

標準 RL モデルにおける誤差項の 等分散性の仮定についての一考察

倉内 慎也¹

¹正会員 愛媛大学大学院准教授 理工学研究科 (〒790-8577 愛媛県松山市文京町3番)
E-mail: kurauchi@cee.chime-u.ac.jp

Fosgerau et al.(2013)によって提案された Recursive Logit モデルは、逐次選択型の構造を採用することで経路列挙の問題を回避しつつ、合成効用項により目的地までの効用も明示的に考慮することができるという点で、極めて実用性の高いモデルである。しかし、同モデルが前提としている効用のリンク加法性が誤差項にも成立するとした場合、誤差項については必然的に目的地までの距離等が長いほど分散が大きくなるような異分散性を示し、I.I.D.ガンベル分布の等分散性の仮定に反することになる。本稿では、仮定の格子状ネットワークにおいて、リンクごとに誤差効用が生ずるような状況を想定して生成した人工的な経路選択データを用いて、同モデルを適用した場合の問題点について検証した結果を報告する。

Key Words: discrete choice model, route choice, recursive logit model, heteroscedasticity

1. はじめに

都市圏における環状道路や鉄道路線の整備効果の分析、幹線道路の整備に伴う生活道路への流入交通量の変化の分析など、交通需要分析が対象とするネットワークは年々大規模になっている。同時に、時々刻々と変化する交通需要の把握や、リアルタイム情報の提供効果の分析など、分析に求められる要求性能も高くなっている。加えて、ETC2.0データをはじめとする移動軌跡データの継続的収集や、計算機の演算能力の向上等により、よりミクロなレベルでの分析環境も整いつつある。

一方で、交通ネットワーク分析において重要な役割を担う経路選択モデルについては、研究レベルでは様々な進展がなされてきたものの、実務的にはロジットモデルを中心としたシンプルな離散選択モデルの適用に留まっている。その最大の障壁は、経路列挙の問題であろう。すなわち、離散選択モデルが網羅的な選択肢集合からの選択行動を記述するものである一方、大規模なネットワークでは、経路選択肢集合の列挙は事実上不可能である。この問題を解決するアプローチとしては、ノードごとに接続する下流リンクを選択すると仮定した逐次選択型モデルの適用が挙げられるが、近視眼的であるがゆえに周回経路を生成したり、リアルタイム情報に対する反応を

過大に評価してしまうなどの問題が生じてしまうため、根本的な解決には至っていない。

このような状況を打破するモデルとして、Fosgerau et al.¹⁾はRecursive Logitモデル（以下、標準RLモデル）を提案している。後述するように、同モデルは逐次選択型の構造を採用することで経路列挙の問題を回避しつつ、合成効用項により目的地までの効用も明示的に考慮することができるという点で極めて実用性の高いモデルであると言える。加えて、同モデルから計算される経路選択確率は、無数の経路が含まれる選択肢集合を仮定した場合の多項ロジットモデルと数学的に等価であるという理論的特長も有している。そのため、近年では、さらなる理論的拡張²⁾³⁾が図られると共に、大規模ネットワークへの適用⁴⁾⁵⁾もなされつつある。

さて、標準RLモデルは、次章で述べるように、効用のリンク加法性を前提に、効用関数は下流リンクの即時効用と、その下流端から目的地までの合成効用の和として表している。また、前者については確定項と誤差項によって構成され、後者については、ログサム型の確定項によって表されている。しかしながら、即時効用に誤差が生ずる場合、リンク加法性の仮定に基づけば、合成効用にも誤差が生ずるものと考えられる。加えて、その誤差は、必然的に目的地までの距離等が長いほど分散が

大きくなるような異分散性を示すものと考えられる。他方、標準RLモデルの推定においては、選択経路を構成する各リンクの選択確率の積を尤度とするため、本来は異分散性を示す誤差に対して、等分散性を暗に仮定して推定しているものとみなすことができる。本稿では、この問題を理論的に考察すると共に、仮想的格子状ネットワークにおいて、リンクごとに誤差効用が生ずるような状況を想定して生成した人工的な経路選択データを用い、同モデルを適用した場合の問題点について検証した結果を報告する。

2. 標準RLモデルの概要と異分散性の問題

(1) 標準 RL モデルの概要

ここでは、Fosgerau et al. (2013)⁹⁾に従い、主に標準 RL モデルの効用関数の定式化と経路選択確率に着目して、その概要を示す。ベルマン方程式の解法や、無数の経路が含まれる選択肢集合を仮定した場合の多項ロジットモデルとの等価性の証明等については、原著論文を参照されたい。

標準 RL モデルでは、図-1 に示す有方向ネットワークを想定し、リンク k から下流に位置するリンク a に移動した際に得られるリンク効用 $u(a|k)$ を次式のように定義する。

$$u(a|k) = v(a|k) + V(a) + \mu\varepsilon(a) \quad (1)$$

ここで、

$$V(k) = E \left[\max_{a \in A(k)} \{v(a|k) + V(a) + \mu\varepsilon(a)\} \right] \quad (2)$$

である。 $v(a|k)$ は、リンク k からリンク a に移動した際に得られる効用、すなわち即時効用の確定成分、 $V(a)$ はリンク a の下流端から目的地まで移動した際に得られる最大効用の期待値、 $\varepsilon(a)$ はスケールパラメータ μ が 1 である IID.ガンベル分布に従う即時効用の誤差項、 $A(k)$ はリンク k の下流端に接続するリンクの集合、である。なお、式(2)において、誤差項に IID.ガンベル分布を仮定しているため、 $V(k)$ は次式のようにログサム変数で表すことができる。

$$V(k) = \begin{cases} \mu \ln \sum_{a \in A(k)} \delta(a|k) \exp \left[\frac{1}{\mu} \{v(a|k) + V(a)\} \right] & \forall k \in A \\ 0 & k = d \end{cases} \quad (3)$$

ここに、 $\delta(a|k)$ はリンク a がリンク k の下流端に接続するリンクである場合は 1、そうでない場合は 0 をとるダ

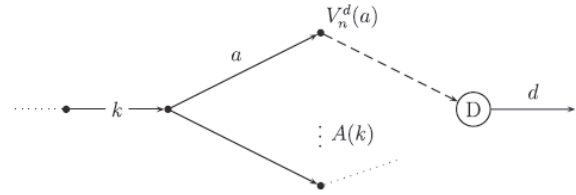


図-1 標準 RL モデルにおけるネットワーク設定

ミー変数、 A は対象とするネットワークを構成する全リンクの集合、 d は目的地の下流端に接続する便宜上設定されたダミーリンクである。

式(1)~式(3)より、誤差項に IID.ガンベル分布を仮定しているため、リンク k からリンク a に移動する確率は、以下のようにロジットモデルで表すことができる。

$$P(a|k) = \frac{\exp \left[\frac{1}{\mu} \{v(a|k) + V(a)\} \right]}{\sum_{a' \in A(k)} \exp \left[\frac{1}{\mu} \{v(a'|k) + V(a')\} \right]} \quad (4)$$

なお、選択経路を構成するリンクの集合を $\sigma = \{k_0, k_1, \dots, k_i, \dots, k_I\}$ とすると、経路選択確率は、式(4)で表される各リンクの選択確率の積として次式のように表すことができる。

$$P(\sigma) = \prod_{i=0}^{I-1} P(k_{i+1}|k_i) \quad (5)$$

紙面の都合上、説明は割愛するが、式(5)で表される経路選択確率は、無数の経路が含まれる選択肢集合を仮定した場合の多項ロジットモデルと数学的に等価であることが原著論文にて証明されている。加えて、モデル推定に際しては、式(4)および式(5)で表されるロジット型選択確率の積を尤度として、最尤推定法により即時効用の確定項に含まれる未知パラメータが推定できる。ゆえに、標準 RL モデルは、ランダム効用最大化に則しつつ、逐次選択型モデルの操作性をも兼ね備えた、極めて実用性の高いモデルであると言える。

(2) 標準 RL モデルの誤差項の問題

さて、1 章で述べた異分散性の問題を考えるために、今、式(1)のリンク a の下流端が目的地ノード D である場合を考える。この場合、目的地ノードの下流に位置するリンクは、ダミーリンク d のみであるため、定義上、式(1)は合成効用項を含まない形で次式のように表される。

$$u(a|k) = v(a|k) + \mu\varepsilon(a) \quad (6)$$

次に、リンク a の上流側に接続するリンク k の効用を考えると、式(6)にはスケールパラメータ μ の IID.ガンベ

ル分布に従う誤差項が含まれるため、リンク k の下流端から目的地まで移動した際に得られる効用の最大値もまたスケールパラメータ μ の IID.ガンベル分布に従う確率変数となる。すなわち、リンク k のさらに上流側に接続するリンクを m とおけば、リンク m からリンク k に移動した際に得られる効用は次式のように表すことができる。

$$u(k|m) = v(k|m) + U(k) + \mu\varepsilon(k) \quad (7)$$

ここで、

$$U(k) = \max_{a \in A(k)} \{v(a|k) + \mu\varepsilon(a)\} = V(k) + \mu\varepsilon'(a) \quad (8)$$

であり、 $\varepsilon'(a)$ はスケールパラメータが 1 の IID.ガンベル分布に従う誤差項である。式(8)を式(7)に代入すると共に、式(1)との比較のため、式(7)の k を a 、 m を k に置き換えると次式ようになる。

$$u(a|k) = v(a|k) + V(a) + \mu\varepsilon(a) + \mu\varepsilon'(a) \quad (9)$$

式(9)の右辺における誤差成分 $\mu\varepsilon(a) + \mu\varepsilon'(a)$ は、スケールパラメータこそ同一ではあるが、ガンベル分布に従う確率変数の和はガンベル分布には従わない。従って、リンク選択確率はもはやロジット型確率で表すことはできない。加えて、誤差成分 $\mu\varepsilon(a) + \mu\varepsilon'(a)$ の分散は、理論上、 $\mu\varepsilon(a)$ の分散よりも大きくなるという異分散性を有することになる。

ここで、仮に式(9)が正しいものとする、式(1)で表される標準 RL モデルでは、あたかも多段階ネステッド・ロジットモデル⁹⁾と同様の仮定をおいているものと解釈できる。すなわち、 $\mu\varepsilon'(a)$ 自体はガンベル分布には従わないものの、 $\mu\varepsilon'(a)$ は $\mu\varepsilon(a) + \mu\varepsilon'(a)$ がガンベル分布になるような分布に従うとの仮定をおくことで経路選択確率がロジット型確率の積として表すことが保たれる。ただし、その場合でも、上流側のリンク効用については分散が大きくなるような異分散性を考慮する必要がある。このような問題にも関わらず、出発地側のリンク効用と目的地側のリンク効用の等分散性仮定した標準 RL モデルを適用した場合、平均的な誤差分散に基づく推定結果が得られるものと推察される。ここで、実際のモデル推定においては、即時効用の確定項に含まれる未知パラメータの推定結果にスケールパラメータが乗じられた形で推定される。ゆえに、等分散性を仮定してモデルを推定した場合には、特に上流側では確定項の影響が本来よりも小さくなるように、逆に言えば、誤差項が相対的に卓越するように推定されるものと予想される。従って、推

定モデルを用いて選択経路をシミュレートした場合、特に出発地付近で周回経路が生成されがちになる恐れが懸念される。

3. 人工経路選択データによる異分散性の問題の検証

前章で考察したような問題が生ずるのかを検証するために、リンク効用に異分散性が生ずるような状況を設定した上で、人工的な経路選択データを生成し、モデルの推定精度の検証を行った。

(1) 仮想ネットワークの設定

本研究では、図-2 のような $n \times n$ の仮想の格子状ネットワークを構成して分析を行った。簡略化のため、1OD のケースのみを対象とし、即時効用としては所要時間のみを考慮することとした。ここで、格子状のネットワークでは、OD 間距離は全ての経路で同一になってしまうため、経路間での所要時間が異なるよう、主要幹線道路(太線)、補助幹線道路(破線)、生活道路(実線)の3種類の道路を模して異なる走行速度を設定し、リンク所要時間を生成した。各道路は起点ノードから水平方向に補助幹線道路、主要幹線道路、生活道路の順に、垂直方向には補助幹線道路、生活道路の順に交互に、規則的に配置した。なお、リンク所要時間の誤差については、道路種別によらず、全ての道路リンクにおいて、同一のスケールパラメータをもつガンベル分布に従って生成した。これにより、所要時間の誤差は目的地からの距離が大きくなるほど分散が大きくなるように設定した。ただし、誤差成分の相対的な大きさの影響を検証するため、スケールパラメータは、0.1~1.9 まで 0.2 刻みで設定した。なお、本ネットワークを構成するリンク数は $2n(n+1)$ 、OD 間の経路選択枝数は ${}_{2n}C_n$ であり、OD 間距離の影響

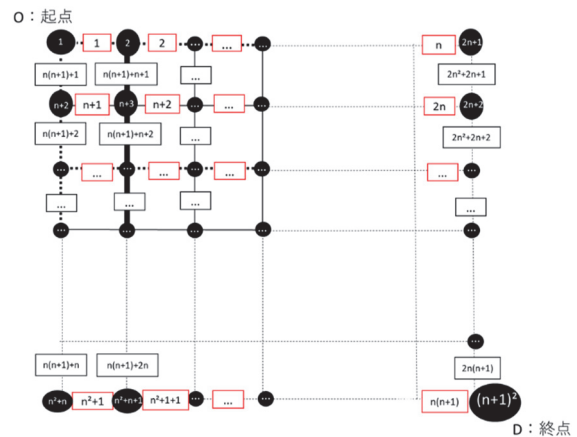


図-2 格子状ネットワークの設定

を把握するため、 n を 1~25 まで 2 刻みで変化させた。

(2) 経路選択データの作成

所要時間のパラメータ β の真値を -1 とおいた上で、誤差を含むリンク所要時間を全てのリンクについて算出し、次いでダイクストラ法にて最短経路を特定し、それを経路選択データとした。なお、リンク所要時間の誤差項については、乱数の影響を排除するため 1000 回生成した。ゆえに、各設定ごとに、リンク所要時間の誤差が異なる 1000 人分の経路選択データを生成した。

(3) 標準 RL モデルの推定精度の検証

生成した経路選択データと、各リンクの平均所要時間のみが与えられているものとし、標準 RL モデルにより所要時間のパラメータを推定した。

まず、図-3 に、スケールパラメータ μ が 1、所要時間のパラメータ β の真値を -1 として生成した経路選択データに対して、標準 RL モデルを用いてパラメータを推定した結果を示す。図からわかるように、真値を -1 と設定しているため、所要時間のパラメータは常に過小推定される結果となった。特に $n=1$ の段階で既にパラメータは -0.59 と大きく過小推定されている。これは、格子状ネットワークの形状により、 $n=1$ のときには経路を構成するリンク数が 2 となっているため、起点における所要時間の分散が個々のリンク所要時間の誤差の 2 倍程度になっていることによるものと考えられる。なお、以降は n が大きくなってもそれほど推定値に変化がないように見受けられるが、ほぼ単調に 0 に近づく傾向にあることがわかる。これは前章で考察したように、OD 間距離が増大し、起点側と終点側での異分散性が大きくなるにつれ、確定項の影響が相対的に小さくなるように推定される傾向にあることを示唆している。

次に、スケールパラメータ μ の値を変化させた場合のモデル推定結果を図-4 に示す。ここで、 μ が大きくなると、誤差項の影響が相対的に大きくなるため、経路選択が平均所要時間によらなくなることを踏まえると、誤差の影響が小さければ過小推定の水準は低い、一方でネットワークが大規模になるにつれて過小推定の度合いが大きくなる傾向にあることが明らかとなった。

4. おわりに

本研究では、効用のリンク加法性を仮定した場合、目的地までの距離等が長いほど経路効用の誤差成分の分散が大きくなるような異分散性が生ずるのに対し、標準 RL モデルは、それを無視して等分散性を仮定しているとの問題提起を行うと共に、異分散性が存在する状況下での経路選択データを人工的に生成した上で、標準 RL

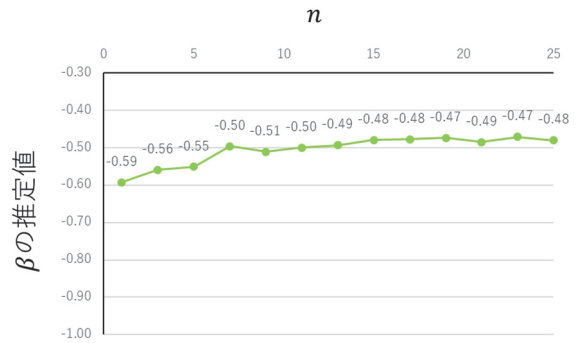


図-3 $\mu = 1, \beta_t = -1$ のデータに対する推定結果

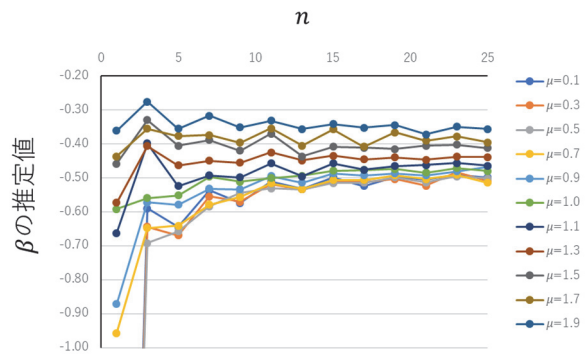


図-4 μ を変化させたデータに対する推定結果

モデルの推定精度を検証した。その結果、効用パラメータが過小推定され、特に出発地付近で周回経路が生成されてしまうなどの問題が生ずる可能性があることを確認した。異分散性の問題を緩和する方法としては、1) それを明示的に考慮できるようなガンベル分布以外の誤差分布を仮定したモデルを構築する、2) ガンベル分布の仮定はそのままに、ミックスト・ロジットモデル η のような合成分布を採用する、等のアプローチが考えられるが、いずれもクロズド・フォームにならないなどして、ベルマン方程式の求解が困難になることが容易に想像される。よって、簡便かつ現実的な対応としては、効用パラメータには誤差分散の影響が含まれることを踏まえ、効用パラメータを距離等に応じて構造化するなどの方法が挙げられる。今後は、それらによる問題の緩和効果等を明らかにすると共に、様々な OD を含むような多様な状況下で同様の検証を行う必要がある。

参考文献

- 1) Fosgerau, M., Frejinger, E. and Karlstrom, A.: Link based network route choice model with unrestricted choice set, *Transportation Research Part B: Methodological*, Vol.56, p.70-80, 2013.
- 2) Mai, T., Fosgerau, M. and Frejinger, E.: A nested recursive logit model for route choice analysis, *Transportation Research Part B: Methodological*, Vol.75, pp.100-112, 2015.
- 3) 大山雄己, 羽藤英二: 一般化 RL モデルを用いた災

- 害時の経路選択行動分析, 交通工学論文集, 第 3 巻, 第 5 号, pp.1-10, 2017.
- 4) 金子法子, 福田大輔: リンクベース再帰的ロジットモデルの首都圏道路網における経路選択分析への適用第 54 回土木計画学研究発表会・講演集, pp.1399-1404, 2016.
 - 5) 力石真, 田名部淳, 大口敬: プローブデータを用いた貨物車経路選択行動のモデル分析, 第 54 回土木計画学研究発表会・講演集, pp.1411-1416, 2016.
 - 6) Ben-Akiva, M. and Lerman, S.: *Discrete Choice Analysis: Theory and Application to Travel Demand*, The MIT Press, 1985.
 - 7) McFadden, D. and Train, K.: Mixed MNL models for discrete response, *Journal of Applied Econometrics*, Vol.15, p p.447-470, 2000.

(2020. 10. 2 受付)