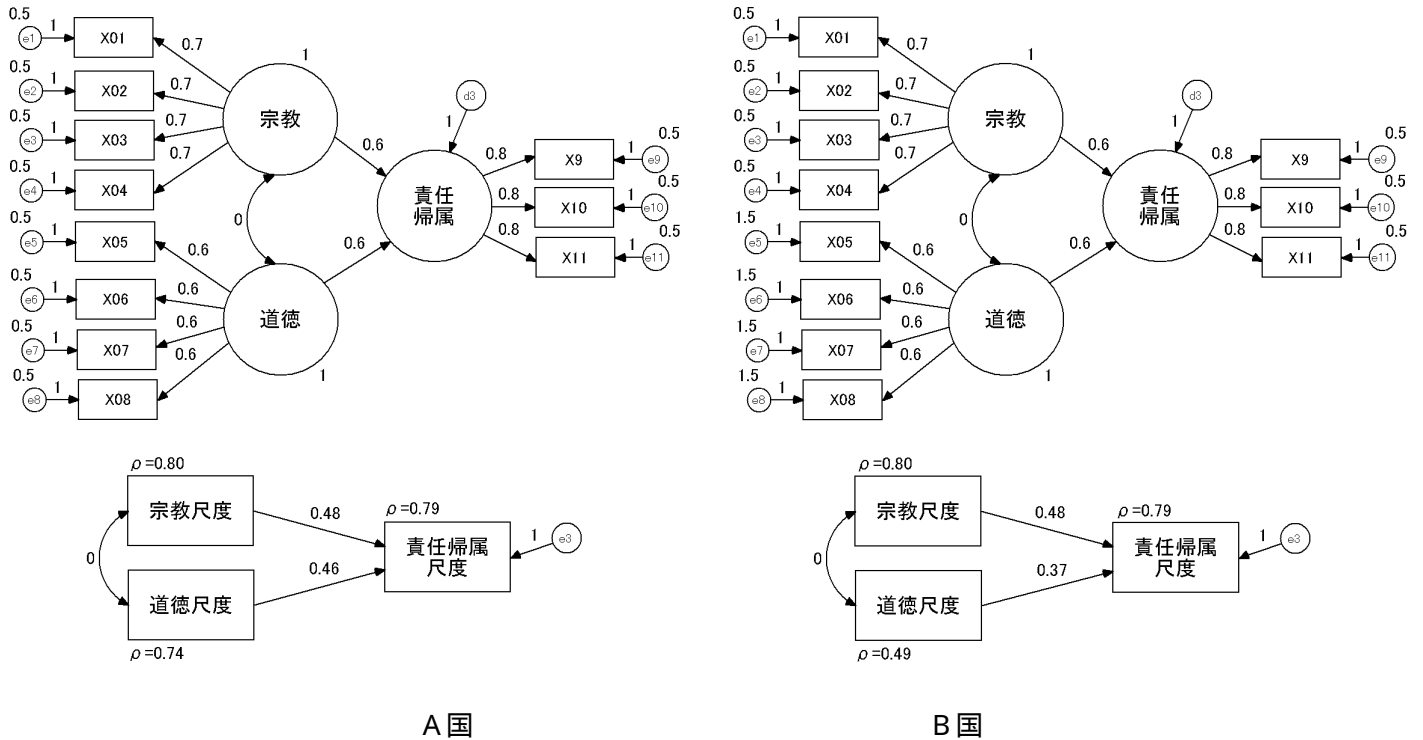


図 8: 構造方程式モデル (上) versus 尺度化 + 回帰分析 (下)

図 8 のモデルは、「宗教」と「道徳」の人工妊娠中絶に関する「責任帰属」への影響を，A 国と B 国で比較することを目的として作成されている<sup>24</sup>．3 つの潜在変数の測定モデルは A 国と B 国との間で等しく測定不変が成立している．また，潜在変数間のパス係数もすべて等しく 0.6 となっている．国間で異なる点は「道徳」の誤差である．B 国の誤差分散が A 国のそれと比べて 3 倍になっている．

図 8(下) には，尺度化された 3 変数を用いて回帰分析を行った結果がパス図で示されている．分析結果から，A 国の「尺度化 + 回帰分析」はそれなりに SEM の結果を近似するが，B 国では「道徳」の偏回帰係数が 0.37 と小さな値として推定され，他の偏回帰係数が 0.48, 0.46 と推定されているの比べると，無視できないバイアスが生じていることがわかる．このことから，B 国の「道徳」の影響は，B 国の「宗教」の影響より小さく，また，A 国の「道徳」と「宗教」よりも小さいという誤った結論が導かれる可能性がある．

その原因は誤差の大きさ，すなわち尺度の信頼性にある．回帰分析結果の各変数 (尺度)

<sup>24</sup>社会心理学者から相談を受けたモデルである．ただし，パス係数等は筆者が与えた仮定量である．

には尺度の信頼性係数  $\rho$  の値が示されている<sup>25</sup>。B 国の「道徳尺度」の信頼性係数が  $\rho = 0.49$  と低く、これによって著しい希薄化が行われたため、このような問題が生じたことが理解される。なお、潜在変数である「宗教」と「道徳」との相関が 0 に設定されているが、これは計算を簡単にするためであり本質的ではない。

SEM の大きな特徴は、構成概念を規定する測定モデルと研究対象である構造モデルの分離にある。測定モデルは、潜在変数として構成概念を定義すると同時に誤差を分離する。誤差の分離は、従来の方法論では、尺度化して高い信頼性を確保し誤差を小さくすることに対応する。尺度の信頼性が高ければ、従来の方法でも適切な結論を得ることは可能である。しかし、常には内的整合性の高い尺度項目を十分に用意できるとは限らず、また、複数個の尺度を扱うときはそれらの信頼性が揃っている保証はない。これらを調整して分析するには SEM が必須であろう。

最後に、多母集団の同時分析では非標準解がより有用であることを注意しておく。

## 4 「尺度化 + 分散分析」と SEM

理論的には、多くの実験計画によるデータを SEM で分析することができる。実際、豊田 (2000) にはそれらの実行例が紹介されている。SEM であればセルごとに分散が不均一という場合も分析することができる。しかし、現在のところ、SEM は分散分析 (ANOVA) を得意とはしていない。特に一変量の ANOVA では汎用ソフトウェアで簡単に分析できる場合は SEM を持ち出すまでもない。というのは、構造方程式モデルの「ソフトウェアの仕様」が ANOVA を主な分析と位置付けていないためモデルプログラムに工夫が必要だからである。要因が 2 水準である場合は比較的扱いやすい。しかし、3 水準以上の要因や交互作用が存在する場合は、SEM ではデザイン行列をデータとして入力しなければならない。

多変量分散分析 (MANOVA) でも同様の困難がある。しかし、MANOVA の場合、SEM を使うメリットとして、観測変数の共分散行列に種々の構造を簡単に設定できることがある。構造としてたとえば因子分析モデルを想定することができる。通常の MANOVA では、観測変数の共分散行列にはすべての分散・共分散を推定するという非効率的なモデル<sup>26</sup>が設定されている。もし、適切な構造を導入することができれば、検定の検出力が向上する可能性がある。

本節では、主に多変量データを尺度化して次元に落としてから一変量の ANOVA を実行する状況を考え、「尺度化 + 分散分析」の代案である SEM による因子の分散分析と比較検討する。

---

<sup>25</sup> 誤差共分散がなく因子負荷量が等しいので、essential  $\tau$ -equivalence が成立している。したがって  $\alpha = \rho = \rho'$  である。

<sup>26</sup> 飽和モデルという。

#### 4.1 ANOVA・MANOVA の SEM での実行方法

ここで分析するのは6個の心理テストのデータである<sup>27</sup>。テストの内容は「視覚的認知」「空間視覚」「方向認知」「文章理解」「文章完成」「語彙」であり、分析の目的は男女差の検討である。

基本的な分析の道具は、性差を全変数で同時にみる MANOVA と各変数ごとにみる ANOVA ないしは t-検定である。汎用統計ソフトウェア (SPSS) によって MANOVA を実行すると

$$F_{138}^6 = 2.107, p \text{ 値} = 0.056 \quad (4.1)$$

となり、性差が見られるかどうかは微妙である。各変数ごとの ANOVA では、「文章理解」のみ  $p$  値が 0.022 となり 5% で有意差がみられる。ただし、この結果は検定の多重性を考慮していない。

MANOVA を SEM のフレームワークの中で実行するには、図 9 のようなモデルを考え、このモデルのカイ 2 乗値を評価する。パラメータ  $m_i$  は観測変数の切片と見なすことができ、パス係数の有意性  $H_0 : m_i = 0$  の検定は、各変数ごとの性差の検定 (ANOVA) に他ならない。ついで、

$$m_1 = m_2 = m_3 = m_4 = m_5 = m_6 = 0 \quad (4.2)$$

としたモデルのカイ 2 乗値を求める。二つのカイ 2 乗値の差を取って自由度 6 のカイ 2 乗分布と比較する<sup>28</sup>ことで MANOVA が実行できる。

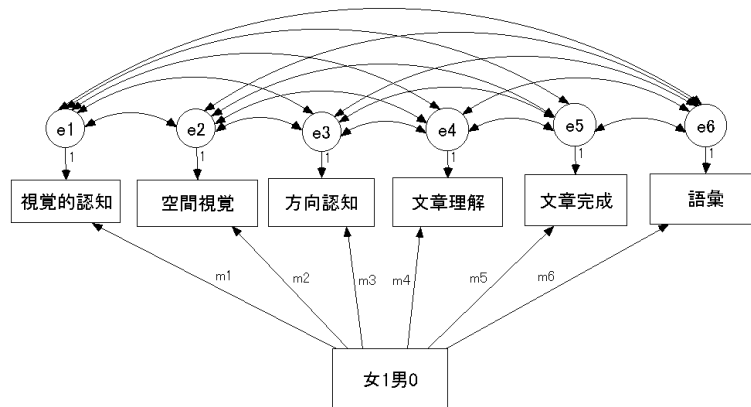


図 9: 構造方程式モデルによる ANOVA・MANOVA

6つの誤差間にはすべて共分散を許す双方向のパスが描かれている。これによって観測変数間に構造を入れない飽和モデルが設定されたことになる。したがって、 $m_i$  を自由なパラメータとして推定されたモデルは、性別を表す変数を込みにしても飽和モデルになりその

<sup>27</sup> AMOS マニュアル (Arbuckle & Wothke, 1999) や狩野 (1997) には同データの多母集団の同時分析の結果が報告されている。

<sup>28</sup> これを difference test という。

カイ 2 乗値は 0 である．よって，difference test は，(4.2) の下でのカイ 2 乗値と等しくなる．SEM による性差の検定は，

$$\chi_6^2 = 12.624, \quad p \text{ 値} = 0.049 \quad (4.3)$$

となる．各変数ごとの ANOVA は，パス係数の Wald 検定を行うか，各変数について  $m_i = 0$  としたモデルと図 9 のモデルとで difference test を実行すればよい．

さて，MANOVA による検定 (4.1) と SEM による検定 (4.3) とでは少し結果が異なる．ここではその理由を調べてみることにする．いま， $p (= 6)$  を観測変数の数， $N_b (= 73)$ ， $N_g (= 72)$  を男子と女子のサンプルサイズ， $\bar{X}_b$  と  $\bar{X}_g$  をそれぞれの標本平均ベクトル， $S$  をプールされた標本共分散行列 ( $df = N_b + N_g - 2$ ) とするとき，(4.1) と (4.3) を導く統計量の定義は，それぞれ

$$\text{MANOVA: } F = \frac{N_b N_g (N_b + N_g - p - 1)}{p(N_b + N_g)(N_b + N_g - 2)} (\bar{X}_b - \bar{X}_g)' S^{-1} (\bar{X}_b - \bar{X}_g) \quad (4.4)$$

$$\text{SEM: } X^2 = \frac{N_b N_g}{N_b + N_g} (\bar{X}_b - \bar{X}_g)' S^{-1} (\bar{X}_b - \bar{X}_g) \quad (4.5)$$

である<sup>29</sup>．MANOVA では，検定統計量 (4.4) が正確に自由度  $(p, N_b + N_g - p - 1)$  の  $F$ -分布に従うことを利用して検定を行う．

(4.4) と (4.5) とから，2 つの検定統計量の間には

$$X^2 = \frac{N_b + N_g - 2}{N_b + N_g - p - 1} (pF) \quad (4.6)$$

という関係が成り立つことがわかる．また，MANOVA の  $F$ -分布は

$$\frac{\chi_p^2 / p}{\chi_{N_b + N_g - p - 1}^2 / (N_b + N_g - p - 1)}$$

と書ける．以上の二点から

$$X^2 = \frac{\chi_p^2}{\chi_{N_b + N_g - p - 1}^2 / (N_b + N_g - 2)} \quad (4.7)$$

を導くことができる．このことから， $X^2$  を自由度  $p$  のカイ 2 乗分布と比較して  $p$  値を求めることは，(4.7) の分母を 1 で置き換えることと同等であり，これは  $F$ -分布の分母の自由度が  $\infty$  であることと同じと考えてよい<sup>30</sup>． $F$ -分布表を眺めてみると，分母の自由度の増加と共に上側 5%点や 1%点が小さくなっていくことがわかる．すなわち， $F$ -分布において分母の自由度を  $\infty$  とすると， $p$  値が小さくなり検出力が上がったようにみえてしまう．SEM での  $p$  値が MANOVA のそれと比べてやや小さく出力される理由はここにある．

SEM は多種多様なモデルが実現可能なため，それらを統一的に扱うためにはこのような近似の理論 (漸近理論) を用いざるを得ない．一般に，MANOVA で分析できる程度の比

<sup>29</sup>たとえば，Anderson (1984, Chapter 5) を参照されたい．SEM の統計量はソフトウェアにより自由度の扱いが微妙に異なるので，カイ 2 乗値の表現が (4.5) に一致しないことがある．

<sup>30</sup>観測変数ベクトルの共分散行列を既知とすることとも同等になる．

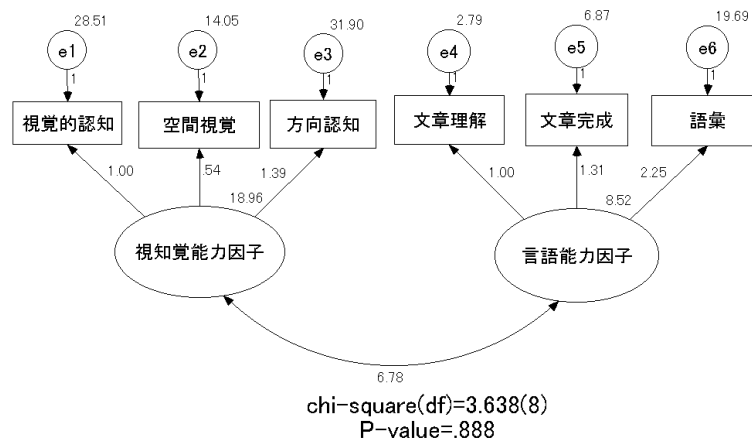


図 10: 検証的因子分析モデル

較的単純なモデルに対しては正確な分布を計算することができるが、SEMのソフトウェアは、正確な分布が求まるかどうかを判定し、もし可能ならばそれを出力するということはしない。

逆に、(4.6)を用いてSEMのカイ2乗統計量からF-値を算出しF-検定を行うこともできる。このときp値は正確になりMANOVAの検定結果と一致する。

本節最後になったが二値変数について補足しておく。男女を表す変数のような二値変数を、基本的に多変量正規性を前提としたSEMに組みこんでもよいのかという疑問がわくかもしれない。二値変数が独立変数である限りこの設定は許されることが知られている(e.g., Kano, 2000)。二値変数を与えた下での条件付統計的推測を行なうことで、男女間で平均の異なるモデルを当てはめていることになる。独立二値変数からのパス係数は従属変数への平均効果を表す。ただし、前提として観測変数の共分散行列が男女間で等しいことが仮定されている。この仮定は通常MANOVAと同一である。

## 4.2 拡張と比較

前節で紹介したのは、通常ANOVA・MANOVAをSEMのフレームワークの中で実行する方法であった。これらは次のように拡張することができる。一つ目は、SEMを用いると観測変数間の共分散行列にSEMで実現できる任意の共分散構造が設定可能なことである。一方、MANOVAには飽和モデルが仮定されている。この心理テストにCFAを想定すると、分析結果は図10となり、「視知覚能力因子」と「言語能力因子」を抽出することができる。共分散行列にこの因子モデルを仮定した下で、男女の平均差を検定するモデルは図11となる。 $m_i$ を自由に推定するモデルのカイ2乗値と(4.2)の下でのモデルのカイ2乗値との差は

$$\begin{array}{lll}
m_i \text{ を推定するモデル} & \chi_8^2 & = 3.734 \\
(4.2) \text{ の下でのモデル} & \chi_{14}^2 & = 16.262 \\
\text{MANOVA(difference test)} & \chi_6^2 & = 12.528
\end{array}$$

(4.8)

となる．difference test の  $p$  値は 0.051 である．

以上の分析結果をまとめると表 6 となる． $p$  値に大きな違いはなく分析方法間に優劣はなさそうである．

表 6:  $p$  値の比較

従来法 (MANOVA)	SEM (共分散構造：飽和モデル)	SEM (共分散構造：因子モデル)
0.056	0.049	0.051

従来の MANOVA のもう一つの拡張は，男女間で異なった共分散行列を許すことである．これは多母集団の同時分析の枠組みで実行されるがここではふれない．

さて，先ほど 6 つの心理テストを因子分析した．次にしばしば行なわれることは，因子分析の結果に基づいて尺度化を行ない次元縮小を図り，そして，それぞれの尺度に関して ANOVA を実行することであろう．CFA の結果に従い，「視知覚能力因子」と「言語能力因子」ごとに 3 項目ずつを尺度化し ANOVA を行なった．結果は， $F$ -値=0.214,  $p$  値=0.242(視覚的認知)， $F$ -値=2.370,  $p$  値=0.126(言語能力) となり有意水準 5% で有意差が確認できなかった．

二つの尺度を二次元確率ベクトルと考えて同時に MANOVA によって分析することも可能である．このような中途半端な分析は実際にはあまり行なわれなないと思うが，ここでは

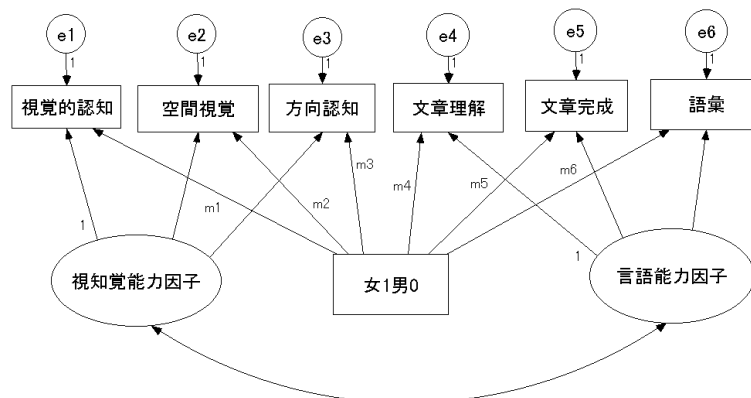


図 11: 検証的因子分析モデルの下での ANOVA・MANOVA: 観測変数の切片が異なるモデル

SEM との比較のため分析してみたところ、 $p$  値は 0.042 となった。

「尺度化 + 分散分析」に対応する SEM は、潜在変数の平均についての検定である。これは、性別を表す変数は（観測変数ではなく）二つの共通因子に効果を及ぼすというモデルによって実行できる（図 12）。すなわち、図 12 において  $sex\_visual=sex\_verbal=0$  の下でのカイ 2 乗値とこの等式制約をおかないモデルのカイ 2 乗値との差をとることになる。分析結果は

$$\chi_2^2 = 8.323, p \text{ 値} = 0.016 \quad (4.9)$$

となる。

尺度化のように、各因子ごとに性別の効果を検討することもできる（図 13）。検定結果は、 $\chi_1^2 = 1.722$   $p$  値=0.189(視知覚能力)、 $\chi_1^2 = 3.240$   $p$  値=0.072(言語能力)となる。

以上の結果をまとめたものが表 7 である。従来方法である ANOVA・MANOVA と SEM である因子平均の推測の結果とを比較すると、SEM の方が一様に  $p$  値が小さくなっている。しかし、この違いは、前節で指摘したように、SEM は漸近理論による近似を用いていることから生じたものと考えられる。

「尺度化 + 相関分析・回帰分析」では、相関係数や偏回帰係数が真の値と比して希薄化（縮小）されることは既に指摘してきたとおりである。では、「尺度化 + 分散分析」ではどのような問題が生じるのだろうか。結論を申し上げますと、希薄化の問題は生じないのである。その理由は、原因変数が固定変数であり誤差を含まないこと、共分散の違いではなく平均の違いをみていることである。

では、「SEM は誤差を分離してより精度の高い分析ができる」という観点から、検出力について検討してみよう。Kano(2000) は、1 因子モデルで構造パラメータが既知という限られた状況ではあるが、MANOVA、SEM による因子平均モデル、そして、回帰タイプの因子得点（重みつき合計）と尺度得点（単純合計）による分散分析の 4 つの方法における検出力を理論的に比較した。結論として、因子平均モデルと因子得点分散分析は同等のパフォーマ

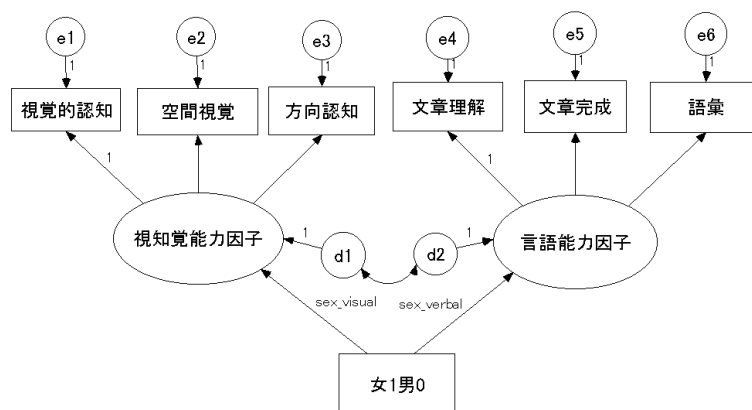


図 12: 検証的因子分析モデルの下での ANOVA・MANOVA: 因子平均が異なるモデル

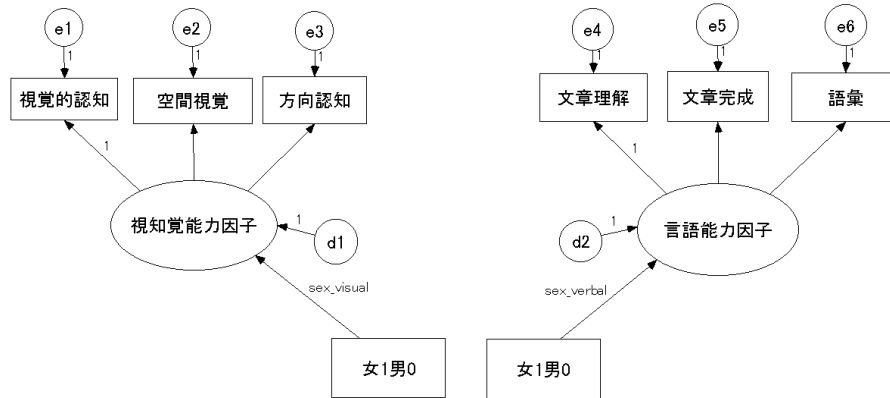


図 13: 因子ごとの効果

表 7:  $p$  値の比較

尺度または因子	従来の方法 ANOVA・MANOVA	構造方程式モデル 因子平均の推測 (図 13・図 12)
視覚能力	0.242	0.189
言語能力	0.126	0.072
2つの尺度または 因子の同時分析	0.042	0.016



ンスを示し他の2つより良いということが示された。MANOVAが劣る理由は平均に構造を設定しないことによる非効率性である。尺度得点が劣る理由は重みがそろった「単純」合計だからであり、重みを適切に選ぶと因子得点となりベストなパフォーマンスを示す。もし、 $\lambda_i/\psi_i$ が変数によらず一定であるならば<sup>31</sup>、MANOVA以外の3つの検定の検出力は一致する。検出力が一致することと信頼性の大きさは関係がない。このように、分散分析においては「誤差を分離する」という意味でのSEMの価値は見出せない。

重みつき合計得点である因子得点は単純合計である尺度得点よりも検出力が高いということを指摘したが、このことは直ちに因子得点の利用が望ましいということを示唆しない。というのは、分散分析での基本仮定であるデータの独立性が崩れているからである。1因子モデルにおいて因子負荷ベクトルを $\lambda$ 、因子分散を $\phi$ 、独自分散行列を $\Psi$ とすると、因子得点で用いられる重みは $(\phi\lambda'\Psi^{-1}\lambda + 1)^{-1}\lambda'\Psi^{-1}$ であり構造パラメータの関数である。通常、構造パラメータは未知でありデータから推定しなければならない。このとき因子得点は $(\hat{\phi}\hat{\lambda}'\hat{\Psi}^{-1}\hat{\lambda} + 1)^{-1}\hat{\lambda}'\hat{\Psi}^{-1}X_i$ となっており、各観測変数 $X_i$ が独立であっても因子得点は独立に分布しないのである。独立性の崩れがF-統計量の分布にどのような影響を及ぼすかが未知である限り、単純合計である尺度得点を用いる方が無難ともいえる。一方、SEMによる因子平均モデルでは4.1節で指摘したように検定は近似的なものに留まる。したがって、これらの3つの方法論の間の優劣は決しがたいのである。

### 4.3 SEMの価値：平均効果の分解と因子分析の基礎仮定

性別変数のような二値変数が独立変数としてモデリングされる時、二値変数からのパス係数は従属変数への平均効果を表すことは既に述べた。特に、二値を0,1でコーディングするとき、パス係数は両母集団における平均効果の差を表す。

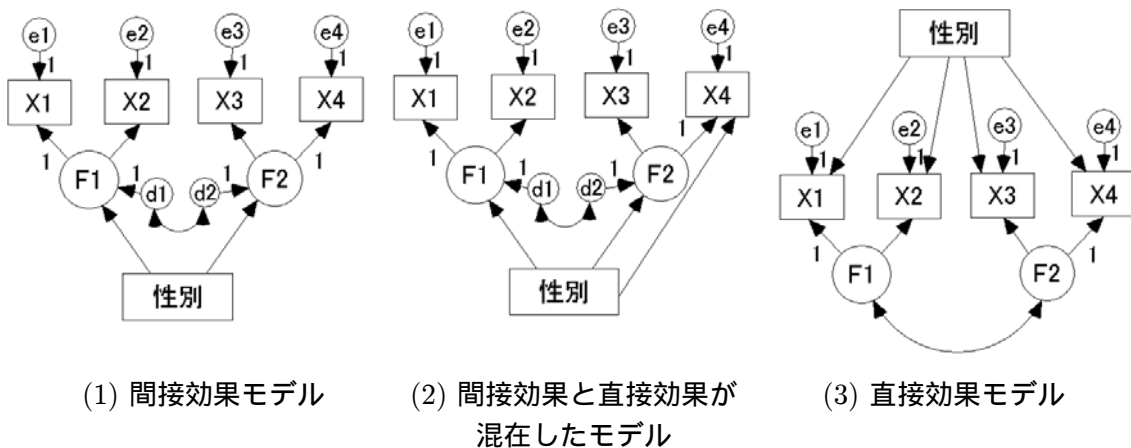


図 14: 種々の平均構造モデル

さて、前節で紹介したモデルは、図 14 で考えれば(1)と(3)であった。すなわち、男女

<sup>31</sup>この条件は、観測変数の分散が標準化してある場合は、 $\lambda_i$ が一定、それゆえ各観測項目の信頼性が一定という条件になる。

で因子の平均 ( $E(F_1), E(F_2)$ ) に違いにあり、それが観測変数  $X_1$ - $X_4$  の平均差に反映するというモデル (1) と、因子の平均は 0 であるが観測変数の切片 (平均) が男女で異なるとするモデル (3) である。別の見方をすると、モデル (1) では性別変数の観測変数への影響は因子を経由する間接効果のみが存在するのに対して、モデル (3) では性別変数の直接効果のみが存在する。

これらのモデルの中間的な存在として図 14(2) がある。このモデルでは、性別変数からの直接効果と間接効果の両者が並存する (e.g., Muthén, 1989)。また、それらの和として総合効果も定義できる。このように、観測変数の背後に潜在変数が想定できるならば、平均効果に関する潜在変数の媒介の程度が評価できる。このような解析は平均効果の分解といってよいだろう。MANOVA にはこのような意味での潜在変数という概念が無いため、平均効果の分解は実行できない。なお、平均の推定に用いられるパラメータ数は、モデル (1), (2), (3) の順に 2, 3, 4 となり、この点でも、モデル (2) は中間的な存在といえる<sup>32</sup>。

さて、因子分析においては、ある母集団から採られた独立同一分布に従うと見なせる標本に基づいて分析する。因子分析を行い尺度を構成し、尺度得点において、たとえば性差があるという結論を導いたとしよう。これは、社会科学で行われる典型的な分析の流れであろう。しかし、「性差がある」という結論と「独立同一分布に従うと見なせる標本」という仮定は矛盾している。この矛盾が分析全体に何らかの影響を与えないだろうか。

男性と女性の母平均ベクトルを  $\mu_b, \mu_g$ 、共分散行列は共通で  $\Sigma$  とするとき、男女同サイズからなるサンプルに基づく標本共分散行列  $S$  の期待値は

$$E(S) = \Sigma + \frac{1}{2}(\mu_b - \mu_g)(\mu_b - \mu_g)'$$

となり、 $E(S)$  に平均の違いが現れる。各母集団において同一の因子分析モデル  $\Sigma = \Lambda\Phi\Lambda' + \Psi$  が適合しているときでも、合併したサンプルでは、上記の右辺第二項の影響で種々の問題が生じる。図 14 で考えれば、(2) と (3) のモデルの場合は、因子分析モデルが適合しないとか、観測変数のグルーピング (因子の指標変数) が変化するということが起こり得る。

一方、(1) の場合は、因子分析モデルは適合するのだが推定値にバイアスが生じる。具体的に計算してみよう。(1) の場合は、 $\mu_b = m + \Lambda c_b$ 、 $\mu_g = m + \Lambda c_g$  と表すことができる。ここで、 $m$  は観測変数の切片で男女に共通、 $c_b$  と  $c_g$  は男女の因子平均である。このとき、 $E(S)$  は次のようになる。

$$E(S) = \Lambda\left\{\Phi + \frac{1}{2}(c_b - c_g)(c_b - c_g)'\right\}\Lambda' + \Psi$$

つまり、男女で因子平均が違えば、因子間の共分散行列が影響を受け、その結果、因子間相関が変化し、そして因子負荷量も影響を受ける。また、男女のサンプルサイズが変われば  $E[S]$  が変化する。したがって、因子間相関や信頼性係数が信用できないのである。

このような問題を避けるには、男女で平均を調整してから因子分析を行うしかない。いくつか属性変数があるとその調整は面倒であるし、どの属性変数で調整するかで因子分析結果が変化し尺度が変わるかもしれない。堂堂巡りを繰り返す可能性もないとはいえない。

<sup>32</sup>性別変数から出ているパスの数である。因子負荷量や誤差分散は共分散行列から定められるため平均構造の推定に関するパラメータには数えない。

実は、図 10 の分析は誤っている可能性がある。平均における性差は微妙であったが、もし性差があると判断するのならば、図 10 の分析はまずいのである。平均が異なるということ的前提で分析する、すなわち、図 11 か、または図 12 のモデルで分析した結果を報告しなければならない。信頼性係数が必要ならば、これらのモデルに基づいて計算することになる。

図 14 のようなモデリングはグループ間による平均の調整を自動的に行いつつ、因子分析と平均における男女差の分析を同時に実行している。また、平均の違いに直接の興味がないときでも、図 14 のようなモデリングを行うことによって、グループによる平均差を調整してから、共分散構造の分析を行うことができる。

SEM は分散分析を得意とはしていない。検出力の観点でも誤差を分離する効果が見えず、「尺度化 + 分散分析」に勝るとは言いえない。ANOVA や MANOVA に関連する分析における SEM の価値は、SEM の平均構造モデルを利用することで、平均の効果の分解が可能になること、属性変数に起因する平均の差異を調整しつつ「正確な」共分散行列の分析が簡単に実行できることである。

本節では議論しなかったが、これらのメリットに加えて、因子分析と分散分析の統合モデルが注目されている (豊田 2000; 豊田&真柳 2001)。今後の発展を待ちたい。

## 5 パス解析と SEM

パス解析は、潜在変数を導入せず観測変数間の因果関係を解明することを目的とする。パス解析の解法には、従属変数ごとに重回帰分析を行ないパス係数を求める従来の方法と、全因果関係を同時に分析する SEM による方法とがある。しかし、SEM が手軽に実行できるようになった昨今では、従来法である重回帰分析の繰返しを使う理由の一つもない。常に、全因果関係を同時に分析する SEM によってパス解析を実行すべきである。理由は以下のとおり<sup>33</sup>。

1. SEM では、モデルの適合度の吟味が可能である。

科学的な研究としてモデルの検証可能性は必須であろう。Popper の反証主義を持ち出すまでもなく、パス解析は因果関係の仮説を立てて因果の大きさを評価するから、仮説の妥当性を必ずチェックすべきである。

2. SEM では、モデルの適合度や LM 検定・Wald 検定を参考にすることで、因果モデルを容易に改善することができる。

多くの変数間の因果関係は一朝一夕で解明できるものではない。先行研究の知見などを基にして因果モデルを作ってもデータに合わないことが頻繁に起こる。その際、統計的な観点と実質科学的な観点の両面をふまえながら因果モデルを改善していくことが重要である。その作業が正に因果分析である。従来法はモデルを改善していくための指針がなく、どのようにモデル改善を図るべきかの方略にとぼしい。

3. SEM では、効果の分解が容易である。

パス解析においては変数から変数へいくつものパスが引かれれているという状況が多い。その中で直接効果と間接効果を分離し、また、総合効果を求めることは現象理

---

<sup>33</sup> 豊田 (2000, 第 8 章) も参照されたい。

解に大いに役立つ。従来法でも手計算は可能であるが面倒である。SEM プログラムであればコマンド一つでこれらの効果を出力できる。

- SEM では、誤差共分散（偏相関）を入れたモデリングが可能であり、また、誤差共分散が必要かどうかについての情報も容易に得られる。

観測変数間の因果モデルは直接効果を念頭においてパスを引く。したがって、観測変数間の相関は十分に説明されないかもしれない。つまり偏相関が残る可能性が、潜在変数を入れた分析よりも高くなる。偏相関は誤差共分散によってモデル化される。従来法では誤差共分散を扱えないばかりか、誤差共分散が必要かどうかもわからない。

- SEM では、双方向因果モデルで容易に推定を行うことができる。
- SEM では、観測変数が尺度であるとき、もし尺度の信頼性  $\rho$  が既知であれば、 $\rho$  を考慮した分析が可能である。
- 重回帰分析では説明変数間の分散・共分散は常に飽和モデルであり、構造を設定できない。したがって、従来法では説明変数間に設定された因果構造を分析に活かすことができない。

飽和モデルは推定すべきパラメータが多いためモデル全体の推定効率が落ちる。因果モデルが適切であれば、従来法ではパス係数が有意ではない場合でも SEM で分析すれば有意性が確認できることがある。

## 6 結論

ここで紹介した議論をまとめると表 8 となる。

因子分析の枠組みで SEM が活躍するのは、EFA が実行できないときや、EFA が使っても統計的証拠が必要なときである。また、SEM は「尺度化 + 相関分析・回帰分析」のように多段階の分析において活躍する。重回帰分析の繰返しで解くパス解析も多段階である。SEM では、このような多段階の分析を一つのモデルで表現し、一回の run (submit) で分析することができる。一つのモデルで表現することから、希薄化が生じないなどさまざまなメリットが生まれる。「尺度化 + 分散分析」では希薄化が生じないため、希薄化の修正という意味での SEM のメリットはない。しかし、平均効果の分解や平均を調整した下での因子分析が簡単に実行できるという別のメリットがある。

本論文では、回帰分析や（多変量）分散分析とは、一般線型モデル (GLM) で実行できる分析を念頭においた。ここではふれなかったが、SEM とは別の選択として（線型）混合モデル (mixed model; Littell et al. (1996)) がある。混合モデルはランダム効果要因を「正式」に扱うことができるという点で、一般線型モデルの拡張と考えることができる。本論文に関連するモデルとしては、誤差分散がセルごとに異なるモデルや、集団間で異なった共分散行列<sup>34</sup>を仮定した MANOVA を実行することができる。混合モデルでは分析の主眼が平均にあり、SEM では共分散にあるという違いがあるもののよく似た下位モデル群を提供する。混合モデルと SEM の類似点や相違点は別の機会に議論したい。

<sup>34</sup>ソフトウェアが提供するいくつかのパターンの中から選択する。

表 8: 比較のまとめ

分析方法	分析状況	従来法	構造方程式モデル
因子分析	因子についての仮説がない 統計的証拠 (適合度, 因子負荷量の零仮説) パラメータに関する仮説検証 (等値性など) 誤差共分散 因子が多いモデル 多母集団の統計的同時分析 因子平均の分析 拡張モデル	探索的因子分析  × × × × × ×	検証的因子分析  ×
構成概念の 因果分析	信頼性が高い 信頼性が低い 単純性を崩す項目	尺度化 + 相関・回帰分析  ×	潜在変数の相関・回帰分析
構成概念の 実験的分析	3水準以上 交互作用 多重比較 平均効果の分解 平均の調整 異なる分散や共分散行列	尺度化 + 分散分析  × × <sup>†</sup>	潜在変数の分散分析  ×
パス解析	モデルの吟味 モデル改善のしやすさ 効果の分解 誤差相関 双方向因果モデル 信頼性を考慮した分析	重回帰分析の繰返し × × × ×	SEM による同時分析

○ : 十分対応できる . ◐ : 部分的に , もしくは , 状況によっては対応できる . × : 対応できない .  
<sup>†</sup> 混合モデルを用いると実行可能である .

## 謝辞

本特集を企画して頂いた豊田秀樹氏 (早稲田大学) に感謝の意を表します . また , 本原稿に対して適切なコメントと提言を下された査読者にも感謝をいたします .

## 参考文献

- Anderson, T. W. (1984). *An introduction to multivariate statistical analysis, 2nd ed.* Wiley: New York.
- Anderson, T. W. & Rubin, H. (1956). Statistical inference in factor analysis. *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, **5**, 111-150. University of California Press: Berkeley.
- Arbuckle, J. L. & Wothke, W. (1999). *AMOS 4.0 user's guide*. SmallWaters Corporation: Chicago.
- Bentler, P. M., Poon, W. Y. & Lee, S. Y. (1988). Generalized multimode latent variable models: Implementation by standard programs. *Computational Statistics and Data Analysis*, **7**, 107-118.
- Bollen, K. A. (1989). *Structural equations with latent variables*. Wiley: New York.
- Chan, W. (2003). Analyzing ipsative data in psychological research. *Behaviormetrika* **30**, 99-121.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, **16**, 297-334.
- 南風原朝和 (1992). 知能検査の因子構造をさぐる — 因子分析 『心理・教育のための多変量解析法：事例編』 (渡部 洋 編著) . pp.71-81. 福村出版 .
- Jöreskog, K. G. (1969). A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika*, **34**, 183-202.
- Jöreskog, K. G. (1973). A general method for estimating a linear structural equation system. In *Structural equation models in the social sciences*, (A. S. Goldberger & Duncan, Eds.) pp.85-112. Seminar Press: New York.
- 狩野 裕 (1997). グラフィカル多変量解析 — 目で見る共分散構造分析 . 現代数学社 (狩野 裕・三浦麻子 (2002) 増補版) .
- Kano, Y. (1997). Exploratory factor analysis with a common factor with two indicators. *Behaviormetrika*, **24**, 179-197.
- 狩野 裕 (1998a). 不適解の原因と処理：探索的因子分析 . 大阪大学人間科学部紀要 . **24**, 303-327.
- 狩野 裕 (1998b). タレント好感度データの分析 『共分散構造分析— 構造方程式モデリング — [事例編]』 (豊田秀樹編) . pp. 9-21. 北大路書房 .
- Kano, Y. (2000). Structural equation modeling for experimental data. In *Structural equation models: Present and future* [A Festschrift in honor of Karl Jöreskog] (Eds., Bob Cudek, Stephen du Toit, & Dag Sörbom), pp.381-402. SSI: Chicago.
- Kano, Y. & Azuma, Y. (2003). Use of SEM programs to precisely measure scale reliability. In Yanai, H. et al. (Eds), *New developments in psychometrics*. pp.141-148. Springer Verlag: Tokyo.
- 児玉 省・品川不二郎・茂木茂八 (1978). 日本版 WISC-R 知能検査法 . 日本文化科学社 .
- Ledermann, W. (1937). On the rank of the reduced correlational matrix in multiple factor analysis. *Psychometrika*, **2**, 85-93.
- Littell, R. C., Milliken, G. A., Stroup, W. W., & Wolfinger, R. D. (1996). *SAS system for mixed models*. SAS Institute Inc: NC.

- Muthén, B. (1989). Latent variable modeling in heterogeneous populations. *Psychometrika*, **54**, 557-585.
- Muthén, B. (2002). Beyond SEM: General latent variable modeling. *Behaviormetrika*, **29**, 81-117.
- 丘本 正 (1986). 因子分析の基礎. 日科技連.
- Sato, M. (1987). Pragmatic treatment of improper solutions in factor analysis. *Annals of the Institute of Statistical Mathematics*, **39**, 443-455.
- 芝 祐順 (1979). 因子分析法 (第 2 版). 東京大学出版会.
- 清水和秋 (1994). Jöreskog と Sörbom によるコンピュータプログラムと構造方程式モデル. 関西大学社会学部紀要, **25**, 1-41.
- 豊田秀樹 (1992). SAS による共分散構造分析. 東京大学出版会.
- 豊田秀樹 (1998). 共分散構造分析 — 構造方程式モデリング — [入門編]. 朝倉書店.
- 豊田秀樹 (2000). 共分散構造分析 — 構造方程式モデリング — [応用編]. 朝倉書店.
- 豊田秀樹 (2001). 探索的ポジショニング分析 — セマンティック・デファレンシャルデータのための 3 相多変量解析法. 心理学研究, **72**, 213-218
- 豊田秀樹・真柳真誉美 (2001). 繰り返し測定を伴う実験のための因子分析モデル — アイスクリームに関する官能評価. 行動計量学, **28**, 1-7.
- Van Driel, O. P. (1978). On various causes of improper solutions in maximum likelihood factor analysis. *Psychometrika*, **43**, 225-243.
- 柳井晴夫・繁榎算男・前川眞一・市川雅教 (1990). 因子分析: その理論と応用. 朝倉書店.
- 吉田光雄・狩野裕・原田章 (in press). SAS による統計解析. 科学技術出版.