

共分散構造分析 (SEM¹) は、パス解析、因子分析、分散分析の すべてにとって代わるのか？

狩野 裕 (大阪大学人間科学部)

目次

オープニング	2
1 データと探索的因子分析・検証的因子分析	2
2 探索的因子分析とSEM	3
2.1 EFA, CFA ともに可能な場合	4
2.2 EFA では分析できないが CFA では分析できる場合	5
2.2.1 2変数にしか影響しない因子がある場合	6
2.2.2 因子が多すぎる場合 (Ledermann の限界)	6
3 「尺度化+相関分析・回帰分析」とSEM	8
3.1 相関分析	8
3.2 回帰分析	9
4 「尺度化+分散分析」とSEM	11
4.1 ANOVA・MANOVA のSEMでの実行方法	11
4.2 拡張と比較	13
5 パス解析とSEM	16
6 さいごに	17

本タイトルは「春セミ」実行委員会から与えられたお題である。なかなか魅力的で気に入っている。この問いに対してだれも「YES」とは言わないと思うが、少しトーンと落して、「パス解析、因子分析、分散分析があるのになぜ共分散構造分析が必要なのか」という疑問は多くの研究者・実務家もっているだろう。若い院生なら、「最近、共分散構造分析が流行っているらしい」という噂だけで、あまり抵抗なくこの新しい分析方法を勉強するであろうが、従来の分析方法で十分に実績を積んだ中堅は、既存の方法論にはない「ミソ」がなければ、つまり、既存の方法論では分析できないが、共分散構造分析では可能になる、そういう状況を目の当たりにしないと、新たに勉強しようという気にはならないだろう。逆に、アンチ共分散構造分析になるかもしれない。

本コースでは、このような目的意識をもち、既存の分析方法と共分散構造分析の relative advantage について話題提供する。以下の内容を主に議論する：

探索的因子分析	vs	検証的因子分析
探索的因子分析+尺度化+回帰分析・相関分析	vs	潜在変数の回帰分析・相関分析
探索的因子分析+尺度化+分散分析	vs	潜在変数の分散分析
パス解析（重回帰分析の繰り返し）	vs	共分散構造分析でのパス解析

パス図やモデルファイルの作成方法など初級者向けの講義はしないが、従来の方法との比較を通して初級者にも十分理解できるよう配慮する。

中級者にとっては「ミソ」が理解できる、また、初級者には共分散構造分析を勉強してみようという意欲がわく、そのようなコースにしたい。

¹呼称 SEM については 6 節を参照のこと。

オープニング

共分散構造分析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。これは正しい事実であるが数学的な議論である。普通の回帰分析や分散分析は、共分散構造分析でも実行できるが、やはり既存のソフトウェアが使いやすい。共分散構造分析では残差分析、多重共線性のチェックや多重比較ができないという問題もある。探索的因子分析 (EFA) も共分散構造分析では普通実行できない。共分散構造分析のソフトウェアを購入したからといって、一般的な統計ソフトウェアが不必要ということではない。

共分散構造分析が活躍するのは、社会科学でしばしば行なわれる「尺度化」というステップをふむ統計分析である。また、因子分析に限っても“有意な”拡張がなされている。

本コースの目的は、共分散構造分析を既存の分析方法と比較することで relative advantage をはつきりさせ、統計ユーザーが分析方法を適切に選択できるようガイダンスすることである。

二時間の話題提供、そして、この分野の著名な先生方の討論をゆっくりとお楽しみ下さい。

1 データと探索的因子分析・検証的因子分析

図 1 は、6 科目（代数・幾何・解析・英語・国語・古文）のテストデータを、(相関行列から) 探索的因子分析 (EFA; Exploratory Factor Analysis) と検証的因子分析 (CFA; Confirmatory Factor Analysis) を実行するための SAS プログラムである。因子数は 2 である。探索的因子分析では最尤法を採用しプロマックス斜交回転解を求めている²。検証的因子分析で採用したモデルは図 2 に示されている³。

分析結果は表 1 に示してある。因子 F1 は数学的能力、因子 F2 は言語的能力と解釈できる。数学的能力因子は英語にも影響している。因子間相関は両者とも約 0.64 である。適合度 (有意確率) や AIC をみると、検証的モデルの方がややよいが、解釈に違いはなさそうである。ということであれば、共分散構造分析 (検証的因子分析) を実行する意義はないようにみえる。

この疑問が本コースの原点である。

探索的因子分析の目的は大きく分けて、(i) 因子分析の結果自身に興味がある場合、(ii) 因子分析後、尺度化などを行ない次元縮小し相関分析や回帰分析に持ちこむ場合、の二通りがある。ここでは、(i) を 2 節で、(ii) を 3 節以降で議論する。

表 1: 探索的・検証的分析結果

	探索的因子分析 (プロマックス斜交回転)		検証的因子分析 (図 2 でのモデル)	
	F1	F2	F1	F2
代数	.634	.077	.698	0
幾何	.672	-.022	.646	0
解析	.732	.008	.733	0
英語	.390	.543	.413	.523
国語	-.059	.819	0	.769
古文	.023	.634	0	.658
因子 相関	1	.641	1	.642
	.641	1	.642	1
$\chi^2(df)$	3.724(4)		6.307(7)	
有意確率	0.4441		0.5044	
AIC	-4.276		-7.693	

²1 節～3 節の多くの部分は、近刊「吉田・狩野・原田 (2000). SAS による統計解析. 科学技術出版」による。

³パス図は AMOS によって描いている。

```

DATA kamoku(TYPE=CORR);
  INPUT _TYPE_ $ _NAME_ $ X1-X6;
  LABEL X1='代数' X2='幾何' X3='解析'
        X4='英語' X5='国語' X6='古文';
  CARDS;
#      .   250   250   250   250   250   250
CORR X1 1.000 .       .       .       .       .
CORR X2 0.412 1.000 .       .       .       .
CORR X3 0.521 0.495 1.000 .       .       .
CORR X4 0.538 0.499 0.525 1.000 .       .
CORR X5 0.334 0.293 0.364 0.607 1.000 .
CORR X6 0.346 0.248 0.323 0.517 0.506 1.000
;
TITLE '*** 探索的因子分析 n=2 ***';
PROC FACTOR DATA=kamoku METHOD=ml NFACTORS=2 PRIORS=smc
          ROTATE=promax;
RUN;
;
TITLE '*** 検証的因子分析 ***';
PROC CALIS DATA=kamoku METHOD=ml ALL;
  LINEQS
  X1 = L_11 F1          + E1,
  X2 = L_21 F1          + E2,
  X3 = L_31 F1          + E3,
  X4 = L_41 F1 L_42 F2 + E4,
  X5 =          L_52 F2 + E5,
  X6 =          L_62 F2 + E6;
  STD
  E1 - E6 = DEL1 - DEL6,
  F1 - F2 =2*1.00;
  COV
  F1 F2 = PHI12;
RUN;

```

図 1: 探索的&検証的因子分析の SAS プログラム

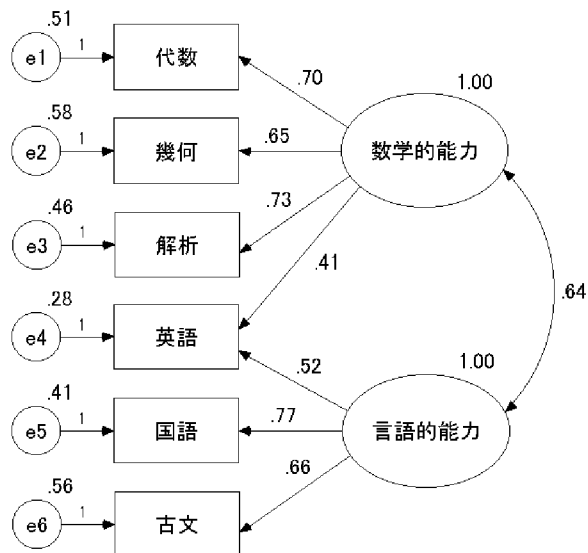


図 2: 検証的因子分析結果

2 探索的因子分析とSEM

探索的因子分析 (EFA) と検証的因子分析 (CFA) を比較する際、CFA は可能だが EFA は実行できないという状況がある。前節で紹介したデータのように両分析が可能である場合、どちらに軍配があがるのか

という問題がある。前者を 2.2 節で議論し、後者を 2.1 節で紹介する。

2.1 EFA, CFA とともに可能な場合

探索的因子分析 (EFA) は観測変数の背後に潜む構造を探し、潜在因子を発見することにその主目的がある。一方、検証的因子分析 (CFA) は、観測変数の背後にある因子についての仮説があり、その仮説を検証することに目的がある。探索的分析は帰納的であり検証的分析は演繹的と形容できよう。このような哲学的な議論は方法論者には好まれるが、統計学のユーザーにとってみれば、「じゃ、実際どのように違うのか」という疑問を解消する方が重要であろう。

まず、因子についての「良質な仮説」がなければ検証的分析は実行できない。探索的分析にはそのような仮説は必須ではない。この意味で検証的分析の方が仮定が強く制限的であり、探索的分析の方が適用範囲が広い。したがって、結果が同じであるならば、仮定が弱い探索的因子分析に軍配が上がる。

「良質な仮説」とは、小修正によって適合の良い最終モデルに到達できるような仮説をいう。実際、図 2 のモデルは初期モデルではなく、当初は、「数学的能力」から「英語」へのパスを引かないモデルであった (図 8 モデル 3)。しかし、その当初モデルはデータへの適合が悪かったので、図 2 のモデルへ小修正した。共分散構造分析 (SEM) は基本的に検証的分析なので、モデルの大々的な探索・構築には不向きである。ここでのように、小修正で最終モデルに到達できるような「良質な仮説」からスタートできることが要求される。

ここまで EFA の利点を取り上げてきたが、CFA には統計的証拠という大きな利点がある。SEM のプログラムには、LM 検定・ワルド検定・修正指標というモデルを修正するためのオプションが備えられている。これらを用いることによって採用されたモデルの部分評価ができる⁴。

図 3 をみてみよう。これは、図 2 の分析における SAS Calis による LM 検定とワルド検定の結果であ

Lagrange Multiplier or Wald Index				

Probability		Approx Change of Value		

	F1		F2	
X1	126.651	[L_11]	0.644 0.422	0.093
X2	105.083	[L_21]	0.238 0.625	-0.055
X3	142.771	[L_31]	0.108 0.743	-0.039
X4	24.920	[L_41]	38.561	[L_42]
X5	0.475 0.491	-0.109	144.781	[L_52]
X6	0.475 0.491	0.093	106.690	[L_62]

図 3: 修正後のモデルにおける LM 検定とワルド検定

⁴AMOS の修正指標は簡易版であり、検定には使わない方がよい。LM 検定・ワルド検定については EQS が使いやすい。

る。具体的には、因子 (F) から観測変数 (X) へのパス係数がゼロであるかどうかの検定結果が示されている。検定統計量としてカイ 2 乗値 (t-値の 2 乗) が報告されており、この値が $\chi^2_1(0.05) = 3.841$ 以上であれば 5% 有意となり、そのパスを引く価値が統計的に認められる。一方、非有意の場合は、パスを引かない理由が統計的に確認できることになる。これらの検定を用いて、分析者が作成したモデルにおいて、引いたパスの有意性と引かなかったパスの非有意性を確認したい。本例題では、引いたすべてのパスが有意であり、引かなかったすべてのパスが非有意であることが分かる。なお、引いたパスについての検定をワルド検定、引かなかったパスについての検定を LM 検定といい、これらの検定は、モデル修正に使うだけでなく想定したモデルの局所評価にも利用できる。

一方、探索的分析では、いくら因子負荷量の値が大きくても統計的に有意であるかどうかは不明であり、絶対値が小さくても非有意であるかどうかは分からない。因子負荷量の値が無視しづらいほど大きいのが解釈しにくいという状況は頻出である。そのようなとき、(偶然) 大きな値に推定されたが再現性のない artifact であるか、有意性が確認できるほど重要な結果かという判断は極めて重要である。この点では、そのような判断を与えてくれる CFA の方が EFA より有用であるといえよう。

では、なぜ探索的分析ではワルド検定ができないのであろうか (LM 検定は不要)。1970 年代にそのための理論は完成されているにもかかわらず、著者の知る限り、どの市販汎用統計プログラムにもそのようなオプションはない。その理由は因子回転の問題にあると思われる。つまり、推定値や検定結果が因子回転の方法に依存するのである。例えば、プロマックス回転には `power=*` というオプションがある。ここで * は

表 2: power の値と推定値

	power の値			
	2	3	4	5
因子相関	0.50	0.64	0.70	0.73
λ_{12}	0.14	0.08	0.04	0.02

1 以上の整数であり、SAS のデフォルトは `power=3` である⁵。表 2 には、この値を 2~5 と変化させたときに因子相関と F2 から X1 への因子負荷量 λ_{12} の推定値を示している。斜交回転は種々あるが、たとえプロマックス回転にしぼったとしてもオプション `power=*` の値によって推定値がかなり変化することが分かるだろう。したがって、たとえ有意性検定を行なったとしても、回転の種類によって結果が異なることが予想される。探索的分析の結果を細かく解釈することは難しいのである。

2.2 EFA では分析できないが CFA では分析できる場合

EFA では分析できないが CFA では分析できるというケースは、大きく分けて次の 4 つの場合がある。

- (i) 2 変数にしか影響しない因子がある場合
- (ii) 因子が多すぎる場合
- (iii) 複数の集団の因子分析結果を統計的に比較したい場合 (多母集団の同時分析)
- (iv) 因子平均に興味がある場合

上記に含まれない場合として、三相データの因子分析 (PARAFAC) や非線型因子分析 (含 交互作用がある場合) などもある⁶。今回は、(i) と (ii) を取り上げる。

⁵SPSS ではこのパラメータを κ (カッパ) といいデフォルトは 4 である。

⁶これらを詳しく勉強したい受講生は「豊田秀樹 (近刊). 共分散構造分析 [応用編]. 朝倉書店」を参照のこと。

2.2.1 2変数にしか影響しない因子がある場合

6科目のテストデータで「X4:英語」を除いて探索的因子分析を行なうとどうなるだろうか. SAS データセット中から一部の変数のみを使いたいときは, VAR コマンドを使う:

```
PROC FACTOR DATA=kamoku M=m1 N=2 PRIORS=smc R=p;
VAR X1 X2 X3 X5 X6;
```

このジョブを走らせると「共通性が1を越えました」というエラーが出て止まってしまう:

```
Iter Criterion Ridge Change Communalities
3 0.0000838 0.000 0.08313 0.44219 0.39467 0.62076 0.32762 1.05049

ERROR: Communality greater than 1.0.
```

実際, X6 の共通性が 1.05049 となっている. 理論的には共通性は 0 ~ 1 に収まるはずであるから, このメッセージが出たときは, データかモデルに問題がある. 今回の場合はモデルに問題がある. 「英語」を分析から外したので「言語的能力」が大きな因子負荷量をもつ観測変数は2つしかない. 探索的因子分析は2つしか大きな因子負荷量をもたない因子があるときには適用できないことが分かっている. 一方, 「英語」を落して検証的因子分析を実行すると図4のようになり, 美しく分析できる. 因子相関は 0.64 と推定されており, 観測変数の変更に対して極めて安定していることが分かる.

このように, 2変数にしか影響しない因子がある場合は, 探索的因子分析では分析できないが検証的因子分析は問題なく実行できる.

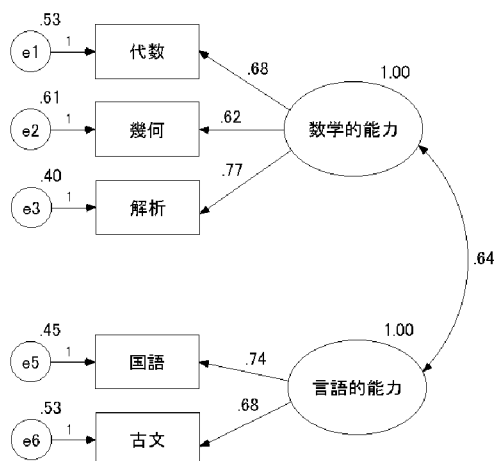


図4: 「X3: 英語」を落とした下での検証的因子分析

2.2.2 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

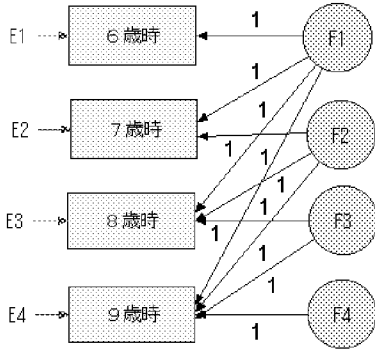
$$\text{因子数} \leq (2p + 1 - \sqrt{8p + 1}) / 2 \quad (2.1)$$

である. これを Ledermann の限界という. いくつかの p について抽出可能な因子数の上限を表3に示してある:

表 3: 抽出可能な因子数の上限

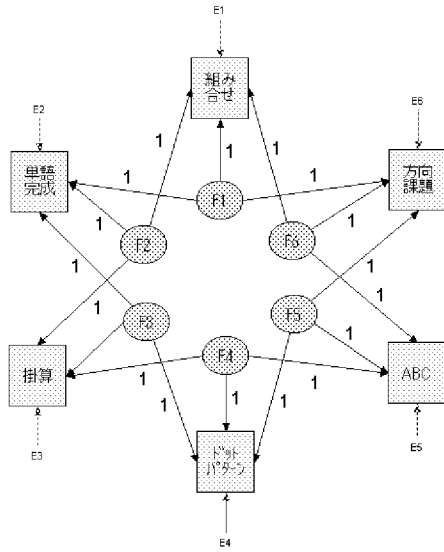
観測変数の数 p	2	3	4	5	6	7	8	9	10	...
因子数の上限	0	1	1	2	3	3	4	5	6	...

Ledermann の限界を超えた因子数に対しても、何らかの事前情報から因子負荷量に関する仮説があれば、検証的方法でもって因子分析することができる。以下に、シンプレックスモデル (図 5)、円環モデル (図 6)、潜在曲線モデル (図 7) の例を紹介している。



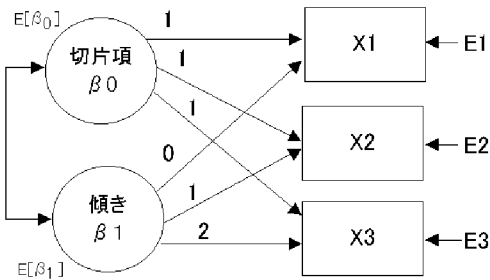
$$\begin{bmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \end{bmatrix}$$

図 5: シンプレックスモデル



$$\begin{bmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \\ X_5 \\ X_6 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \\ f_5 \\ f_6 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \\ e_6 \end{bmatrix}$$

図 6: 円環モデル



$$\begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ e_3 \end{bmatrix}$$

図 7: 潜在曲線モデル

3 「尺度化+相関分析・回帰分析」とSEM

探索的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。ひと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（はずな）ので、そこから因子の命名が行なわれる。観測変数群を、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。因子分析から始まって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易くなる。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ていないような気になる。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。

3.1 相関分析

調査・分析の目的が「数学的能力」と「言語的能力」の相関を求めることとし、6科目のデータを使って尺度化を行ないこれらの間の相関係数を計算してみよう。

尺度化において考慮すべき問題がある。「英語」の扱いである。EFA であれ CFA であれ、分析結果は英語が両因子と係わっていることを示している。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しばしば行なわれる。いわゆる単純構造を目指すのである。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、英語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）変数を抹殺する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。数学的能力を表す尺度を $X_1+X_2+X_3$ とし、言語的能力を表す尺度を X_5+X_6 とし、尺度間相関を求めると以下のようなになる：

$$\text{Cor}(X_1 + X_2 + X_3, X_5 + X_6) = 0.45$$

この値は探索的分析や検証的分析での因子相関 (0.64) よりずいぶん小さい。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。尺度の信頼性とは、尺度 X を真の得点 f 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う。すなわち、

$$\rho = \frac{V(f)}{V(X)} = 1 - \frac{V(e)}{V(X)} \quad (3.1)$$

である。

「相関係数の希薄化」、すなわち、相関係数の絶対値が小さく推定されるというバイアスは、この尺度の信頼性が高くなく誤差が無視できないほど大きいことによって生じる。

このようなバイアス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」と仕込まれる理由はここにある⁷。

このバイアスは、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。検証的因子分析のメリットは、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が構成できず信頼性

⁷ α 係数は信頼性の簡便推定値である。

が低くとも、CFA でモデル化すれば、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。

なお、上記二つの尺度の信頼性は次のようになる：

$$\begin{aligned} \text{数学的能力の尺度 } (X_1 + X_2 + X_3): & \quad 0.73 \\ \text{言語的能力の尺度 } (X_4 + X_5): & \quad 0.67 \end{aligned}$$

希薄化された相関（尺度化して求めたもの）と真の相関（CFA による相関）の間には、信頼性を介して次の関係がある：

$$0.45 \approx 0.64 \times \sqrt{0.73 \times 0.67} \quad (3.2)$$

上記の計算方法を活用すれば、信頼性の低さから希薄化の程度を見積もることができる。

最後に、単純構造を乱す変数の扱いについてふれておきたい。上記の例では、数学的能力と言語的能力の両方に負荷をもつ「英語」の扱いである。尺度化して分析するには、このような変数を外さざるを得なかった。しかし、CFA ではこのような変数を含めたまま分析することができる。図 8 には3つのモデルと

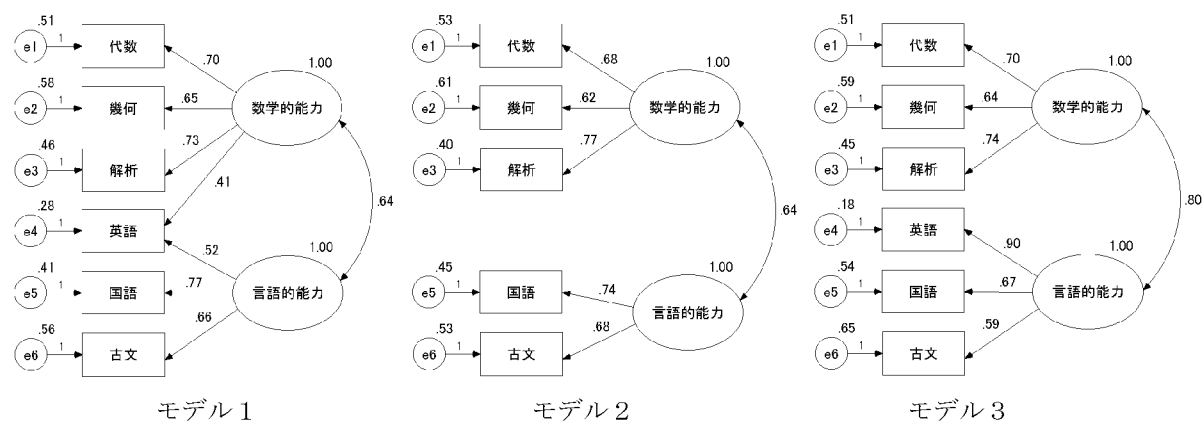


図 8: 単純構造を乱す変数のコントロール

それらに基づく分析結果が示されている。モデル 1 は「英語」を含めたもの、モデル 2 は「英語」を外したもので、どちらで分析しても興味の対象である因子相関は安定して推定できる。このことは既に指摘した通りである。モデル 3 は、数学的能力から英語へのパスを消したモデルである。ここには報告していないが、実は、モデル 3 は適合度が低く棄却されてしまう。因子間相関は 0.8 と異常に高い値に推定されている。数学的能力から英語への直接効果を因子相関で説明しようとした結果、このような不当に高い因子相関になってしまっているのである。モデル 1 は、数学的能力から英語へパスを引くことでその影響をコントロールし、因子間相関を適切に推定することに成功している。

3.2 回帰分析

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、「言語的能力」から「数学的能力」への回帰モデルを考えると、分析結果は図 9 のようになる。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。

もうすこし尤らしい例をあげよう。図 10 のモデルは、「F1 : 宗教」と「F2 : 道徳」の「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⁸。潜在変数 F1 には 3 個の観測変数が、F2 には 4 個の観測変数が用意されている。

⁸ 高名な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。

CFAの結果によると、両国とも、「F1:宗教」からのパス係数が0.5、「F2:道徳」からのパス係数が0.8と推定されているから、国間で差異がなく、道徳の影響がより大きいということになる。ただし、測定モデル間には差がある。道徳の測定モデルにおいてA国の方が因子負荷量が小さい。

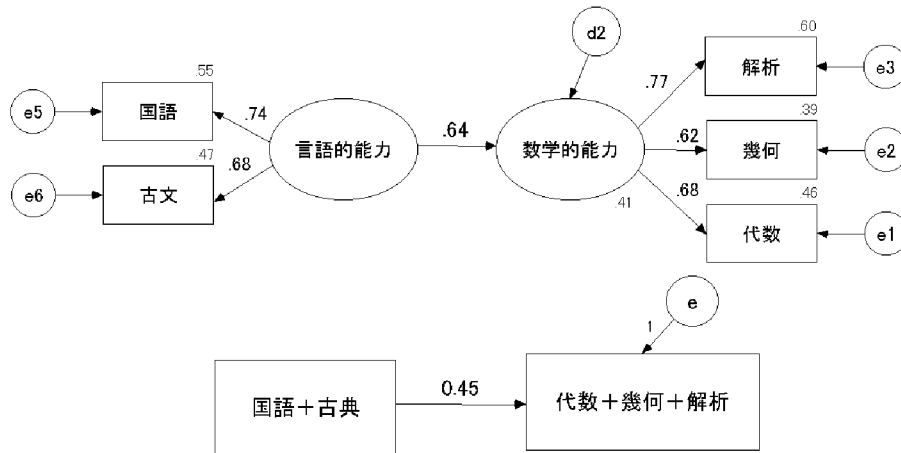


図 9: 共分散構造分析 (上) versus 尺度化+回帰分析 (下) (すべての分散を1に標準化した標準解)

次に、これらを尺度化して尺度間の回帰分析を実行すると、B国のデータはCFAの結果を近似するが、A国は、道徳からのパス係数が大きく下がり0.43となる。道徳の影響は宗教の影響とほぼ同じか幾分小さくなる。このように、尺度化+回帰分析とCFAとでは結論が大きく異なることがある。

原因はA国の道徳尺度にある。 $\alpha = 0.28$ であり、道徳尺度の信頼性が極めて低いのである。このように信頼性の低い尺度を分析に用いることは問題になることがある。一方、CFAは信頼性を低める誤差の影響を取り除いて適切に分析してくれる。なお、因子負荷量が0.3の下位項目を合算した尺度で、0.8の信頼性を確保するには、40の下位項目が必要である⁹。

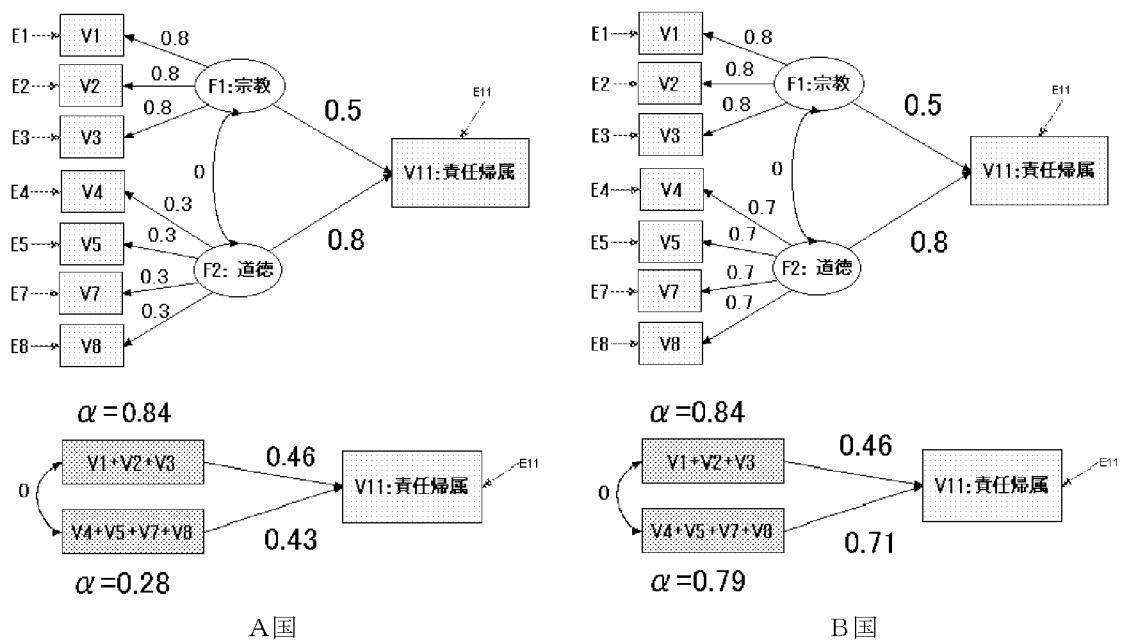


図 10: 共分散構造分析 (上) versus 尺度化+回帰分析 (下)

⁹ 因子負荷量が0.4のときには必要な下位項目は20、0.5のときには12となる。しかしながら、たとえ0.8の信頼性を確保してもパス係数は1程度過小評価される。

共分散構造モデルの大きな特徴は、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、従来の方法でもそれなりに適切な結論を得ることは可能である。しかし、常に、内的整合性の高い下位項目を十分に構成できるとは限らない。また、多標本問題では、母集団間で信頼性が異なるかもしれない。これらを調整して分析するには共分散構造分析が必須である。

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共分散構造モデルで分析することは可能である¹⁰。例えば、共分散構造分析であればセルごとに分散が不均一という場合も扱うことができる。しかし、現在のところ、共分散構造分析は分散分析(ANOVA)を得意とはしていない。特に一変量の分散分析ではSAS PROC GLMなどで簡単に分析できる場合は共分散構造分析を持ち出すまでもない。というのは、共分散構造分析ソフトウェアの仕様が分散分析を念頭においていないためモデルプログラムに相当の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が、交互作用がある場合や3水準以上の要因に対する共分散構造分析は複雑である¹¹。

多変量分散分析(MANOVA)¹²でも同様に、水準数が多いと分析しにくいという問題がある。しかし、共分散構造分析では、観測変数の共分散行列に構造を簡単に設定できるというメリットがある。構造として例えば因子分析モデルを想定することもできる。通常の変量分散分析では、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¹³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。

さて、ここでは、多変量データを尺度化して次元に落としてから一変量分散分析を実行する状況を考える。「尺度化＋分散分析」の alternative である因子の分散分析(共分散構造分析の一つのモデル)との relative advantage を議論したい。

4.1 ANOVA・MANOVA のSEMでの実行方法

ここで分析するデータは図11の心理データである¹⁴。

心理テストの内容は「visual perception (視覚的認知)」「cubes(空間視覚)」「lozenges (方向認知)」「paragraph comprehension (文章理解)」「sentence completion (文章完成)」「word meaning (語彙)」である。分析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。SPSSによってMANOVAを実行すると有意確率は0.056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」のみ有意確率が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を共分散構造分析のフレームワークの中で実行するには、図12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、

$$m_1 = m_2 = m_3 = m_4 = m_5 = m_6 = 0 \quad (4.1)$$

¹⁰「豊田秀樹(近刊). 共分散構造分析 [応用編]. 朝倉書店」にはそれらの実行例が紹介されている。

¹¹デザイン行列をデータとして入力しなければならない。

¹²MANOVAは Multivariate Analysis of variance の短縮形。ANOVAも同様。

¹³飽和モデルという。

¹⁴「AMOS マニュアル」や「狩野(1997) グラフィカル多変量解析. 現代数学社」には同データの多母集団の同時分析による分析が報告されている。

	visperc	cubes	lozenges	goraetap	sentence	wordmean	sex	b	視知覚	言語能力	
1	33.00	22.00	17.00	8.00	17.00	10.00	女1男0	1.00	1.00	72.00	35.00
2	30.00	25.00	20.00	10.00	23.00	18.00	1.00	1.00	75.00	51.00	
3	36.00	33.00	36.00	17.00	25.00	41.00	1.00	1.00	105.00	83.00	
4	28.00	25.00	9.00	10.00	18.00	11.00	1.00	1.00	62.00	39.00	
5	30.00	25.00	11.00	11.00	21.00	8.00	1.00	1.00	66.00	40.00	
6	20.00	25.00	6.00	9.00	21.00	15.00	1.00	1.00	51.00	46.00	
7	17.00	21.00	6.00	5.00	10.00	10.00	1.00	1.00	44.00	25.00	
8	33.00	31.00	30.00	11.00	23.00	18.00	1.00	1.00	94.00	52.00	

図 11: 6 科目の心理テストのデータ

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¹⁵ことで MANOVA が実行できる。

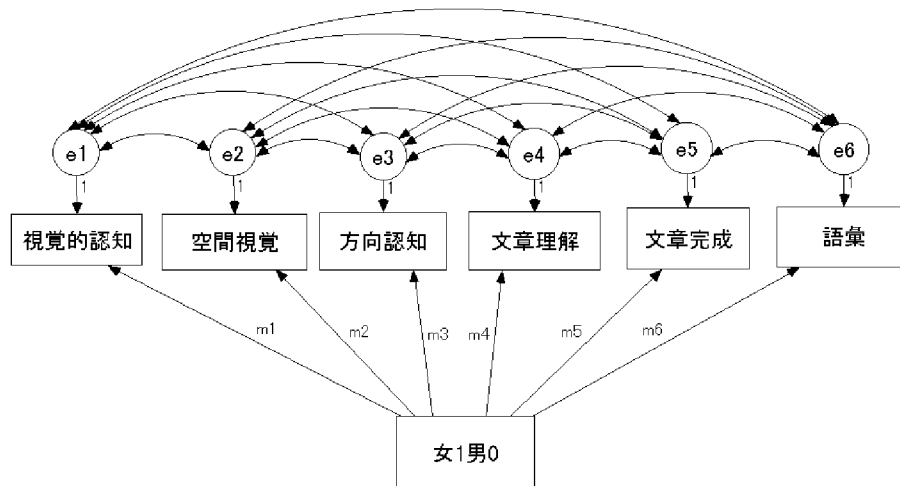


図 12: 共分散構造分析による ANOVA・MANOVA

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は構造を入れない飽和モデルが仮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カイ 2 乗値は 0 である。よって、difference test は、(4.1) の下でのカイ 2 乗値と等しくなる。共分散構造分析による性差の検定は、

$$\chi_6^2 = 12.624, \text{ 有意確率} = 0.049 \quad (4.2)$$

となる。SPSS による通常の MANOVA との違いはない。

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共分散構造分析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分かっている。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

¹⁵これを difference test という。

当てはめていることになる。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。この仮定は通常の MANOVA と同一である。

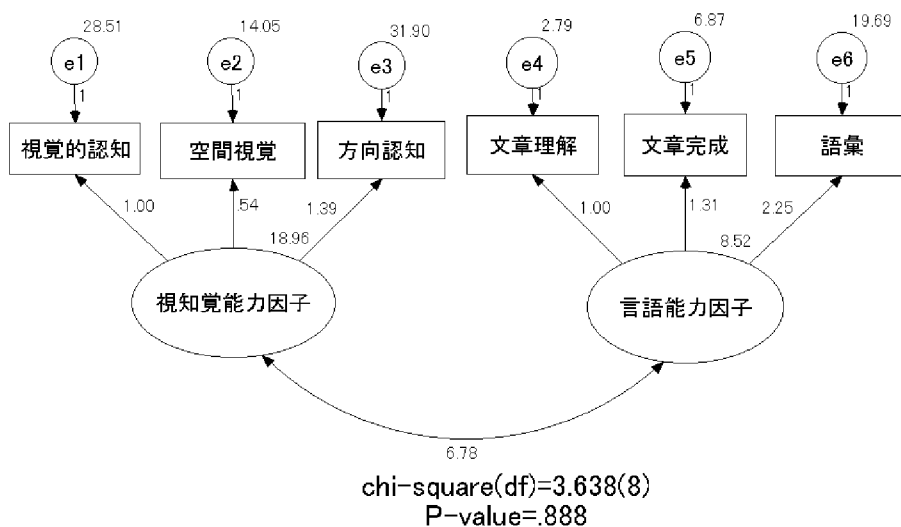


図 13: 検証的因子分析モデル

4.2 拡張と比較

前節で紹介したのは、通常の ANOVA・MANOVA を共分散構造分析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。一つ目は観測変数間の共分散行列には飽和モデルが仮定されていたが、任意の共分散構造モデルが設定可能なことである。この心理テストの検証的因子分析結果は図 13 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 14 となる。 m_i を自由に推定するモ

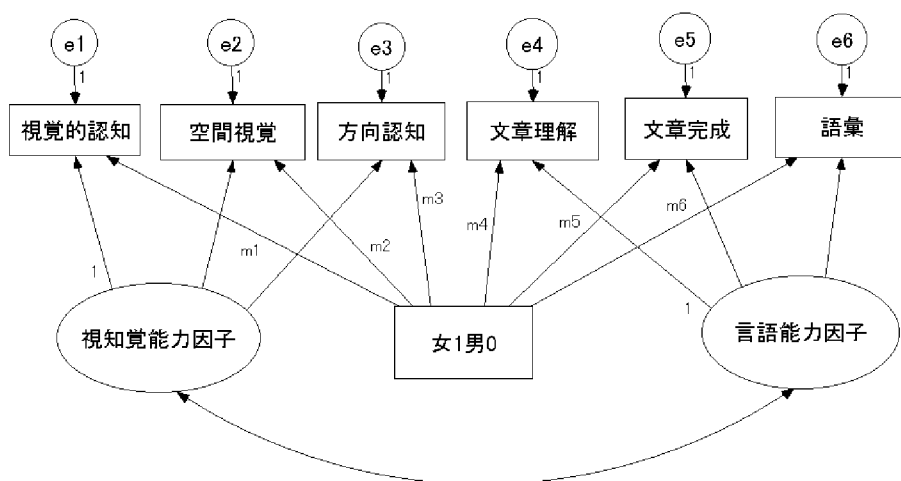


図 14: 検証的因子分析モデルの下での ANOVA・MANOVA

デルのカイ 2 乗値と (4.1) の下でのモデルのカイ 2 乗値との差は

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

(4.3)

となる。difference test の有意確率は 0.051 である。

以上の分析結果をまとめると表 4 となる。有意確率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 4: 有意確率の比較

従来法 (SPSS)	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 項目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。検証的因子分析結果に従い、「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, 有意確率=0.242(視覚的認知), F-値=2.370, 有意確率=0.126(言語能力) となり有意水準 5% で有意差は確認できなかった。

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れないと思うが、ここでは共分散構造分析との比較のため分析してみた。有意確率は 0.42 となった。

「尺度化+分散分析」に対応する共分散構造分析は、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

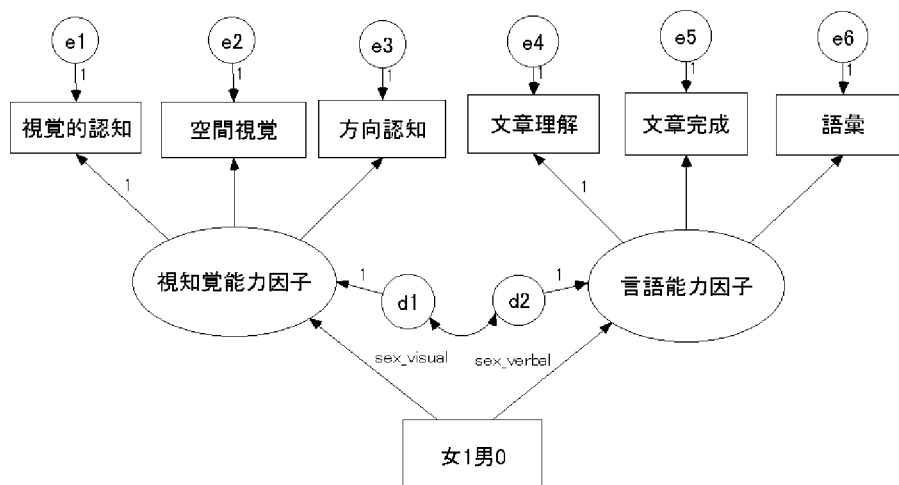


図 15: 検証的因子分析モデルの下での ANOVA・MANOVA

(図 15). すなわち, 図 15 において $\text{sex_visual}=\text{sex_verbal}=0$ 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. 図 15 の分析においては

$$\chi_2^2 = 8.323, \text{ 有意確率} = 0.016 \quad (4.4)$$

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 16). 検定結果は, $\chi_1^2 = 1.722$

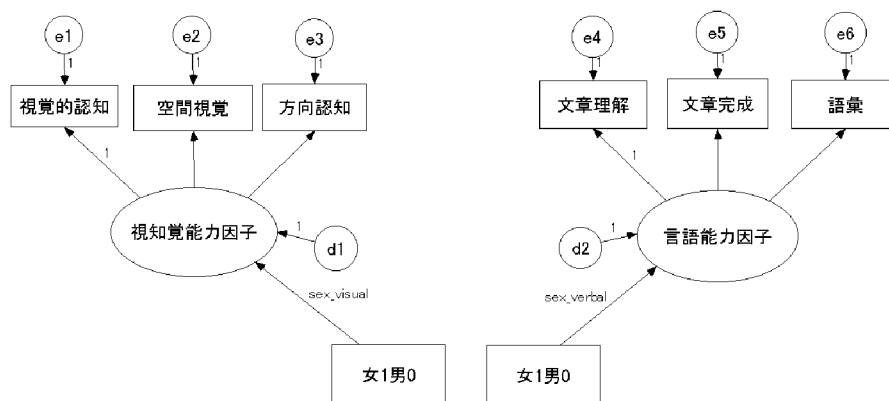


図 16: 因子ごとの効果

有意確率=0.189(視覚能力), $\chi_1^2 = 3.240$ 有意確率=0.072(言語能力) となる.

表 5: 有意確率の比較

尺度または因子	従来の方法 ANOVA・MANOVA	共分散構造分析 因子平均の推測 (図 16・図 15)
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

以上の結果をまとめると表 5 となる. 従来の方である ANOVA・MANOVA と共分散構造分析である因子平均の推測の結果とを比較すると, 確かに, 共分散構造分析の方が一様に有意確率が小さくなっている. このことから, 共分散構造分析の方が検出力が少し高いかもしれないが, 二つの分析方法間に“有意な差”はないとみるのが modest であろう¹⁶.

¹⁶ MANOVA は正確な F-検定であるが, 共分散構造分析では近似的なカイ 2 乗検定が用いられている. 正確な F-検定の方が有意確率が大きくなる傾向があるので, 表 5 の違いは (分析方法の優劣ではなく) この傾向を示しているだけかもしれない.

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共分散構造分析による方法とがある。

結論を先に申し上げると、共分散構造分析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は一つもない。常に、全因果関係を同時に分析する共分散構造分析によって、パス解析を実行すべきである。理由は以下のとおり。

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、仮説の妥当性は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・ワルド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を出さず、信じるものは自分自身の仮説のみである。暗中模索の因果分析といえよう。

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は多い。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解に大いに役立つ。従来法でも手計算は可能であるが、極めて面倒である。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。

4. SEM では、誤差相関（偏相関）を入れたモデリングが可能であり、また、誤差相関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できないかもしれない。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。偏相関は誤差相関によってモデル化される。従来法では誤差相関を扱えないばかりか、誤差相関が必要かどうか分からない。

5. SEM では、双方向因果モデルが記述可能である。

従来法では、数学的に双方向因果モデルを構築できない。

6. SEM では、観測変数が尺度であるとき、もし尺度の信頼性 ρ が既知であれば、 ρ を考慮した分析が可能である。

具体的な方法を図 17 で示す。

7. 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母数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。

最後に、信頼性を調整したパス解析（ここでは単回帰分析）を解説する。まず、入力を標準化する、つまり、相関行列を入力にして全観測変数の分散を 1 にする。尺度の誤差分散の値を「1.00 - 信頼性」に固定する。図 17 がパス解析を実行する SAS プログラムである。図 18 には非標準解と標準解をパス図で示してある。標準解で言語的能力から数学的能力へのパス係数が 0.64 と推定されていることが分かる。この値は、図 9 (上) にある共分散構造分析による結果と一致している¹⁷。

¹⁷ 正確に一致しているわけではない。測定モデルの因子負荷量が同じ大きさならば完全に一致する。


```

DATA kamoku1(TYPE=CORR);
  INPUT _TYPE_ $ _NAME_ $ X1-X2;
  LABEL X1='代数+幾何+解析'
        X2='国語+古文';
  CARDS;
N      .   250  250
CORR X1 1.00 .
CORR X2 0.45 1.00 /* 尺度間相関 0.45 は事前に計算しておく */
;
TITLE '*** 信頼性の補正 ***';
PROC CALIS DATA=kamoku1 METHOD=ml ALL;
  LINEQS
  X1 = F1          + E1,
  X2 = F2          + E2,
  F1 = BETA12 F2 + D1;
  STD
  E1 = 0.27,
  E2 = 0.33,
  D1 = PSI1,
  F2 = PHI2;
RUN;

```

図 17: SAS プログラム：信頼性を調整した分析

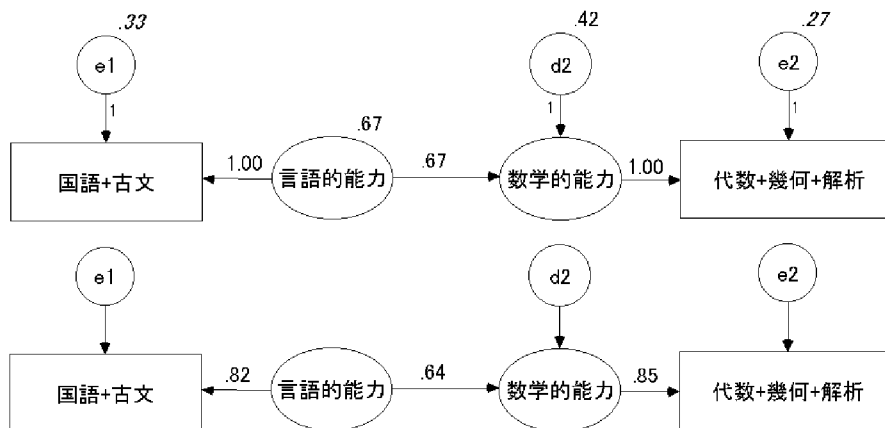


図 18: 非標準解 (上) と 標準解 (下) (誤差変数 e1, e2 の分散 (0.33, 0.27) は固定パラメータ)

6 さいごに

ここで紹介した議論をまとめると表 6 となる。

EFA の枠組みで共分散構造分析が活躍するのは、EFA が実行できないときや、EFA が使えても統計的証拠が必要なときである。また、共分散構造分析は、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。共分散構造分析では、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。

希薄化の修正とは尺度の信頼性の問題であった。尺度には、信頼性の他にさまざまなチェックポイントがあり、その中の一つに内容的妥当性がある。内容的妥当性は、尺度に含めるべき項目がすべて含まれているかどうかの指標である。言い換えれば、尺度に含めた項目にバラエティがあるかどうかを問うており、

表 6: 比較のまとめ

分析方法	分析状況	従来法	共分散構造分析
因子分析		探索的因子分析	検証的因子分析
	因子についての仮定	○	×
	統計的証拠	×	○
	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	×	○
構成概念の		尺度化+相関・回帰分析	潜在変数の相関・回帰分析
因果分析	信頼性が高い	○	○
	信頼性が低い	×	○
	単純性を崩す項目	×	○
構成概念の		尺度化+分散分析	潜在変数の分散分析
実験的分析	3水準以上	○	△
	交互作用	○	△
	多重比較	○	×
	異なる分散や共分散行列	×	○
パス解析		重回帰分析の繰返し	SEMによる同時分析
	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	×	○

信頼性や内的整合性とはやや矛盾する妥当性規準である。尺度項目をそれなりの個数集める理由は、信頼性を高めるという理由の他に、内容的妥当性を高める意味もある。

本稿では、この内容的妥当性についてはまったく考慮していない。筆者の勉強不足で、内容的妥当性が乏しい項目だけで分析することの問題点を具体的に捉えられていないから、議論することができなかった。指定討論者の方々に補足説明をお願いしたい。

最後に名称について一言。共分散構造分析は構造方程式モデリング (Structural Equation Modeling) とも呼ばれる。外国では、こちらの方がメジャーである。共分散構造分析は共分散の構造だけでなく平均の構造も分析できるから、構造方程式モデリングの方が妥当な名称ということであろう。Structural Equation Modeling の短縮形である SEM もよく用いられる。

共分散構造分析に関する (国際) メイリングリストがあり SEMNET という。筆者もメンバーの一人である。SEM に興味のある方には有益な情報源になる。以下に URL をあげておくので参考にされたい：

<http://www.gsu.edu/~mkteer/semnet.html>

参考ソフトウェア一覧

AMOS <http://www.spss.co.jp/product/amos40.htm>

EQS <http://www.mvsoft.com/> 日本代理店：<http://www.informatiq.co.jp/EQS/EQS1.html>

SAS <http://www.sas.com/offices/asiapacific/japan/>

SPSS <http://www.spss.co.jp/index.htm>