

# 共分散構造分析における数値解析上の課題に関する一考察

目黒 達也<sup>1</sup>・Borith LONG<sup>2</sup>・中辻 隆<sup>3</sup>

<sup>1</sup>学生会員 北海道大学大学院工学院（〒060-8628 札幌市北区北13条西8丁目）

E-mail: meg@eng.hokudai.ac.jp

<sup>2</sup>学生会員 北海道大学大学院工学院（〒060-8628 札幌市北区北13条西8丁目）

E-mail: borith\_long@hotmail.com

<sup>3</sup>正会員 北海道大学大学院教授 工学研究院（〒060-8628 札幌市北区北13条西8丁目）

E-mail: naka@eng.hokudai.ac.jp

土木計画学研究における市民の意識構造分析に、共分散構造分析が広く用いられるようになってきている。本研究では、4種類のSEMプログラムと3種類のサンプルデータを用いて、その共分散構造分析における数値解析上の課題である、①残差分散が負値になったり、制約条件によってパス係数の推定値が変動すること、②初期値の正負を変えると、パス係数の推定値の符号が反転すること、③使用するソフトウェアの違いにより解の有無や値自体が変化することに加え、解の多様性について検証分析を通して考察し、信頼度の高い解を得るには、与える初期値と残差分散の制約条件、解の多様性を十分理解した上で分析を行い、結果の解釈をする必要があることがわかった。

**Key Words** : Structural Equation Modeling, Genetic Algorithm, Solver, Amos

## 1. はじめに

土木計画学研究に関わる市民の意識構造を分析する際に、共分散構造分析（Structural Equation Modeling，以下SEM）が広く用いられるようになってきている。SEMは、パス図の自動描写などの高い利便性から、その分析にはAmos<sup>1)</sup>をはじめとする専用ソフトウェアが広く用いられており、その計算過程はブラックボックスとなっている。汎用SEMソフトを用いて意識構造の分析を行った際に、解が存在しないこともしばしば経験することである。また、SEMでは、解の存在や一意性についても、初期値により解の有無が起こること<sup>2)</sup>、残差分散の値が負値になること<sup>3)</sup>が指摘されてきた。さらに、著者らによる経験でも、以下のような現象が観測されている。

- ① 残差分散が負値になったり、制約条件によってパス係数の推定値が変動すること。
- ② 初期値の正負を変えると、パス係数の推定値の符号が反転すること。
- ③ 使用するソフトウェアの違いにより解の有無や解の値自体が変化すること。

本研究では、SEMにおける数値解析上の課題を分析するため、4種類のSEMプログラムと3種類のサンプルデータ

を用い、様々な条件下における数値解析を行った。

## 2. 分析方法

### (1) 分析プログラム

本研究では、分析の信頼性を高めるために4種類のSEMプログラムを用いて分析を行った。以下にその4種類を示す。

- ① Amos18.0<sup>1)</sup> (P1)
- ② 参考文献<sup>2)</sup>に付属したSEMソフト (P2, Kソルバー)
- ③ 自作したSEMプログラム (P3, 自作ソルバー)
- ④ GA<sup>4)</sup>による非線形最適化プログラム (P4)

なお、P2およびP3は、Excelのアドイン機能の1つである「ソルバー」を使って未知数を求める数値計算を実行し、未知数部分に与える初期値を変化させることができる。P2に対しP3は、初期値に乱数を与える機能および推定値の範囲に制約条件を設ける機能を付加している。また、これらすべてのプログラムで条件を適切に設定した上で分析を行った場合、ほぼ同値の解が得ることを確認している。

## (2) 使用データ

本論文で使用するサンプルデータは、参考文献<sup>3</sup>より引用した①と②と、筆者らが実施したアンケート調査から得られた③の3種類である。

### ① 参考文献<sup>3</sup>図6.1の例題

「自分や家族の自分の趣味に対するイメージ」を表すモデル。(3章(1) 図-1, 図-2, (2) 図-3で使用)

### ② 参考文献<sup>3</sup>図6.3の例題

「収入や学歴が社会的地位に及ぼす影響の程度, 社会的地位が人脈や知名度に及ぼす影響の程度」を表すモデル。(3章(4) 図-5, 図-6, 図-7で使用)

### ③ 著者らが実施したアンケート調査のデータ

学生を対象とした飲酒問題に関するアンケート調査により得られたデータ。(3章(3) 図-4で使用)

## (3) 分析手法

4種類すべての分析プログラムにおいて、最尤法と最小2乗法により数値計算を実行している。なお、本論文での解は、すべて最尤法によるものである。

## 3. 共分散構造分析における数値解析上の問題点

1章に示したように、SEMにおける問題点の検証分析を行った結果、大きく3つの数値解析上の問題点が明らかになった。以下にその3つの問題点を示す。さらに、解の多様性や一意性についてより深く分析するため、SEMに非線形最適化計算である遺伝的アルゴリズム(GA)を適用した。

### (1) 残差分散の制約条件による問題

1つ目の課題は、残差分散が負値になったり、制約条件によってパス係数の推定値が変動することである。2章(2), ①のデータを用いて、P1 (Amos18.0, 上段太字) とP2 (Kソルバー, 下段斜字) により分析したパス係数および残差分散の値を図-1に示す。e<sub>2</sub>およびe<sub>6</sub>に注目すると、残差分散の値が負値になっていることがわかる。これら負値は残差分散の値としてふさわしくなく、不適解であるといえる。そこで、P3 (自作ソルバー) を用いてこの残差分散の値を0-1の間で解を導くよう、制約条件を加えて分析を行い、パス係数行列や残差分散行列になにも制約をかけずに分析した結果との比較を図-2に示す。上段太字が制約なし、下段斜字が制約ありのパス係数と残差分散の値である。制約条件の有無により、残差分散だけでなくパス係数の値にも差が生じていることがわかる。このような場合、制約を加えたほうがより信頼できる解であると考えられる。なお、ここでソルバー実行前に与えた初期値は、パス係数、残差分散とも0.5で

あり、パス係数は標準化解、残差分散は非標準化解である。

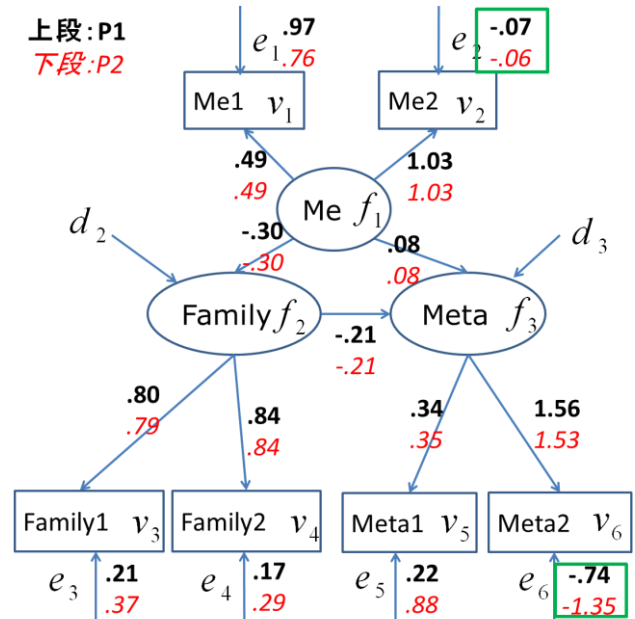


図-1 データ①のP1とP2による分析結果

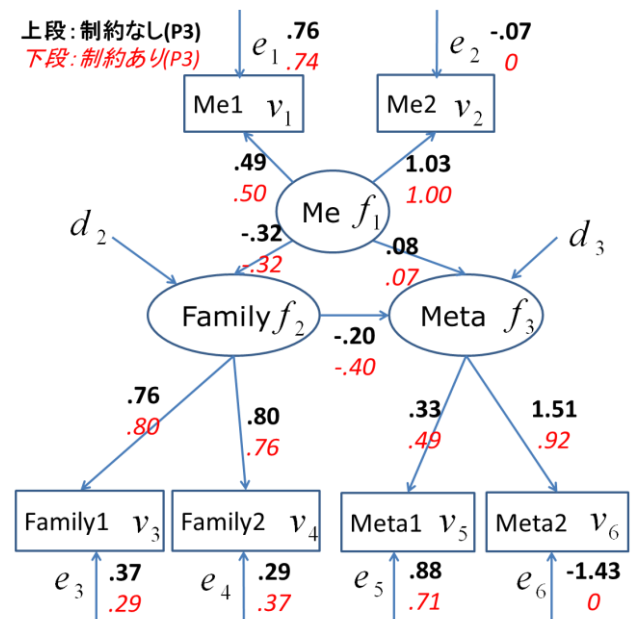


図-2 制約条件の有無による分析結果の比較 (P3)

### (2) 初期値による問題

2つ目の課題は、初期値の正負を変えると、パス係数の推定値の符号が反転することである。2章(2), ①のデータを使用し、P2およびP3を用いて様々な値を初期値に

与えて分析を行った。図-3に、P3を用いてパス係数、残差分散とも0.5を与えた初期値A（上段太字）およびパス係数は-1~1間の乱数を与え、残差分散には0~1間の乱数を与えた初期値B（下段斜字）の2パターンにおけるパス係数の推定値（標準化解）を示す。これらの初期値は、解として導かれる可能性のある範囲の値であり、初期値として妥当な値であるといえる。表-1に、それぞれに与えた初期値を示す。図-3に示すように、初期値Bは初期値Aに比べて絶対値は一致しているが、一部符号が反転した結果となっている。他のモデルでも同様の検証を行った結果、負値の初期値が入る位置に応じて、一部に符号の反転した解が生じることがわかった。また、初期値Aのパターンのパス係数の初期値を-0.5に変更した場合も、ソルバー実行後の目的関数が同値であるにも関わらず、同様の現象が起ることを確認している。

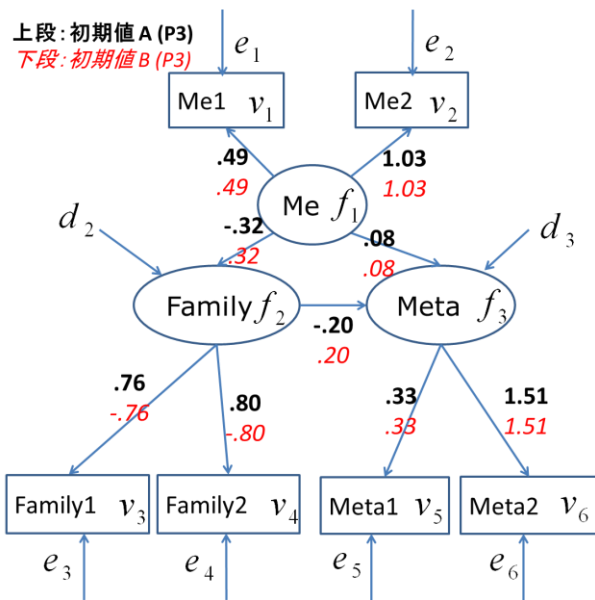


図-3 初期値Aと初期値Bの分析結果の比較 (P3)

表-1 初期値Aと初期値Bの値

	パス係数	残差分散
初期値A	0.5	0.5
初期値B	- 1~1間の乱数	0~1間の乱数

### (3) 使用するソフトウェアの違いによる問題

SEMで解が求まらないことは屢々経験することであるが、2章(2)、③のデータに対してもAmos18.0による分析を行ったところ、解が存在しないという経験をした。ここでの分析モデルを図-4に示す。そこで、P2を用いて

同様のデータ、モデルを使用し、初期値を変更しての再分析を試みたところ、いくつかの初期値の組み合わせで解を得ることができた。表-2は、解が得られた初期値の組み合わせと、得られなかった組み合わせの例を示している。また、P3においても同様の傾向が見られた。なお、最尤法で解がでないモデルでも、最小2乗法を適用することで解が得られるというケースもいくつか確認された。

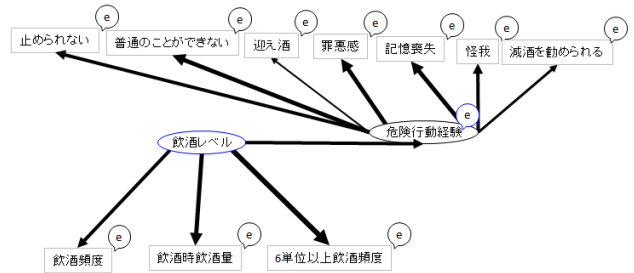


図-4 データ③の分析モデル

表-2 解の有無に対する初期値の組み合わせ

	解あり			解なし		
パス係数	0.2	0.3	0.6	0.1	0.1	0.2
残差分散	0.2	0.5	0.6	0.7	0.8	0.8

### (4) 遺伝的アルゴリズムによる非線形最適化

解の多様性や一意性についてより深く考察するため、遺伝的アルゴリズムを適用してSEMを行った。単純交叉数20、単一算術交叉数20、全体算術交叉数を20とし、一様突然変異数20、境界突然変異数20、非一様突然変異数は40である。その他のパラメータは、GAのデフォルト値を使用している。

SEMにGAを適用したプログラムP4を使用し、いくつかのモデルを用いて、様々な設定条件の下に解の多様性や一意性について更なる分析を行った。どのモデルにも共通して、繰り返し回数に応じてP2、P3でのソルバー実行後の目的関数の値よりも観測変数の数に限りなく近づくが、推定値の制約条件をかけない、もしくはあまりに範囲が広ければ、激しく解が変動することがわかった。

また、2章(2)、②のデータ(図-5のモデル参照)を用いて、表-3に示す制約条件および初期値の下でパス係数と残差分散を求めた。このとき、母数を100、繰り返し回数を500回とした。図-6は、図-5に示すパス係数 $\alpha_{411}$ と目的関数の値が、繰り返し回数によってどのように変化するかを表すグラフである。横軸は繰り返し計算回数を対数目盛でとり、左縦軸はパス係数の推定値、右縦軸は

目的関数の値を示している。なお、最尤法により計算を実行した。図-6を見ると、繰り返し回数が10回程度で目的関数は収束に近づくが、収束に近づいてからもなお、目まぐるしく解が変動していることがわかる。これは、他のパス係数や残差分散についても同様の傾向が見られた。

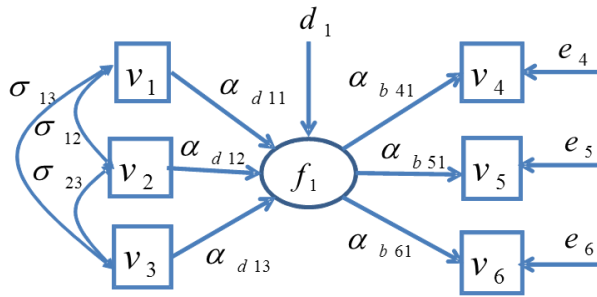


図-5 データ②の分析モデル

表-3 制約条件および初期値

	パス係数	残差分散 (対角)	残差分散 (非対角)
制約条件	- 2~2	0~1	- 1~1
初期値	- 2~2間の乱数	0~1間の乱数	- 1~1間の乱数

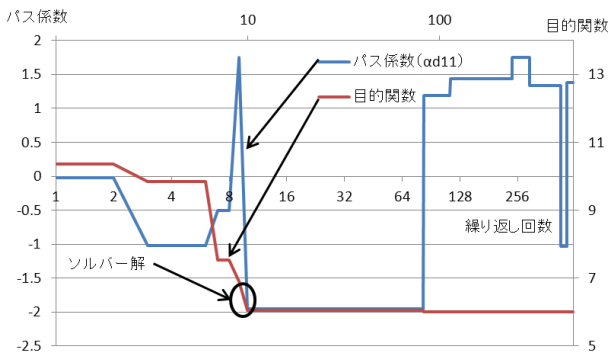


図-6 繰り返し回数に対するパス係数  $\alpha_{d11}$  と目的関数の値 (P4)

次に、P1, P2 (上段太字), P3 (中段太斜字) およびP4 (下段斜字) で分析したパス係数の値の比較を図-7に示す。P2およびP3の初期値は、パス係数と残差分散の非対角要素が0.3, 残差分散の対角要素が0.5であり、推定値の制約条件は設けていない。P4の初期条件は、表-3に示したものと同条件である。P2とP3のソルバー実行後

の目的関数値はともに6.32452, P4の最終的な目的関数の値は6.00022であり、収束に非常に近い値である。両者を比較すると、その結果は大きく異なり、1)ソルバー解がまだローカルミニマムに留まっていること、2)パス係数や残差分散がわずかな目的関数値の変化で大きく変動することを表している。

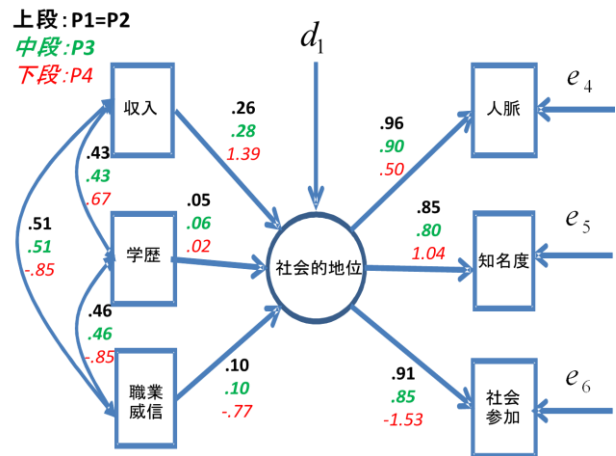


図-7 P1, P2, P3の解とGA解 (P4) の比較

#### 4. 結論

本研究では、4種類のSEMプログラムを用いて、SEMにおける数値解析上の課題を分析し、さらに解の多様性について言及した。初期値や制約条件、使用するソフトによって解が変動したり、有無があることがSEMの弱点であり、そのような条件下で導いた解がどの程度信頼できる結果であるかはわからないのが現状であると考えられる。信頼度の高い解を得るには、与える初期値と残差分散の制約条件に加え、多様な解が存在する可能性があることを十分理解した上で分析を行い、結果の解釈をすることが必要であるといえる。

#### 参考文献

- 1) James L. Arbuckle : IBM SPSS Amos19 ユーザーズガイド, IBM, 2011.
- 2) 小島隆矢 : Excel で学ぶ共分散構造分析とグラフィカルモデリング, オーム社, 2003.
- 3) 豊田秀樹 : 共分散構造分析ー構造方程式モデリングー[入門編], 朝倉書店, 1998.
- 4) Zbigniew Michalewicz : Genetic Algorithms + Data Structures = Evolution Programs, Springer, 1996.

(?????.?? 受付)