

潜在的態度変数を用いた潜在セグメントモデルと LISREL モデルの同時推定 Simultaneous Estimation Method for the LISREL and Latent Segment Model Considering the Heterogeneity of Decision Maker

佐々木邦明
Kuniaki SASAKI

1. はじめに

非集計離散型選択モデルに個人の嗜好の違いに関する情報を取り入れる方法論として、外生変数による古典的なアプリオリセグメント手法や、非観測要因による異質性を考慮したマスポイント手法¹⁾などがあげられる。これらは個人の異質性をモデルの仮定を変えることで考慮するものであり、明快な行動論的背景を持っていない。これに対し、筆者らは、消費者意思決定パスダイアグラムに基づき、潜在的態度変数の同定に LISREL モデルを、効用関数のパラメータの変動に潜在セグメントモデルを用いて、適合度の高いモデルが構築できることを示した^{2) 3)}。

この手法の問題点として、態度指標の調査時点と選択が行われた時点の時間的な差による、認知的不協和解消行動の影響、段階推定時に潜在的態度変数のフィッティッドバリューを確定変数として用いることによる統計的な性質の問題、潜在セグメントの数を分析者が与えることの恣意性などがあげられた。

本研究では認知的不協和解消行動による問題と、段階推定による統計的な性質の問題を同時に解決するために、モデルシステム全体の同時推定を行う手法を提案する。同時推定を行うことにより、態度指標と選択結果が、すべてのパラメータの推定値に影響を与えるため、認知的不協和解消によるバイアスの問題は回避可能であると考えられる。そして、提案したモデルを実際のデータに適用し事例研究を行う。

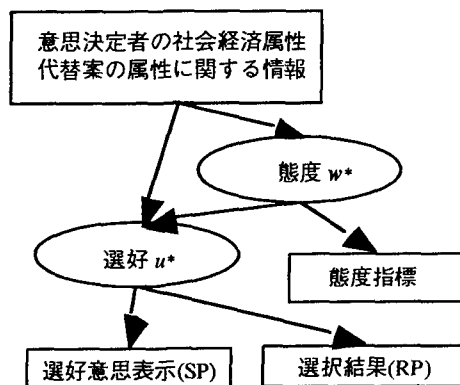


図1 態度を考慮した選択モデル

2. モデルのフレームワーク

図1は潜在的態度を離散型選択モデルに組み入れるモデルのフレームワークをパスダイアグラムで示したものである。このパスダイアグラムを以下のように定式化した。

まず、意思決定者の社会経済属性と代替案の属性に関する情報と態度の関係を(1)式で表す。

$$w^* = Bs + \zeta \quad (1)$$

ただし

w^* : 潜在的態度変数ベクトル

s : w^* を形成する客観的説明変数ベクトル

B : 未知パラメータ行列

$\zeta \sim MVN(0, \Psi)$ に従うランダム項 (Ψ は共分散行列)

続いて態度と態度指標の関係を(2)式で表す。

$$Y = \Lambda w^* + \varepsilon \quad (2)$$

ただし

Y : アンケートで得られた態度指標ベクトル

Λ : 未知パラメータ行列

$\varepsilon \sim MVN(0, \Theta)$ に従うランダム項 (Θ は共分散行列)

一般的に態度指標 Y は、1-0などの離散指標

キーワード: 交通行動分析, 交通手段選択

正員 博士(工学) 名古屋大学大学院工学研究科土木工学専攻

〒464-8603 名古屋市中千種区不老町

Tel.052-789-3565, Fax.052-789-3738

で得られることが多い。つまり、この要因を重視しますか?などの質問にイエス、ノーのような回答になることが多い。このとき連続的な変数を想定する(2)式は妥当性を欠き、(3)、(4)式のようにモディファイする必要があると考えられる。

$$\mathbf{Y}^* = \Lambda \mathbf{w}^* + \varepsilon \quad (3)$$

$$\mathbf{Y} = \begin{cases} 1: & \text{if } \mathbf{Y}^* \geq \mu \\ 1: & \text{if } \mathbf{Y}^* < \mu \end{cases} \quad (4)$$

ただし、

$\mathbf{Y}^*$: 潜在的な態度指標

μ : 態度指標が選択される閾値

(3)、(4)式は各態度指標が選ばれる確率が離散型選択モデルと同様に定式化可能であることを示している。ここで、 $\mathbf{w}^*$ が与えられたときの各態度指標の同時出現確率は

$$\prod_k \left[\Phi \left(\frac{\mu - \lambda_k \mathbf{w}^*}{\theta_k} \right)^{1-y_k} \times \left\{ 1 - \Phi \left(\frac{\mu - \lambda_k \mathbf{w}^*}{\theta_k} \right)^{y_k} \right\} \right] \quad (5)$$

となる。

ただし、 Φ は標準正規分布の分布関数を表す。

(1)式より $\mathbf{s}$ が与えられたとき $\mathbf{w}^*$ は多変量正規分布になるので、各態度指標の同時出現確率は以下のように表すことができる。

$$\int \prod_k \left[\Phi \left(\frac{\mu - \lambda_k \mathbf{w}^*}{\theta_k} \right)^{1-y_k} \times \left\{ 1 - \Phi \left(\frac{\mu - \lambda_k \mathbf{w}^*}{\theta_k} \right)^{y_k} \right\} \right] \times \phi \left(\frac{\mathbf{w}^* - \mathbf{B}\mathbf{s}}{\varphi} \right) d\mathbf{w}^* \quad (6)$$

ただし、 ϕ は標準正規分布の密度関数を表す。

また、態度を考慮した離散型選択モデルは、効用関数のパラメータをノンパラメトリックに分布させる潜在セグメントモデルを適用する。潜在セグメントモデルとは、潜在クラス分析の手法⁴⁾を離散型選択モデルへ適用したものであり、異なる効用パラメータをもついくつかの(潜在的な)セグメントが母集団に存在し、各個人はいずれかのセグメントへ帰属していると仮定している。しかし、帰属するセグメントは直接観測が不可能であるため、セグメント帰属を決定する構造式に確率項を導入し、セグメントへの帰属は確率的な値になるとする。

帰属するセグメント s , $s=1, \dots, S$ が与えられたとき、ランダム効用理論に基づく効用関数は

$$U_{nls} = \alpha_s X_{ni} + v_{nls} \quad (7)$$

ただし、

U_{nls} : セグメント s に帰属した個人 n の選択肢 i に対する効用

α_s : セグメント s の効用パラメータベクトル

X_{ni} : 個人 n の属性と選択肢 i の属性のベクトル

v_{nls} : セグメント s に帰属した個人 n の選択肢 i に対する効用の誤差項

になる。またあるセグメント s への個人の帰属確率は、帰属度関数を

$$\Gamma_s \mathbf{w}_{ns}^* + \delta_s \quad (8)$$

ただし、

Γ_s : 未知パラメータベクトル

δ_s : セグメント帰属の確率項

と定義し、各セグメントへの帰属確率の誤差項に独立で同一のガンベル分布を仮定すると、個人の態度変数 $\mathbf{w}_{ns}^*$ が与えられたという条件の下、次式のようなロジット型で表される。

$$H_n(s | \mathbf{w}_{ns}^*) = \frac{e^{\Gamma_s \mathbf{w}_{ns}^*}}{\sum_{k=1}^S e^{\Gamma_k \mathbf{w}_{nk}^*}} \quad (9)$$

このとき個人 n が選択肢 i を選ぶ条件付き確率は、

$$P_n(i | \mathbf{w}_{ns}^*) = \sum_{s=1}^S \left[P_{nls}(i) H_n(s | \mathbf{w}_{ns}^*) \right] \quad (10)$$

となる。

ここで、ある個人 n が選択肢 i を選択する同時確率は、潜在的態度変数 $\mathbf{w}_{ns}^*$ の確率分布を用いて、次式で得られる。

$$P_n(i) = \int_{-\infty}^{+\infty} P_n(i | \mathbf{w}_{ns}^*) f(\mathbf{w}_{ns}^*) d\mathbf{w}_{ns}^* \quad (11)$$

ただし、

$f(\mathbf{w}_{ns}^*)$: $\mathbf{w}_{ns}^*$ の同時確率密度関数

(1)式より $\mathbf{s}$ が与えられた場合の $\mathbf{w}_{ns}^*$ の分布は多変量正規分布であることから、(6)式と(11)式の関係を用いて、選択モードと態度指標 $\mathbf{Y}$ の同時出現確率を、以下のように表すことができる。

$$P_n(i, \mathbf{Y}) = \int_{-\infty}^{+\infty} P_n(i | \mathbf{w}_{ns}^*) \times \prod_k \left[\Phi \left(\frac{\mu - \lambda_k \mathbf{w}_{ns}^*}{\theta_k} \right)^{1-y_k} \times \left\{ 1 - \Phi \left(\frac{\mu - \lambda_k \mathbf{w}_{ns}^*}{\theta_k} \right)^{y_k} \right\} \right] \times \phi \left(\frac{\mathbf{w}_{ns}^* - \mathbf{B}\mathbf{s}}{\varphi} \right) d\mathbf{w}_{ns}^* \quad (12)$$

この(12)式を各個人の尤度として、サンプル全体の同時確率を最大にすることにより、すべての未知パラメータが同時に推定される。

3. 事例研究

(1) 使用したデータ

本研究で用いるデータは、首都圏の湾岸海浜観光地域を訪れていた人を対象に、無作為にアンケート調査票を配布して郵送回収されたものである。アンケートの内容としては、主に次の3種類に分類することができる。

- 1) 個人の社会経済属性に関するデータ
 - ・性別、年齢、職業、自家用車の有無
- 2) 今回の旅行の手段、経路、旅行内容に関するデータ (RPデータ)
 - ・目的、旅行人数、旅行ルート、旅行費用、交通機関選択理由(態度指標)等
- 3) 新規航路の選択に関するデータ (SPデータ)
 - ・当該湾岸海浜観光地域と対岸を結ぶ仮想の新規高速船(フェリー)が就航したという仮定のもと、新規高速船の所要時間は一定とし、各被験者に料金と運行間隔の組み合わせを示し、新規高速船を利用するか否かの2項選択データを得ている。なお、実際の利用交通手段が車である人と車以外である人では、それぞれ新規高速船へのアクセス・イグレスは現状の利用交通手段を変更しないという条件をつけ、異なる料金設定でSP質問を行った。

アンケートに回答した全サンプル852人のうち、今回の旅行と同条件(旅行人数、旅行目的が同一)の場合に実際に新規高速船の利用が可能であると考えられ、上記の各データについて有効な回答が得られたサンプルは211人であった。ここで利用不可能な人とは、新規高速船の運賃・所要時間・運行頻度などの条件に関わらず利用しないと回答した人、当該地域在住の人、観光バス利用者で交通機関選択が不可能な人である。このうち、全ての人々がSP質問に回答したわけではないのでRPデータ211、SPデータ819、計1030の観測データが得られている。

(2) 推定結果

(1)で示したデータに2.のフレームワークで示したモデルを、推定パラメータを減らすために ϵ や ξ を標準正規分布とすることや、閾値を係数で割って基準化した値をもちいて統一化を図るなどの簡略化を施して適用した。また、同様の理由で、SP、RPデータの違いや同一個人から繰り返し採取したことは考慮せずに、互いに独立なデータとして推定を行った結果が表1及び表2に示してある。これらは便宜上4つの表に分割してあるが、同時に推定されたものである。

表1 LISRELモデルと閾値の推定結果

		w_1^*	w_2^*	
$\hat{B} =$		0.122	0	(cong)
		0.139	0.0495	(trsf)
		0	0.0420	(comf)
		0	0.0677	(ease)
		0	0.220	(selt)
		-3.49	1.11	(chep)
		0.0485	0	(shot)
		0.374	0	(rela)
		w_1^*	w_2^*	
$\hat{A} =$	-0.145	-1.23	(female)	
	0.547	1.62	(age20)	
	-0.0111	-0.191	(age50)	
	0.395	1.70	(job)	
	-0.361	0.679	(sight)	
	-0.206	0.275	(sprt)	

$\mu = [2.16 \ 6.83]$

ただし、

- female: 女性1, その他0のダミー変数
- 客 age20: 年齢20歳以下のダミー変数
- 観 age50: 年齢50歳以上のダミー変数
- 変 job: 有職者のダミー変数
- 数 sight: 旅行目的が観光のダミー変数
- sprt: 旅行目的がスポーツのダミー変数
- cong: 混雑が少ない
- trsf: 乗換えが少ない
- 態度 comf: 快適である
- ease: 荷物を運ぶのが楽だ
- selt: 静かである
- chep: 料金・運賃が安い
- shot: 所要時間が短い
- rela: 所要時間が正確である

である。

このLISRELモデルおよび離散型選択モデルのスペシフィケーションは、筆者らが以前に行った線形構造方程式モデルと離散型選択モデルの段階

表2 潜在セグメントモデルの推定結果

変数名		セグメント1	セグメント2
選 択 モ デ ル	鉄道定数 RP	1.34 (3.8)	0.221 (1.0)
	フェリー定数 SP	-2.40 (-6.4)	-2.13 (-4.2)
	総費用	-0.0552 (-1.6)	-0.157 (-3.7)
	所要時間	-0.0201 (-0.1)	0.940 (3.4)
帰 属 モ デ ル	運行頻度	0.760 (2.0)	4.27 (5.8)
	w_1	-0.507 (-1.6)	
	w_2	-0.196 (-0.7)	
観測数		1030	

() 内は t 値

推定³⁾と同一である。

ここで、LISREL モデルの推定値より、2つの潜在的態度指標は、観光やスポーツといった非日常交通であるかどうかで変動し、費用の安さの態度指標に敏感に反応することが特徴としてあげられ、そのほかのパラメータの値も考慮して、定時性重視変数と費用安楽性重視変数を示していると考えられる。この前提のもとで、潜在セグメントを考慮した離散型選択モデルの推定結果をみると、帰属モデルの推定パラメータから、セグメント1は定時性重視変数が大きくなるほど帰属確率が高まるセグメントで、セグメント2が費用安楽性重視変数が大きくなるほど帰属確率が高まるセグメントであると解釈できる。これらの潜在変数の解釈と、得られたセグメントのパラメータを比較すると、直感と矛盾しない値が得られた。たとえば、セグメント2はセグメント1よりも総費用のパラメータが大きいことや、セグメント1の鉄道定数のパラメータが大きな正の値をとったことなどである。ただ、セグメント2の所要時間のパラメータが正の値をとった理由は、費用を重視することは一般的に、より遅い交通機関であることが多く、そのためにこのような直感に矛盾する結果になったためであると考えられる。

4. おわりに

本研究は、筆者らが先に行った態度を考慮した交通機関選択モデルの発展として、段階的に推定を行っていたものを、同時推定することを主眼としたものである。提案した同時推定法を用いて行

った事例研究は、段階推定モデルと LISREL モデルの誤差項に標準正規分布を仮定するなど、多少の差異はあるが、基本的な構造を同一のスペシフィケーションとしてモデルを推定したにもかかわらず、LISREL モデル、離散型選択モデルとも段階推定と異なる結果が得られた。即断はできないが、このことは、段階推定したパラメータには認知的不協和解消によるアンケート調査のバイアスが存在している可能性があることと、統計的な性質の違いがバイアスを生んでいた可能性が指摘できる。

意識要因を交通行動モデルに取り込むことは、交通行動を分析するために非常に有効な手法である。ただし、アンケートにより得られた意識データには、調査手法を改善することにより、その軽減をはかることは可能であるが、認知的不協和解消などの様々なバイアスが存在する可能性がある。筆者らが以前に行った、知覚値を考慮した行動モデル⁵⁾も、段階推定と同時推定では推定されたパラメータには大きな違いが生じていた。よって、意識要因を交通行動モデルに取り込む際には、これらのバイアスの存在を考慮したモデル構築手法が必要であると考えられる。

参考文献

- 1) 西井和夫、北村隆一、近藤勝直、弦間俊彦(1995) 観測されていない異質性を考慮した繰り返しデータに関するパラメータ推定法：Mass Point Model と Mixing Distribution Model, 土木学会論文集, No.506/IV-26, pp.25-34.
- 2) 佐々木邦明・森川高行・杉本直(1995)：潜在セグメントを考慮した動的な休日買物目的地選択分析, 土木計画学研究・論文集, No.12, pp.397-404.
- 3) 佐々木邦明(1996)：意識構造モデルによって求められた態度変数を用いた RP・SP データの潜在クラスター分析, 土木計画学研究・講演集, No.19, pp.799-802.
- 4) 柳井晴夫、繁舩算夫、前川眞一、市川雅教(1990) 因子分析、その理論と方法、朝倉書店、第5章。
- 5) 森川高行・佐々木邦明(1993)：主観的要因を考慮した非集計離散型選択モデル, 土木学会論文集 IV, No.470, pp.115-124.