

RP/SP融合推定法における系列相関のモデル化

Modeling Serial Correlation in the Context of the RP/SP Combined Estimation

森川高行*・田中小百合**

By Takayuki MORIKAWA and Sayuri TANAKA

1. はじめに

非集計タイプの選択モデルを推定する際に、互いに補完的な性質を持つRPデータとSPデータを同時に用いるRP/SP融合推定法が提案されている¹⁾。このとき、SP質問の回答が実際の行動に影響を受ける「状態依存」と、同一個人内の効用関数の誤差項間に「系列相関」が存在することが確認され、その対処法が提案された²⁾。

状態依存については、SPモデルの説明変数としてRPデータの選択結果を用いることにより、その影響を考慮することができることが確かめられている。また系列相関においても、通常の誤差項を分解し、RPモデルとSPモデルで共有する「個人・選択肢に共通なシステムティックな誤差項」を明示的に導入することによって対応した。しかし、先に提案された方法では、後述するように、RPモデルとSPモデルの共分散行列に拘束条件が存在し、パラメータの推定に影響を及ぼす可能性がある。本研究では、共分散行列内の拘束条件を取り除く、より一般的なモデルの定式化を行い、そのモデルの実証的研究を行った。

2. RP/SP融合推定法の概要

RP/SP融合推定法のフレームワークでは、RPデータを発生させるRPモデルとSPデータを発生させるSPモデルの2つを考え、主要な属性間のトレードオフ関係はRPモデルとSPモデルで共通と考えている。より詳しい説明は文献1)および2)を参照されたい。

キーワード：交通手段選択、意識調査分析

* 正会員 Ph.D. 名古屋大学助教授 工学部土木工学科

** 学生会員 名古屋大学大学院工学研究科土木工学専攻

〒464-01 名古屋市千種区不老町

TEL 052-789-3564 FAX 052-789-3738

RPモデル

$$U_{in}^{RP} = \beta x_{in}^{RP} + \alpha' w_{in}^{RP} + \varepsilon_{in}^{RP} \\ = V_{in}^{RP} + \varepsilon_{in}^{RP}$$

SPモデル

$$U_{in}^{SP} = \beta x_{in}^{SP} + \gamma' z_{in}^{SP} + \varepsilon_{in}^{SP} \\ = V_{in}^{SP} + \varepsilon_{in}^{SP}$$

ただし、

U_{in} : 個人 n の選択肢 i に対する総効用

V_{in} : 総効用 U_{in} の確定項

ε_{in} : 総効用 U_{in} の誤差項

x_{in}, w_{in}, z_{in} : 個人 n の選択肢 i の説明変数ベクトル

α, β, γ : 未知パラメータベクトル

この定式化において、 x はRPモデルとSPモデルで共通の係数ベクトル β を持つ説明変数ベクトルであり、 w, z はそれぞれRPモデル、SPモデルで異なる係数を持つ説明変数ベクトルである。

効用関数の誤差項には、個人、選択肢に特有な誤差項がかなり卓越していることが考えられる。例えば「車が好きだ」というような潜在的態度、所要時間や費用に対する重要度の個人内のばらつき（異質性）、端末交通の時間や費用に関する測定誤差などが挙げられよう。この誤差項は、1人につき1つの選択データのみを用いて推定するときには、誤差項の独立性が保たれるため、大きな問題を引き起こさない。しかし、RP/SP融合推定法を行うときには、同一個人からRPデータ及び複数のSPデータを同時に用いて推定を行うため、個人、選択肢に特有な誤差項の存在がもたらす系列相関の影響を無視できなくなる。その中で、重要な属性の係数の符号の反転といった明らかかつ深刻なバイアスも報告されている。こ

れまでに提案された方法は、誤差項を個人および選択肢に共通なシステムティックな部分 λ と、その他の影響によるランダムなホワイトノイズ v に分解することにより、系列相関に対処した²⁾。

$$\begin{aligned}\varepsilon_{in}^{RP} &= \lambda_{in} + v_{in}^{RP} \\ \varepsilon_{in}^{SP} &= \theta_i \lambda_{in} + v_{in}^{SP}\end{aligned}$$

ここでは、RPモデルの λ と v の分散を同じとし、RPデータとSPデータにおける λ の影響力の違いを未知パラメータ θ によって表している。また、ホワイトノイズの分散の大きさの違いは、スケールパラメータ μ で表している。

$$Var(v_{in}^{RP}) = \mu^2 Var(v_{in}^{SP}) \quad \dots (1)$$

ただし、通常のロジットモデルやプロビットモデルにおける誤差項の分散の正規化と同様に、プロビットモデルの場合は $Var(v_{in}^{RP}) = 1$ 、ロジットモデルの場合はガンベル分布のスケールパラメータを1として正規化する。

仮に、1個人につき1つのRPデータと2つのSPデータがあった場合の ε_{in}^{RP} 、 $\varepsilon_{in}^{SP_1}$ 、 $\varepsilon_{in}^{SP_2}$ の共分散行列は、プロビットモデルの場合、次のようになる。

$$\begin{array}{c} RP \quad SP_1 \quad SP_2 \\ \begin{array}{c} RP \\ SP_1 \\ SP_2 \end{array} \begin{bmatrix} 2 & \theta_i & \theta_i \\ \theta_i & 1/\mu^2 + \theta_i^2 & \theta_i^2 \\ \theta_i & \theta_i^2 & 1/\mu^2 + \theta_i^2 \end{bmatrix} \end{array}$$

この共分散行列には、

$$Cov(\varepsilon_{in}^{RP}, \varepsilon_{in}^{SP_1}) = Cov(\varepsilon_{in}^{RP}, \varepsilon_{in}^{SP_2}) = \sqrt{Cov(\varepsilon_{in}^{SP_1}, \varepsilon_{in}^{SP_2})} \quad \dots (2)$$

および、

$$Var(v_{in}^{RP}) = Var(\lambda) = 1 \quad \dots (3)$$

なる拘束条件が存在する。

RPモデルとSPモデルの選択確率は λ が与えられたときの条件付き確率をロジットモデルやプロビットモデルとして表し、その λ に関する数学的期待値をとることにより求められる。また、効用関数中の係数及びスケールパラメータの推定は、RP/SP同時選択確率から得られる尤度関数の最大化問題を解くことにより求められる。

3. 系列相関の新たな定式化

従来の系列相関の定式化における問題点であった共分散行列内の拘束条件を取り除くため、2種類の新たな定式化を行った。モデルAは、従来の誤差項の分解方法を改良することにより、また、モデルBでは、新たに複数回答のSPデータ間のみの系列相関を表すシステムティックな誤差を導入することにより、拘束条件を取り除いている。

(1)モデルA

従来モデルでは、RPモデル内の系列相関を表すシステムティックな誤差 λ_{in} の分散を1に固定して計算を行っていたが、ここではその分散の大きさを表す未知パラメータ θ_i^{RP} を導入することにより、共分散行列内の拘束条件(式(2))を取り除いている。つまり、このモデルにおいては、誤差項の分解は次のように行われる。

$$\begin{aligned}\varepsilon_{in}^{RP} &= \theta_i^{RP} \lambda_{in} + v_{in}^{RP} \\ \varepsilon_{in}^{SP} &= \theta_i^{SP} \lambda_{in} + v_{in}^{SP}\end{aligned}$$

ホワイトノイズの分散の大きさは式(1)と同じと仮定する。

これにより、RPモデルにおいてシステムティックな誤差項とホワイトノイズの間にある拘束条件(式(3))も取り除くことが可能になる。この定式化による共分散行列は、

$$\begin{array}{c} RP \quad SP_1 \quad SP_2 \\ \begin{array}{c} RP \\ SP_1 \\ SP_2 \end{array} \begin{bmatrix} 1 + \theta^{RP2} & \theta^{RP} \theta^{SP} & \theta^{RP} \theta^{SP} \\ \theta^{RP} \theta^{SP} & 1/\mu^2 + \theta^{SP2} & \theta^{SP2} \\ \theta^{RP} \theta^{SP} & \theta^{SP2} & 1/\mu^2 + \theta^{SP2} \end{bmatrix} \end{array}$$

となり、共分散行列内に明らかな拘束条件は存在しない。

(2)モデルB

次に、RP、SPデータ間の系列相関のみでなく、SPデータ間のみの系列相関をも考慮したモデルの定式化を行う。

$$\begin{aligned}\varepsilon_{in}^{RP} &= \theta_i \lambda_{in} + v_{in}^{RP} \\ \varepsilon_{in}^{SP} &= \theta_i \lambda_{in} + \eta_i \tau_{in} + v_{in}^{SP}\end{aligned}$$

このモデルでは、RP、SPデータ間の系列相関を表すシステムティックな誤差 λ の係数を同じにしたため、両モデル間に λ が与える影響は同じとなり、系列相関の両モデル間の影響の違いは、SPデータ間のみの系列相関を表す ϵ を導入することにより対処することとなる。また、共分散行列は

$$\begin{matrix} & RP & SP_1 & SP_2 \\ \begin{matrix} RP \\ SP_1 \\ SP_2 \end{matrix} & \begin{bmatrix} 1+\theta^2 & \theta^2 & \theta^2 \\ \theta^2 & 1/\mu^2+\theta^2+\eta^2 & \theta^2+\eta^2 \\ \theta^2 & \theta^2+\eta^2 & 1/\mu^2+\theta^2+\eta^2 \end{bmatrix} \end{matrix}$$

となり、従来のモデルに存在した拘束条件は取り除かれるものの、以下に示す拘束条件が新たに生じることになることに注意したい。

$$Var(\epsilon_{in}^{RP}) = Cov(\epsilon_{in}^{RP}, \epsilon_{in}^{SP}) + 1 \quad \dots (4)$$

3. ケーススタディ

前章で提案したモデルの実証研究として、表-1に示す2つのデータを用いてケーススタディを行った。両データ共に都市間交通における機関選択データで、各個人からRPデータとSPデータをとっている。データAのSP質問は各交通機関利用者ごとに最大3問、データBでは、回答者ごとのRPデータの属性値を考

表-1 ケーススタディに用いたデータの概略

	データA	データB
調査目的	・2都市間を結ぶ高速交通機関 ・快適性の高い新型車両を用いた特急増便による転換交通の予測	・オランダにおけるナイメゲン-ランドシユタット大都市圏間の高速交通機関の需要予測
代替交通機関	・鉄道(特急) ・高速バス ・自家用車	・鉄道 ・自家用車
調査法	・選択肢別抽出による調査票配布、郵送回収	・ランダムサンプリングによる住民調査 ・ラップトップコンピューターを用いた「応答型意識調査」
サンプル数	・RP: 255 ・SP: 679	・RP: 228 ・SP: 1511
調査時期	・1990年2月(日本)	・1987年(オランダ)
SPデータ	・選択データ	・一対比較5段階評定データ

慮して応答的にSP質問を作成し、1人最大10数問の回答を得ている。

(1)データAの推定結果

表-2にデータAによる推定結果を示す。まず、適合度を見るとモデルAが従来モデルを若干上回っており、モデルBはこれらのモデルより大きく劣っている。モデルBが劣る理由は、本データは、個人当たりのSP質問の数が少なくSP独自のシステムティックな誤差の影響を同定することに問題があったのと、式(4)の拘束条件が不適当であったためと思われる。

効用関数の係数はほとんど同じ値をとっており、例えば時間価値を計算してみると3モデルとも1200円/時前後の値である。 θ の推定値を見ると、従来モデルに課された式(3)の拘束条件には無理があることがわかり、モデルAの優位性が示された。また、モデルAを用いてデータ間の相関係数行列を計算すると、例えば鉄道の効用では、RPとSPの間の相関係数が0.52、SP間の相関係数が0.91となり、複数SP間

表-2 データAの推定結果

	従来モデル	モデルA	モデルB
鉄道定数項(RP)	1.34 (3.8)	1.29 (4.0)	1.22 (4.0)
バス定数項(RP)	1.73 (4.2)	1.58 (4.6)	1.75 (4.1)
鉄道定数項(SP)	2.57 (1.4)	2.54 (0.9)	1.63 (2.2)
バス定数項(SP)	0.956 (1.6)	0.752 (1.9)	0.778 (1.8)
旅行時間(時間)	-0.523 (-2.0)	-0.450 (-2.0)	-0.412 (-1.7)
1人当たり費用(1000円)	-0.412 (-4.6)	-0.381 (-4.5)	-0.321 (-4.2)
新型特急ダミー	0.901 (1.9)	0.773 (1.9)	0.569 (1.4)
鉄道慣性項	-1.36 (-0.9)	-1.28 (-0.5)	-0.462 (-0.8)
バス慣性項	1.72 (1.3)	1.99 (0.7)	1.04 (1.5)
自家用車慣性項	4.03 (2.0)	4.00 (1.2)	2.83 (1.9)
θ^{RP} - 鉄道	1.00 -	0.645 (2.5)	-0.283 (-1.0)
θ^{SP} - 鉄道	1.82 (3.8)	1.76 (2.9)	
θ^{RP} - バス	1.00 -	0.407 (0.8)	1.29 (3.4)
θ^{SP} - バス	1.37 (2.2)	1.53 (2.4)	
η - 鉄道			1.35 (3.0)
η - バス			0.00004 (0.5)
スケールパラメータ μ	1.47 (2.1)	1.73 (2.1)	2.40 (1.4)
N	255	255	255
L(0)	-707.00	-707.00	-707.00
$L(\beta)$	-504.93	-502.58	-528.98
$\frac{\sigma^2}{e^2}$	0.286	0.289	0.252
$\frac{\sigma^2}{e}$	0.267	0.268	0.231

に相関が強いこともわかった。また、これは式(2)の拘束条件が不適切であることも表している。

また、 ρ^2 および $\bar{\rho}^2$ は、モデルAの値が一番大きくなった。逆に、モデルBの値は、従来のモデルのそれよりも小さくなっている。これは、RPモデルとSPモデルのシステムティックな誤差の大きさを同じにしたことにより生じる拘束条件がこのデータにおいては非常に重大であったこと、さらに、1人当たりのデータ数が少なく、誤差のスケールが定まりにくくなっていることなどによると考えられる。

(2)データBの結果

データBの推定結果においても同様に、適合度は、モデルAで最も高くなるが、データAとは異なり、モデルBにおいても従来モデルより高くなっている。この理由として挙げられることの1つは、データBにおいては1人当たりのSP質問の数がデータAよりも多く、SPデータ間のみの系列相関の影響がデータ内に強く存在することである。これは、その係数のt値が4.1であることから明らかである。2つめに挙げられることは、従来モデルにおけるRP、SPデー

タ間におけるシステムチックな誤差の大きさを1に固定することによって生じる拘束条件が、このデータにおいては非常に厳しいものとなっていることによるものと思われる。実際に、モデルAを用いて鉄道の効用の相関係数行列を計算すると、RPとSP間はほとんど無相関であるが、SP内の相関係数は0.43という結果になった。

このようなデータの特性を持っても、モデルBの適合度が、モデルAのそれよりも低いのは、式(4)の拘束条件の影響であると推測される。効用関数の係数推定値は、モデル間でかなり異なるものがあり、例えば車内および車外旅行時間の時間価値はモデルAから計算される値に最も妥当性がある。

4. まとめ

本研究では、今までに提案された系列相関を考慮したRP/SP融合推定モデルに存在していた共分散行列内の拘束条件を取り除くため、2種類の新たな定式化を行い、そのモデルの実証研究を行った。その結果として、RP、SP間の系列相関を考慮し、その影響度をRPモデルとSPモデルで異なるとしたモデルAが、優れたモデルであるという結果が得られた。しかし、この結果はモデルBを否定するものではなく、個人当たりのSPデータの数やSP質問における属性値の定め方にも大きく左右されるものであると思われる。

計算の負荷の面から考えると、導入したシステムティックな誤差の数だけ数値積分の回数が増え、計算時間が長くなるため、実用上あまり多くの誤差を導入するのは望ましくない。この点からも、計算時間が短くモデルの適合度が高いモデルAが優れているのではないと思われる。

今後、さらに多くの事例研究により、系列相関に関する知見の蓄積を行う必要があろう。

参考文献

- 1) 森川高行、M.Ben-Akiva：RPデータとSPデータを同時に用いた非集計行動モデルの推定法、交通工学、Vol.27、No.4、1992
- 2) 森川高行、山田菊子：系列相関を持つRPデータとSPデータを同時に用いた離散型選択モデルの推定法、土木学会論文集、No476、pp11～18、1993

表-3 データB推定結果

	従来モデル	モデルA	モデルB
鉄道定数項(RP)	0.661 (1.5)	0.637 (2.7)	0.759 (2.3)
鉄道定数項(SP)	-1.08 (-1.8)	-1.77 (-2.9)	-1.21 (-2.3)
1人当たり費用 (Guilder)	-0.0510 (-6.3)	-0.0325 (-6.0)	-0.0395 (-4.7)
車内旅行時間 (時間)	-0.456 (-2.0)	-0.530 (-2.8)	-0.585 (-2.4)
車外旅行時間 (時間)	-1.60 (-1.8)	-1.22 (-3.5)	-1.56 (-3.0)
乗り換え回数	-0.285 (-1.3)	-0.141 (-1.2)	-0.180 (-1.2)
業務目的ダミー	1.03 (2.7)	0.483 (2.0)	0.660 (1.8)
女性ダミー	1.02 (3.4)	0.199 (1.1)	0.279 (1.3)
鉄道慣性項	1.69 (3.3)	3.17 (4.1)	2.38 (3.7)
θ^{RP} - 鉄道	1.00 -	0.0368 (0.3)	
θ^{SP} - 鉄道	1.72 (6.5)	1.51 (4.4)	0.736 (3.1)
η - 鉄道			1.44 (4.1)
スケールパラメータ μ	0.513 (5.9)	0.574 (4.1)	0.610 (4.1)
N	228	228	228
L(0)	-1205.38	-1205.38	-1205.38
L(β)	-712.82	-696.82	-703.28
$\frac{\rho}{\sigma^2}$	0.409	0.422	0.417
$\bar{\rho}$	0.400	0.412	0.407