観測誤差分散のリンク固有性に着目した 経路選択モデルの構造推定

大山 雄己¹・羽藤 英二²

¹正会員 スイス連邦工科大学ローザンヌ校 (GC B2 385, Station 18, CH-1015 Lausanne, Switzerland)
E-mail: yuki.oyama@epfl.ch

²正会員 東京大学教授 大学院工学系研究科社会基盤専攻 (〒113-8656 東京都文京区本郷 7-3-1)
E-mail: hato@bin.t.u-tokyo.ac.jp

経路選択モデルの推定には経路情報が必要であり、GPS 技術により自動観測が期待されるものの、GPS 位置情報には機器・空間依存の観測誤差を持ち、ネットワーク上で分散不均一であるという観測不確実性が存在する。本研究は、従前の経路選択分析が、観測モデルに必要な初期パラメータの設定によって推定バイアスを生んでいた点とそのメカニズムに着目する。本研究で提案する「逐次リンク観測モデル」は、GPS の観測誤差分散パラメータをリンク固有変数として推定可能とし、また事前情報として経路選択モデルを導入することにより行動メカニズムを反映した経路の特定を行う。さらに経路選択モデルのパラメータに関する「構造推定」手法を提案し、事前情報に含まれる不確実性を軽減する。最後に数値実験を通じて、本手法の有用性を確認する。

Key Words: route choice model, route measurement model, GPS data, sequential link measurement, structural estimation, pedestrian

1. はじめに

離散選択モデルのパラメータ推定にはモデルに対応する実際の行動データが必要である。経路選択モデルにおいては、ネットワーク上のある出発地-目的地間を結ぶリンクの系列としての経路情報が必要となる。従前では、電子メールや電話による聞き取り調査が行われていたものの、旅行者が正確に経路情報を記憶していることは稀であった。近年では GPS (Global Positioning System) による位置情報の取得が行われるようになり、経路情報の観測を自動的に行うことが期待される。Probe car や Connected vehicle によるリアルタイムモニタリングのみならず、Probe Person 調査¹⁾では市街地の歩行者回遊経路を観測することが可能となり、観測の時空間解像度が高まってきている。位置情報の観測精度も進歩を続けているものの、特に歩行者ネットワークにおいて未だ大きな観測の不確実性が存在する。まず、ネットワークが観測誤差の大きさに対して密であることから、候補となる経路数が増加し、真の状態の特定が困難になる。またリンクごとの空間特性が大きく異なり、観測誤差に影響を与える。例えば広幅員街路や空地沿いの街路では安定した位置情報が記録される一方で、路地やアーケード商店街では観測精度が悪化する。つまりネットワーク上では観測誤差の分散が不均一であると考えられる。しかし経路を確率的に特定する従前の経路観測モデルでは、モデルパラメータである観測誤

差分散はネットワーク上で均一かつ所与の値として扱われてきた^{2),3)}。経路の観測が不確実であれば、その情報を用いて推定される経路選択モデルは必然的にバイアスを伴うことになる。経路観測の不確実性を考慮した確率的な経路観測モデルが近年提案されているものの、モデルに必要な初期パラメータの設定による推定バイアス増幅の可能性があった。観測補正のための事前情報として経路選択モデルを導入するベイジアンアプローチ^{4),5)}では、導入したモデルと推定されるモデルとの一貫性が保たれないことで、誤った方向へ補正がなされてしまう課題が存在する。本研究では、こうした経路観測モデルと経路選択モデルの相互関係に依存した推定バイアス増幅のメカニズムに着目し、バイアス軽減のための推定手法を新たに提案する。

2. 既往研究の整理

本章では、観測の不確実性を持つ経路選択モデル推定のフレームワークについてレビューを行う。はじめに GPS データを用いた経路観測手法を概観し、その後経路観測モデルと経路選択モデルの同時推定手法について言及を行う。

(1) 経路観測モデル

GPS データは経路選択モデルで用いられる記法（ネットワーク）と本来的にデータフォーマットが異なり、ま

た機器・空間依存の観測誤差を持つため、基本的に事前処理なしでは交通行動分析において意味をなさない。こうした理由から、GPS データから利用者の真の状態（経路）を特定するための様々な手法がここ 20 年ほどの間に多く提案されてきている。GPS の位置情報を交通ネットワーク上の点とリンクさせるマップマッチング手法はその代表的な技術であり、幾何学的手法⁶⁾、位相幾何学的手法^{7),8),9)}、確率的手法^{10),11),12)} のように大別される。しかし上述した論文を含むほとんどのマップマッチング手法はオンラインのナビゲーションシステムへの応用を目的としており、観測点ごとの逐次的なマッチングが行われる。そのため、リンクの接続関係が満たされないなど、行動論的に意味をなさない経路が生成されることも多い。一方で交通行動分析（経路選択モデル）の枠組みにおいては、リアルタイムな位置情報よりも寧ろ、リンクの系列としての経路情報が必要となる³⁾。対して、Pyo et al.¹³⁾ や Bierlaire et al.³⁾ の確率的経路観測モデルは、トリップに含まれる全ての点列に対する経路の尤もらしさを評価する。また Chen and Bierlaire¹⁴⁾ は加速度の情報を合わせて用いることで、交通機関判別を含むマルチモードネットワーク上の経路観測モデルを提案している。

さらに近年では、ベイズアンアプローチを用いた発展的な手法^{4),5),14)} が提案されている。ベイズアンアプローチは、確率的経路観測モデルによる経路の観測確率を、事前情報である経路選択モデルによって補正する手法である。位置精度が悪い、あるいはネットワークが密な状況においては観測確率の不確実性が高く、候補経路間で確率の差が小さくなり、経路の特定に十分な情報が得られないことも少なくない⁵⁾。そうした状況において、行動的な情報を用いて補正を行うベイズアンアプローチは有効となる。しかし、逆に言えば事前情報が不確実である場合には、観測が誤った方向へ補正されてしまうことに他ならない。つまり理論上では、事前情報として用いられる経路選択モデルと、特定された経路を用いて推定される経路選択モデルとの一致性が満たされない限り観測・推定のプロセスにバイアスが含まれてしまうことになる。

(2) 経路観測モデル・経路選択モデルの同時推定

Bierlaire and Frejinger²⁾ は観測の曖昧さを残したまま経路選択モデルを推定するフレームワークを提案した。経路選択モデルのパラメータは以下の尤度を最大化させるように推定される。

$$p(\hat{m}; \sigma, \theta) = \sum_{r \in \mathcal{R}} p(\hat{m}|r; \sigma) p(r|\mathcal{R}; \theta), \quad (1)$$

ここで、 r は経路、 $\mathcal{R}$ はその候補集合である。 $p(\hat{m}|r; \sigma)$ は経路の観測方程式であり、「 r が真の経路である場合

にデータ列 $\hat{m}$ が観測される確率」を与える。 σ はそのパラメータであり、主にデータの観測誤差分散として仮定される。 $p(r|\mathcal{R}; \theta)$ は経路選択モデルであり、「選択肢集合 $\mathcal{R}$ から経路 r が選択される確率」を与える。 θ はそのパラメータである。式 (1) において、経路は一意に特定されず潜在変数として扱われる。つまり、候補となる全ての経路について $p(\hat{m}|r; \sigma)$, $p(r|\mathcal{R}; \theta)$ を評価し、それを足し合わせたものを最大化することでパラメータを推定する。しかし、経路を確率的に評価するためには、確率を割り当てる対象となる経路の集合 $\mathcal{R}$ を定義しなくてはならない。経路選択モデルの選択肢集合の定義が困難であることはよく知られた事実であるが、経路観測モデルについても同様の課題が生じる。Bierlaire and Frejinger²⁾ は経路観測モデルの経路候補集合生成のために、データが関連づけられる空間的範囲である DDR (Domain of Data Relevance) の概念を提唱している。GPS データの場合であれば、各観測点（位置座標）から一定の範囲（DDR）内に存在するリンクのみが経路の構成要素となりうると仮定することで、経路候補集合を外生的に定めることが可能となる。ただし、DDR の範囲は候補集合のサイズを左右するため、計算の正確性（真の経路の抜け落ち回避）と計算負荷のトレードオフの関係が生じることが指摘されている^{2),3)}。また、DDR に基づく経路観測モデルはこれまで言及されていない課題を残している。DDR を定義するためには観測点の誤差パラメータである σ の値が既知である必要がある。マッチング前の段階では各観測点の真の位置・空間情報を知ることができないため、 σ は「ネットワーク上で均一」の値として分析者に設定される「所与」のパラメータとみなされる。上述したように実際のネットワークでは観測誤差に分散不均一性が存在すると考えられるため、等分散性を仮定することはすでにモデルが不確実性を含むことを意味する。 σ が不確実である場合には、候補集合及び経路観測確率 $p(\hat{m}|r; \sigma)$ にバイアスがかかることに他ならない。

つまり、同時推定の実行には潜在変数である経路の候補集合の情報が必要であり、計算負荷の増加及び推定バイアスを生じさせていた。式 (1) の最大化は理論的に明快な問題である（詳しくは 7 章 (2) において議論する）ものの、本論文では計算の簡便化と推定バイアス除去を目的とし、別のアプローチをとる。

(3) 本研究の新規性

以上見たように、従前の経路選択モデル推定の枠組みは、観測モデルにおいて必要なパラメータの初期設定（ベイズアンアプローチでは事前情報の行動パラメータ θ 、同時推定の枠組みでは観測誤差分散 σ の仮定）によって、推定のプロセスでバイアスを増幅させてしまっ

ていた。本研究は、逐次リンク選択行動を観測・行動モデル共通の基本単位としたベイジアンアプローチに基づく手法、そして事前情報と推定される経路選択モデルとの一貫性を担保する推定手法を提案する。前者は、観測誤差分散を所与・均一の値とする必要を回避することに加え、リンク固有変数として推定を可能とすることで、経路観測確率を持つ不確実性を解消する。後者は、事前情報として用いられる経路選択モデルのバイアスを取り除くことにより、誤った方向へ観測が補正されることを回避する。つまり、経路の観測・推定のプロセスにおいていずれのパラメータも所与とせずにモデルを通じて推定可能とすることによって、従前の枠組みが抱えていた推定バイアスを軽減する点が、本研究の既往研究にはない新規性であり、貢献である。

3. 研究のフレームワーク

本研究の目的は、観測誤差を持つ GPS データ列を用いて、バイアスを軽減した形で経路選択モデルを推定することである。本章では研究のフレームワークとして、経路選択モデル及び経路観測モデルの定義を行う。

(1) ネットワーク

ネットワークを有向グラフ $G = (N, \mathcal{A})$ として定義する。 $N, \mathcal{A}$ はそれぞれ、ノード及びリンクの集合である。ノード $i \in N$ は位置座標 $x_i = \{x_{i1}, x_{i2}\}$ を属性として持つ。リンク $a \in \mathcal{A}$ は直線であり、始端・終端ノード (u_a, d_a) 及び属性ベクトル y_a によって特徴付けられる。また、 $\mathcal{A}(a)$ をリンク a から接続する（推移可能な）リンクの集合、 $\delta(a'|a)$ を $a' \in \mathcal{A}(a)$ のときに 1、それ以外で 0 をとる指示変数として定義する。つまり、 $\mathcal{A}(a) = \{a' \in \mathcal{A} | \delta(a'|a) = 1\}$ である。

(2) 経路選択モデル

旅行者がネットワーク $G = (N, \mathcal{A})$ 上のある地点 (Origin) から別のある地点 (Destination) までを連続的に移動することを、真の状態であると仮定する。本研究ではさらに、旅行者は O-D 間でプレトリップ的に経路を選択するのではなく、交差点ごとの逐次的なリンク選択行動を行っているものと仮定する。逐次リンク選択モデルにおいて、経路はリンク選択結果の連なりによって出力される。個人は幅を持つ時点 t ごとに状態リンク $a_t \in \mathcal{A}$ をとり、 a_t に基づいて自身の効用を最大化させるよう、次時点の状態リンク a_{t+1} を $\mathcal{A}(a_t)$ の中から選択する。このとき、リンク選択確率は次のように表される。

$$p(a_{t+1}|a_t) = P\left(\tilde{u}(a_{t+1}|a_t) = \max_{a'_{t+1} \in \mathcal{A}(a_t)} \tilde{u}(a'_{t+1}|a_t)\right), \quad (2)$$

where

$$\tilde{u}(a_{t+1}|a_t) = u(a_{t+1}|a_t) + \varepsilon(a_{t+1}). \quad (3)$$

ここで、リンク a_t, a_{t+1} 間を推移する際の効用関数の確定項はリンクの属性ベクトル $y_{a_t}, y_{a_{t+1}}$ 及びパラメータ θ の関数 $u(a_{t+1}|a_t) = u(y_{a_t}, y_{a_{t+1}}, \theta)$ として表される。効用関数の誤差項 $\varepsilon(a_{t+1})$ は i.i.d. ガンベル分布に従うと仮定する。マルコフ性を仮定すれば、経路選択確率はリンク選択確率式 (2) の積として表される。

$$P(r = [a_0, \dots, a_T]) = \prod_{t=0}^{T-1} p(a_{t+1}|a_t; \theta) \quad (4)$$

ここで、 a_0, a_T はそれぞれ出発地リンク、最終到着地リンクである。モデル推定を扱う逐次リンク選択モデルとしては、近年 Fosgerau et al.¹⁵⁾ が Recursive Logit (RL) モデルを提案している。RL モデルは経路の列挙を必要とせずに、ユニバーサルチョイスセットを選択肢集合とする MNL 型の経路選択モデルと等価な結果を導くことが可能である。Mai et al.¹⁶⁾ や Mai¹⁷⁾ は RL モデルを GEV モデル¹⁸⁾ の枠組みに展開し、経路の相関構造を考慮可能にした。また、大山・羽藤¹⁹⁾ は動的離散選択モデルにおける時間割引率のパラメータに着目し、近視眼的な経路選択行動を同一のフレームワークで記述可能なモデルへ発展している。

(3) 経路観測モデル

本研究では移動通信体から得られる GPS データの観測点列を用いて、経路情報の観測及び経路選択モデルの推定を行うことを考える。観測値を $\hat{m}$ のように上付きのハット記号を用いて表すこととし、観測を $\hat{m} = (\hat{x}, \hat{t})$ と定義する。つまり各観測は位置座標 $\hat{x} = \{\hat{x}_1, \hat{x}_2\}$ (一般的に緯度経度 $\{\hat{x}_{lat}, \hat{x}_{lon}\}$) 及び時刻 $\hat{t}$ の情報を持つ。各トリップ z に対して、時系列順に並ぶ N_z 個の観測点列 $\hat{\mathbf{m}} = (\hat{m}_1, \dots, \hat{m}_n, \dots, \hat{m}_{N_z})$ が記録されている。時刻は正確に記録される一方、位置座標 $\hat{x}$ は常に観測誤差を持っており、真の位置からの距離が確率分布 $p(\hat{x}|x; \sigma)$ に従うと仮定する。 σ は確率分布の分散パラメータである。2章で見たように、従前研究では σ は所与かつネットワーク上で均一の値として扱われてきた。GPS データの観測誤差は大気や受信機雑音、遮蔽物、多重反射等のさまざまな要因から発生する各測位点に固有の変数であり、特に建築物による影響が大きいことも知られている²⁰⁾。本研究では、多くの場合において交差点で区切られた街路内で空間特性（建築物の高さや街路幅員等）が似通っていることに着目し、観測誤差の分散をリンク固有の変数 $\sigma = \sigma_a$ とみなす。つまり各リンク内では等分散性、リンク間では分散不均一性を仮定してモデルによる推定を行うことで、従前ではネットワーク上で均一の値として扱われてきた観測誤差の街路空

間（リンク）に応じた差異を考慮することを試みる。

4. 逐次リンク観測モデル

本章では、逐次リンク観測モデルを提案する。上述したように、パスベースの確率的経路観測モデル³⁾は、トリップに含まれるすべての観測点列に対する各候補経路の尤もらしさを評価する。この時、各観測がどのリンク上の点を真の状態として発生されたかどうかは未知であり、計算量の関係からすべてのリンクを考慮することも不可能である。そこでDDRのように各観測点に対応づけられる空間的範囲を、誤差分散のパラメータによってあらかじめ定義することで、候補リンクとその組み合わせである候補経路の限定を行う。

一方本研究では、逐次的なリンクの特定を行うことで、候補リンク集合がモデルによって与えられることに着目する。つまり、ある時点のリンクが a に特定されれば、次時点のリンクは $\mathcal{A}(a)$ に含まれるうちのいずれかとみなすことができる。観測点と候補リンクとの対応関係があらかじめ明らかとなることで、その空間的關係からリンク固有変数としての観測誤差分散パラメータが推定可能となる。また、逐次リンク観測モデルは、ベイジアンアプローチに基づく、逐次リンク選択モデルを採用することで、経路選択肢集合の列挙を必要とせずに行動メカニズムを考慮した観測の補正を行うことができる。以下、モデルの説明を行う。

(1) データの時間構造化

逐次リンク観測モデルは、逐次リンク選択行動を観測・行動モデル共通の基本単位とする。両者を同一のフレームワークに落とし込むため、観測点列の操作をはじめに行う。

離散時間 $(1, \dots, t, \dots, T_z)$ 、 $t^- \leq t \leq t^+$ を定義する。各時点 t の時間幅 $\bar{t} = t^+ - t^-$ は一定として考える。なお、 t^+, t^- はトリップ開始からの経過時間（したがって $\hat{t}_{N_z} - \hat{t}_1 = T_z^+$ ）である。 $\bar{t}$ はネットワークごとに定義する変数であり、基本的にはリンクの接続条件を満たすよう設定される。時点 t 内に含まれる $(t^- \leq \hat{t}_j \leq t^+)$ を満たす観測点列を、 $\hat{\mathbf{m}}^t = (\hat{m}_1^t, \dots, \hat{m}_{j_t}^t, \dots, \hat{m}_{J_t}^t)$ と定義する。 J_t は $\hat{\mathbf{m}}^t$ に含まれる観測点の数であり、 $\sum_{t=1}^T J_t = N_z$ である。1つのトリップに関する観測点列 $\hat{\mathbf{m}}$ は、 $(\hat{\mathbf{m}}^1, \dots, \hat{\mathbf{m}}^t, \dots, \hat{\mathbf{m}}^T)$ のように時間ごとの観測ベクトルの列として分解される。本研究ではこのベクトル $\hat{\mathbf{m}}^t$ を単位として、3章(2)における各時点 t の状態リンク a_t を観測する。つまり、同時点内に含まれるすべての観測点が、同じリンクを真の状態とするという仮定に基づき、 $t = 1$ から時間前進方向に逐次的なリンク観測を行う。

(2) 逐次リンク観測

観測点列 $\hat{\mathbf{m}}^t$ を用いて、すべての時点 t における真のリンク a_t をリンク尤度に基づいて特定する。なお、候補リンクは前時点の状態 a_{t-1} に接続するリンクの集合 $\mathcal{A}(a_{t-1})$ として定義されている。リンク尤度は前時点の状態 a_{t-1} 、観測点列 $\hat{\mathbf{m}}^t$ として、ベイズの法則から以下のように与えられる。

$$p(a_t | \hat{\mathbf{m}}^t, a_{t-1}) \propto p(\hat{\mathbf{m}}^t | a_t; \sigma_{a_t}) p(a_t | a_{t-1}; \theta), \quad (5)$$

ここで $p(\hat{\mathbf{m}}^t | a_t; \sigma_{a_t})$ は真の状態が a_t である場合にデータ列 $\hat{\mathbf{m}}^t$ が観測される確率を与える（リンクの）観測方程式である。 $p(a_t | a_{t-1}; \theta)$ は経路選択モデルであり、前章で定義したように本研究では逐次リンク選択モデルを用いる。 σ_{a_t}, θ は未知のパラメータである。 a_t は離散化された変数であるため、式(5)は以下のように正規化される。

$$p(a_t | \hat{\mathbf{m}}^t, a_{t-1}) = \frac{p(\hat{\mathbf{m}}^t | a_t; \sigma_{a_t}) p(a_t | a_{t-1}; \theta)}{\sum_{a_t \in \mathcal{A}(a_{t-1})} p(\hat{\mathbf{m}}^t | a_t; \sigma_{a_t}) p(a_t | a_{t-1}; \theta)}. \quad (6)$$

リンク尤度に基づき、逐次的なリンクの特定を行う。つまり尤度が最大のリンクを時点 t の真の状態リンクとして特定する。

$$a_t = \arg \max_{a_t \in \mathcal{A}(a_{t-1})} p(a_t | \hat{\mathbf{m}}^t, a_{t-1}). \quad (7)$$

この操作を $t = 1$ から $t = T$ まで繰り返すことによって、最終的にリンクの系列としての経路情報 $r = [a_1, \dots, a_t, \dots, a_T]$ を得る。

(3) 観測方程式

時刻には観測誤差がないという仮定より、真の状態がリンク a_t であるときに $\hat{\mathbf{m}}^t = (\hat{m}_1^t, \dots, \hat{m}_{j_t}^t)$ が観測される確率は以下のように表される。

$$p(\hat{m}_1^t, \dots, \hat{m}_{j_t}^t | a_t; \sigma_{a_t}) = p(\hat{x}_1^t, \dots, \hat{x}_{j_t}^t | a_t; \sigma_{a_t}). \quad (8)$$

また、旅行者は同一リンク上をホワイトノイズを持つ等速直線運動し、リンク上の位置はリンクに到達してからの経過時間によってのみに決まる（式(10)～(12)）と仮定すれば、式(8)は以下のように分解される。

$$\begin{aligned} p(\hat{x}_1^t, \dots, \hat{x}_{j_t}^t | a_t; \sigma_{a_t}) &= \prod_{j=1}^{J_t} p(\hat{x}_j^t | a_t; \sigma_{a_t}) \\ &= \prod_{j=1}^{J_t} \int_{x_j \in a_t} p(\hat{x}_j^t | x_j, a_t; \sigma_{a_t}) p(x_j | a_t) dx_j \end{aligned} \quad (9)$$

where,

$$x_j = l x_{d_{a_t}} + (1 - l) x_{u_{a_t}}, \quad (10)$$

$$l = \frac{\hat{t}_j^t - t^-}{t^+ - t^-} + \eta, \quad (11)$$

$$\eta \sim N(0, \sigma_\eta^2). \quad (12)$$

式(9)は真の位置 x_j がリンク a_t 上に確率的に分布していることを考慮し、リンク上を線積分することを意味

図-1 経路観測の失敗例

している。 $p(\hat{x}_j|x_j, a_t; \sigma_{a_t})$ は GPS 位置情報の観測誤差が従う確率分布である。GPS の水平誤差は一般に独立正規分布に従うことが知られており²¹⁾、真の位置と観測点の水平距離 d はレイリー分布に従う。したがって、リンク a_t 上の点 x_j が真の位置であるときに観測 $\hat{x}_j$ が観測される確率 $p(\hat{x}_j|x_j, a_t; \sigma_{a_t})$ は以下となる。

$$p(\hat{x}_j|x_j, a_t; \sigma_{a_t}) = \frac{\|\hat{x}_j - x_j\|}{\sigma_{a_t}^2} \exp\left(-\frac{\|\hat{x}_j - x_j\|^2}{2\sigma_{a_t}^2}\right). \quad (13)$$

$\sigma_{a_t}^2$ は誤差分散パラメータであり、ネットワーク由来の誤差と機器依存の誤差から構成される^{3), 22)}。上述したように、従前の経路観測モデルでは、誤差分散を所与かつネットワーク上で均一の値としてみなすことによって、観測確率に歪みを生じさせていた。しかし本研究で提案する逐次リンク観測モデルは、リンク $a_t \in \mathcal{A}(a_{t-1})$ の固有変数 σ_{a_t} として誤差分散パラメータを評価することが可能である。 σ_{a_t} は以下の観測尤度最大化によって推定される。

$$\hat{\sigma}_{a_t} = \arg \max_{\sigma_{a_t}} p(\hat{\mathbf{m}}^t|a_{t-1}; \sigma_{a_t}), \quad (14)$$

where

$$p(\hat{\mathbf{m}}^t|a_{t-1}; \sigma_{a_t}) = \sum_{a_t \in \mathcal{A}(a_{t-1})} p(\hat{\mathbf{m}}^t|a_t; \sigma_{a_t}) p(a_t|a_{t-1}; \theta). \quad (15)$$

推定された $\hat{\sigma}_{a_t}$ を式 (6) に代入し、リンク尤度を評価する。

(4) スイッチングによるエラー回避

逐次的なリンクの観測が持つ問題として、一度の観測の失敗が後続の観測に影響を与えてしまうことが挙げられる。図-1 に一例を示す。逐次リンク観測モデルを通じて、まず $t=1$ の真の状態としてリンク $a_1=2$ が特定される。 $t=2$ におけるリンク候補集合は $\mathcal{A}(a_1=2) = \{2, 4, 5, 7\}$ である。各リンクの尤度 $p(a_2|\hat{\mathbf{m}}^2, 2)$ を計算し、尤度が最大のリンク 4 が真の状態として特定されたと

図-2 リンクスイッチングアルゴリズム

する。このとき、 $t=3$ のリンク候補集合は $\mathcal{A}(a_2=4) = \{1, 3, 4, 6\}$ である。しかし観測点列 $\hat{\mathbf{m}}^3 = (\hat{m}_1^3, \hat{m}_2^3, \hat{m}_3^3)$ に着目すれば、 $\mathcal{A}(a_2)$ に尤もらしいリンクが含まれないことが考えられる。しかし、 $t=2$ で $a_2=4$ としてリンクの特定が行われたことで、 $t=3$ の真のリンクは $\mathcal{A}(a_2)$ の中から選ばれなければならない、例ではリンク 6 が特定される。つまり、本来推定されるべき経路が点線で表される $r = [2, 7, 10]$ であったにも関わらず、 $t=2$ における観測の失敗が後続の観測に影響を与え、観測列 $(\hat{\mathbf{m}}^1, \hat{\mathbf{m}}^2, \hat{\mathbf{m}}^3)$ に対して、太線の経路 $r = [2, 4, 6]$ が推定されてしまったケースを図-1 の例は示している。

この逐次リンク観測に起因した問題に対処するため、リンクスイッチングアルゴリズム (図-2) を提案する。アルゴリズムは以下のステップに従う。

Step 1: 時点 t において、全ての候補リンク $a_t \in \mathcal{A}(a_{t-1})$ に対して式 (6) に基づくリンク尤度 $p(a_t|\hat{\mathbf{m}}^t, a_{t-1})$ を計算する。

Step 2: リンク尤度の降順に候補リンクを並び替え、 $[a_{t,1}, \dots, a_{t,k}, \dots, a_{t,|\mathcal{A}(a_{t-1})|}]$ のようにラベリングを行う ($1 \leq k \leq |\mathcal{A}(a_{t-1})|$ は順位のサフィックスである)。このとき、以下の関係式が満たされている。

$$p(a_{t,1}|\hat{\mathbf{m}}^t, a_{t-1}) \geq \dots \geq p(a_{t,k}|\hat{\mathbf{m}}^t, a_{t-1}) \geq \dots \geq p(a_{t,|\mathcal{A}(a_{t-1})|}|\hat{\mathbf{m}}^t, a_{t-1}). \quad (16)$$

Step 3: $k=1$ とする。時点 t の状態ベクトルを $a_{t,k}$ 及び

時点 $t+1$ の観測点列 $\hat{\mathbf{m}}^{t+1}$ に基づいて候補リンクの最大観測尤度 LLm_k を計算する.

$$LLm_k = \log(p(\hat{\mathbf{m}}^{t+1}|a_{t+1}; \sigma_{a_{t+1}})), \quad (17)$$

where

$$a_{t+1} = \arg \max_a p(a|\hat{\mathbf{m}}^{t+1}, a_{t,k}). \quad (18)$$

Step 4: 以下の不等式が満たされれば時点 t の状態リンクを $a_t = a_{t,k}$ として特定し、そうでない場合は Step 5 に進む.

$$\frac{LLm_k}{J_{t+1}} > \gamma, \quad (19)$$

Step 5: $k = |\mathcal{A}(a_{t-1})|$ であれば時点 t の状態リンクを $a_t = a_{t,1}$ とする. そうでない場合, $k = k+1$ として Step 3 へ戻る.

Step 3, 4 は次時点の観測尤度に基づいて現時点の観測の信頼性を評価するための操作であり, 次時点のすべての候補リンク $a_{t+1} \in \mathcal{A}(a_{t,k})$ が観測点から大きく離れている場合に, $a_{t,k}$ が棄却される. すべての $a_{t,k}$ について式 (19) が満たされなければ, 事後尤度最大の $a_{t,1}$ が採択される (Step 5). なお, アルゴリズムは 1 時点のものであり, 各時点についてこの操作を行う.

5. 構造推定

経路選択モデルにおいて, 経路観測モデルを通じて得られた経路情報は経路選択モデル推定におけるデータ・セットとして用いられる. 本章では, 両モデルの関係性に着目し, 観測・推定のプロセスで生じるバイアスを取り除くための推定フレームワークとして, 構造推定手法を提案する.

(1) 経路選択モデルの推定

経路選択モデルの推定は最尤推定法によって行われる. 本研究で扱う逐次リンク選択モデルの対数尤度関数は, パラメータベクトル θ の関数として以下のように定式化される.

$$\begin{aligned} LL(\theta) &= \log \left(\prod_z \prod_{t=2}^{T_z} p(a_t|a_{t-1}; \theta)^{\delta_{a_t}^z} \right) \\ &= \sum_z \sum_{t=2}^{T_z} \delta_{a_t}^z \log(p(a_t|a_{t-1}; \theta)), \end{aligned} \quad (20)$$

ここで, z はトリップ (個人) の識別子であり, $\delta_{a_t}^z$ はトリップ z の経路の t 番目リンクが a_t であれば 1, それ以外で 0 となる指示変数である. $p(a_t|a_{t-1}; \theta)$ は逐次リンク選択モデルであり, 逐次リンク観測モデルにおける事前情報としても同様のモデルが用いられている. θ は対数尤度関数を最大化させるパラメータとして推定

される.

$$\tilde{\theta} = \arg \max_{\theta} LL(\theta). \quad (21)$$

ここで, 動的計画法の再帰式系に従う逐次リンク選択モデル^{15),16),17),19)} の場合は通常的最尤推定法を適用できず, ベルマン方程式の求解を明示的に含んだ NFXP アルゴリズム²³⁾, NPL アルゴリズム²⁴⁾ などが用いられる.

(2) 構造推定のアルゴリズム

ベイジアンアプローチに基づく経路観測モデルは, 事前情報としての経路選択モデルを観測の段階で必要とする. 経路選択モデルの評価にはパラメータ θ の事前情報を与えなければならない. 経路を観測する段階で θ の真値は知り得ないため, 既往研究では他の調査データを用いた推定パラメータ^{4),14)} や任意の値⁵⁾ として与えることが多く, 全く情報がない場合には一様分布とされる^{12),14)}. しかし, いずれの場合においても最終的に推定される経路選択モデルのパラメータと異なることは明らかであり, 経路の観測確率および推定パラメータにバイアスを生じさせてしまうことに他ならない. 本研究では, θ が観測・行動モデルに共通のパラメータであることに着目し, パラメータの一致性を保持した形で両モデルの推定を行う構造推定手法を提案する. 構造推定のアルゴリズム (図-3) は以下のステップに従う.

- Step 1: 初期化. 観測 $\hat{\mathbf{m}}$ 及び初期パラメータ $\bar{\theta}$ をセットし, $h = 1$ とする.
- Step 2: 逐次リンク観測モデルの推定. 逐次リンク観測モデルを通じてリンクごとの観測誤差分散 σ_a および経路 r を推定する. $r^{(h)} = r$ とする.
- Step 3: 逐次リンク選択モデルの推定. $r^{(h)}$ をデータ・セットとし, 最尤推定法によって逐次リンク選択モデルを推定し, パラメータ $\tilde{\theta}$ を得る. $\theta^{(h+1)} = \tilde{\theta}$ とする.
- Step 4: 収束判定. 以下の不等式を満たせば終了し, $\theta = \theta^{(h+1)}$ とする. そうでない場合, $h = h+1$ として Step 2 へ戻る.

$$|\theta^{(h+1)} - \theta^{(h)}| < \xi \quad (22)$$

上記のアルゴリズムは, 経路選択モデルのパラメータ θ が不動点に達するまで, 観測・推定の操作を繰り返すことを意味する. なお, 本研究は観測モデル内で逐次的にリンクを確定させることにより計算の効率化と記憶容量の削減を行っており, 観測・行動モデルは段階的に推定される. そのため各モデルで尤度最大化操作を行っており, 同時推定フレームワークへの展開及び解の存在や一意性といった理論的な証明は今後の課題として最終章で詳述する. 本稿では以降の数値実験により, パラメータの収束性とアルゴリズムの有効性を検証する.

図-3 構造推定のフロー

6. 数値計算

本章では、前章までに提案した推定フレームワークの検証を数値実験を通じて行う。まず、仮想ネットワークを用いてシミュレーションデータを生成し、真値を知った状態でリンクの特定精度及びパラメータ推定結果を検証する。その後、愛媛県松山市で行われたプローブパーソン調査のデータを用いて、現実の歩行者ネットワークへ手法の適用を行う。

(1) シミュレーションデータを用いた数値実験

実験には、図-4のネットワークを用いる。ネットワーク上の括弧内の数字はそれぞれ、リンクコスト（連続変数） CC_a 、リンクコスト（1 or 0） DC_a 、リンク固有の誤差分散パラメータ σ_a を表している。ここでは変数は特定の意味を持たないため、具体的な変数の定義は6章(2)のケーススタディを参照されたい。 σ_a の値がネットワーク上で不均一であることを仮定し、 $x_2 = 60, 30$ 上のリンクの誤差分散パラメータを $\sigma_a = 20, 10$ とし、その他のリンクを $\sigma_a = 5$ とした。図-5は異なる σ_a の経路を通過した場合に生成される位置データをプロットしたものであり、分散の大きなリンクを通過する経路(b)では、経路(a)に比べて観測が大きな誤差を持っている（真の経路から離れて観測されている）ことがわかる。シミュレーションでは歩行者の経路選択行動を想定し、以下の一次マルコフ連鎖に基づく逐次リンク

選択モデルを用いる。

$$u(a_{t+1}|a_t) = \theta_1 LC_{a_{t+1}} + \theta_2 CC_{a_{t+1}} + \theta_3 DC_{a_{t+1}} + \theta_4 UT_{a_{t+1}|a_t}, \quad (23)$$

and

$$p(a_{t+1}|a_t) = \frac{\exp(u(a_{t+1}|a_t))}{\sum_{a_{t+1} \in \mathcal{A}(a_t)} \exp(u(a_{t+1}|a_t))}, \quad (24)$$

ここで、 LC_a はリンク長、 $UT_{a'|a}$ はUターンのダミー変数である。離散化のための時間幅 $\bar{t} = 30s$ 、観測間隔を $\hat{t}_n - \hat{t}_{n-1} = 10s$ で一定とし、また真のパラメータ値を $\bar{\theta} = [-0.1, -2, -1.5, -10]$ (TRUE values) としてシミュレーションを行い、50の経路データ及び各経路に対応する観測点列を発生させた。

複数のモデルを用いて経路観測を行った結果として、リンクの特定精度および σ の推定誤差を表-1に示す。リンクの特定精度はシミュレーションで発生させた時点 t のリンクデータ a_t^* と、特定された時点 t のリンク $\tilde{a}_t$ との一致性に基づいて割合を算出した。 σ の推定誤差は以下の式で計算した。

$$\frac{1}{|I||T|} |\tilde{\sigma}_a - \sigma_a^*|, \quad (25)$$

$|I||T|$ はサンプル数（リンクの観測数）である。表-1において、 $\bar{\theta} = [0, 0, 0, 0]$ であるモデル (Model 1 and 2) は事前情報を用いずに観測確率のみからリンクの特定を行うことを意味する。Model 1 はさらに観測誤差分散パラメータを所与かつネットワーク上で均一の値 $\sigma_a = 20$ ($\forall a \in \mathcal{A}$) とした場合であり、リンクの特定精度は

図-4 実験用ネットワーク

図-5 異なる σ_a を持つ経路のサンプリング例

54.57%であった。一方、Model 2 はリンク固有変数として分散パラメータを推定した場合であり、リンクの特定精度は 76.86%であった。つまりネットワーク上で σ の値に異質性が存在する場合には、 σ を所与とせずに推定することで観測精度が上昇することが示された。また、全てのモデルにおいて、スイッチングアルゴリズムを導入することでリンクの特定精度、 σ の推定精度が共に上昇しており、アルゴリズムの有効性が確認できた。また、Model 3, 4 は事前情報として経路選択モデルを導入して観測の補正を行ったモデルである。事前情報のパラメータ $\bar{\theta}$ に真値を代入した場合、リンクの特定精度が向上し、スイッチングアルゴリズムとの組み合わせによって 91.71%を達成している (Model 4)。一方、真値から大きく外れた値をインプットした Model 3 では、事前情報の導入によりかえってリンクの特定精度が悪化してしまう結果となった。観測頻度が低い、位置精度が悪い、ネットワークが密であるといったように観測の不確実性が大きい状況下では、事前情報の影響が強く働く。つまり、事前情報が正確でない場合には誤った方向に補正がなされてしまい、バイアスを増幅させる可能性があることが結果から示唆された。

Model 2, 3 (スイッチングあり) を用いて、観測モデルから得られた経路情報をもとに経路選択モデルを推定した結果を表-2 に示す。表中の“One-way”は初期パラメータ $\bar{\theta}$ を用いた経路の観測と、その観測結果を用いた経路選択モデルの推定を一度ずつのみ行ったときの結果を示している。“Structural estimation”は図-3 の構造推定を用いて、 θ が収束するまで計算を繰り返したときの結果を示している。収束判定は $\xi = 1e-10$ とした。ここで、サンプル数は (50 のサンプリング経路において) 観測されたリンク選択の回数である。“One-way”の初期尤度 $LL(0)$ がモデル間で異なるのは、事前情報の初期値 $\bar{\theta}$ に依存して観測経路 (データセット) が異なるためである。 $\bar{\theta} = [0, 0, 0, 0]$ (Model 2) および $\bar{\theta} = [-1.5, -0.1, -2, -10]$ (Model 3) のとき、“One-way”の推定結果は真値との差が大きく、いずれのモデルも (特に θ_2, θ_4 に関して) 真値と大きく異なる結果となった。修正済み尤度比 ρ^2 の観点からも良い推定結果とはいえない。表-1 に示したように、Model 2, 3 では真の経路に対する観測失敗が (特に σ の大きいリンク周辺で) 多く、推定される行動パラメータにバイアスが生じてしまったと考えられる。Model 3 では正負の逆転も見られ、また Input ($\bar{\theta}$) に近い値が得られているわけでもなく、事前情報がノイズとして働いてしまったことを示している。反対に、“Structural estimation”の結果は、観測と推定の反復によって事前情報が改善され、最終的な推定値が真値に近い値を示しただけでなく、いずれのケースにおいても $\bar{\theta}$ によらず同じ収束値をとることがわかった。最終尤度や ρ^2 に改善が見られたことから、構造推定の有効性が示されたといえよう。なお、初期値を $\bar{\theta} = [-10, -10, -10, -10], [-100, 0, 0, 0], [-100, -100, -100, -100], [10, 10, 10, 10]$ として推定を行った場合においても、全ての初期値に対して表-2 に見られる収束結果と全く同様の結果が得られた。

また、得られた結果のサンプル依存性を確認するために、上記と同様の手順で、シミュレーションデータの作成及び経路選択モデルの推定を 100 セット繰り返し行った。なお、検証には表-1 における Model 2 (スイッチングあり) を用いた。表-3 は、その結果得られた経路選択モデルのパラメータ値、収束に要した反復回数、計算時間それぞれの平均、標準偏差を示している。パラメータ推定値の平均・分散を見ると、リンク長 (θ_1) に関しては真値に近い値が安定して得られている一方で、その他のパラメータでは真値からのズレが見られる。しかしいずれの場合においてもパラメータ初期値 $\bar{\theta} = [0, 0, 0, 0]$ よりも真値に近づいていることから、観測モデルの精度を向上させることで推定結果の改善が期待される。また、U ターンダミー (θ_4) のパラメータは分散が大きく推定値がバラついている。U ターン

表-1 リンク特定精度および σ 推定誤差

Model	σ	$\bar{\theta}$	Link accuracy(%)		Ave. $ \bar{\sigma}_a - \sigma_a^* $	
			-	Switching	-	Switching
1	Given	[0, 0, 0, 0]	54.57	68.86	-	-
2	Estimated	[0, 0, 0, 0]	76.86	82.86	5.85	4.40
3	Estimated	[-1.5, -0.1, -2, -10]	4.86	38.29	41.99	21.21
4	Estimated	[-0.1, -2, -1.5, -4]	76.86	91.71	7.58	4.06

表-2 経路選択モデルのパラメータ推定結果

Input: $\bar{\theta} = [0, 0, 0, 0]$ (No information)							
	One-way				Structural Estimation		
	TRUE	Estimates	abs(diff.)	t-value	Estimates	abs(diff.)	t-value
θ_1	-0.1	0.002	0.102	0.10	-0.064	0.036	-2.56
θ_2	-2	-0.755	1.245	-4.16	-1.727	0.273	-6.88
θ_3	-1.5	-1.312	0.188	-4.77	-1.046	0.454	-3.52
θ_4	-4	-1.892	2.108	-8.86	-3.519	0.481	-9.74
sample				350			350
$LL(0)$				-373.22			-371.89
$LL(\bar{\theta})$				-269.87			-211.31
$\bar{\rho}^2$				0.266			0.421
iteration							6

Input: $\bar{\theta} = [-1.5, -0.1, -2, -10]$ (Wrong values)							
	One-way				Structural Estimation		
	TRUE	Estimates	abs(diff.)	t-value	Estimates	abs(diff.)	t-value
θ_1	-0.1	-0.097	0.003	-5.31	-0.064	0.036	-2.56
θ_2	-2	-0.419	1.581	-2.71	-1.727	0.273	-6.88
θ_3	-1.5	0.178	1.678	0.96	-1.046	0.454	-3.52
θ_4	-4	-1.204	2.796	-6.77	-3.519	0.481	-9.74
sample				350			350
$LL(0)$				-373.56			-371.89
$LL(\bar{\theta})$				-328.59			-211.31
$\bar{\rho}^2$				0.110			0.421
iteration							8

表-3 パラメータ推定値, 反復回数, 計算時間の平均・標準偏差 (100 回の試行)

	$\bar{\theta}_1$	$\bar{\theta}_2$	$\bar{\theta}_3$	$\bar{\theta}_4$	Iteration	CPU time (s)
Ave.	-0.112	-1.140	-1.006	-2.916	4.086	858.2
Std. err.	0.023	0.663	0.413	4.006	1.007	229.7

の抵抗を大きく設定した ($\theta_4^* = -4$) ことから, 各データセットに含まれる U ターン経路の観測数が少なくなってしまうと考えられる. 真値及びサンプリング数との関係についての詳細な検討が今後の課題として挙げられる. 反復回数, 計算時間の平均はそれぞれ 4.086 回, 858.2 秒であった. また今回の実験において, 構造推定が収束しないケースが見受けられた (100 セットのうち 7 セット). 上述したように, 収束が確認されたデータセットでは初期値によらず同じ値に収束する可能性を示したものの, サンプルによっては 2 つの値を反復してしまう場合があった. このことから, 解の収束性には課題を残しており, 理論的な検証が可能な枠組みへの展開が今後の課題として必要であると言える.

(2) ケーススタディ

最後に, 提案手法を実データに適用する. 用いるデータは, 松山市を対象に行われたプローブパーソン (以下, PP) 調査¹⁾ から得られたデータである. 中心市街地を訪れた歩行者の GPS データを対象に分析を行なった.

前節と同様, 一次マルコフ連鎖に基づく逐次リンク選択モデルを用いる. 歩行者がリンク a_t から a_{t+1} に推移する際の効用を以下のように設定する.

$$u(a_{t+1}|a_t) = \theta_1 LC_{a_{t+1}} + \theta_2 CU_{a_{t+1}} + \theta_3 DU_{a_{t+1}} + \theta_4 UT_{a_{t+1}|a_t}, \quad (26)$$

ここで, LC_{a_t} はリンク長 (m), CU_{a_t} はリンク a_t の歩道幅員 (m), DU_{a_t} はアーケードダミー変数, $UT_{a_{t+1}|a_t}$ は U ターンダミー変数である. 初期値を $\bar{\theta} = [0, 0, 0, 0]$ (事前情報なし, スイッチングあり) とし, ランダムに 30

表-4 実データを用いた経路選択モデルの推定結果

Input: $\tilde{\theta} = [0, 0, 0, 0]$ (No information)				
One-way		Structural Estimation		
	Estimates	t-value	Estimates	t-value
θ_1	-0.007	-2.47	-0.001	-0.43
θ_2	0.088	1.50	0.134	1.58
θ_3	-0.004	-0.01	2.760	4.29
θ_4	0.774	0.53	0.469	3.34
sample		270		270
$LL(0)$		-307.61		-309.07
$LL(\tilde{\theta})$		-302.17		-225.16
$\bar{\rho}^2$		0.005		0.259
iteration				11

トリップ（792 の位置データ，270 回のリンク選択）を抽出して推定を行った．結果を表-4 に示す．なお，収束判定は $\xi = 1e-10$ として計算した．“One-way”は事前情報のない場合の観測結果を用いて経路選択モデルを推定した結果であり，その結果と比較して，構造推定を用いた場合には最終尤度， ρ^2 の観点からも結果の向上が確認できる．事前情報のない場合の推定結果としては，リンク長（旅行時間）のパラメータである θ_1 が負に有意な値をとった．一方で構造推定の結果は， θ_1 は有意ではなく，寧ろアーケードの有無（ θ_3 ）やUターン（ θ_4 ）に有意な傾向を示した．このことは，事前情報のないケースでは観測の不確実性が大きいアーケードのリンクの通過を捉えられず，最短経路が選ばれやすいという誤った因果を導いてしまう可能性を示唆している．Uターンダミーのパラメータが正の値をとった結果については，他の交通手段と比較して歩行者はUターンに対する抵抗が小さいことを表すと考えられるものの，より詳細な検証が必要である．

また，観測モデルを通じたリンク単位の誤差分散パラメータ σ_a の最終的な推定結果を地図上にプロットした（図-6）．太さは分散の大きさを示しており，リンクごとに大きなばらつきを持っていることがわかる．このことから，観測誤差分散をネットワーク上で均一かつ所与の値として扱う従前の手法は，観測・推定のプロセスにバイアスを含んでしまう可能性が示唆される．また，リンクごとに推定結果を確認し，いくつかの直感に従う結果が得られた．図の下方 A の通りは背の低いアーケードを持つ「銀天街」と呼ばれる松山市の中央商店街を示しており，アーケードの存在が観測誤差分散を大きくしたものと考察される．また図中のリンク B，C はそれぞれ松山市役所，愛媛県庁の前面に位置するリンクであり，本研究で用いた PP データにおいて両施設は目的地として多く選ばれている．PP 調査ではトリップ終了報告をモニタが手動で行うため，施設内に入った後の位置情報をトリップ内に含んでいることから，リンク B，C の観測誤差分散が大きくなったものと考えられる．

図-6 観測誤差分散の推定結果のプロット

7. おわりに

(1) 本研究の結論

不確実性を持つ観測を前提とした従前の経路選択モデルの推定プロセスは，観測モデルにおいて必要なパラメータの初期設定によってバイアスを含んでしまっていた．本研究は経路観測モデルと経路選択モデルの相互関係に依存した推定バイアス増幅のメカニズムに着目し，バイアス軽減のための推定フレームワークを新たに提案した．

本研究ではまず，逐次リンク選択行動を観測・構造モデル共通の基本単位とした逐次リンク観測モデルを提案した．逐次的なリンクの特定を行うことで，観測点と候補リンクとの対応関係があらかじめ明らかとなり，従前では所与かつネットワーク上で均一の値として扱われてきた観測誤差分散のパラメータをリンク固有変数として推定することを可能にした．また，ベジアンアプローチに基づき，事前情報として逐次リンク選択モデルを用いることで，行動メカニズムによる観測の補正を経路集合の列挙を必要とせずに行うことができる．一方で，事前情報が最終的に推定される経路選択モデルと不一致な場合に発生する推定バイアスにも目を向け，両者の一致問題を解くための構造推定フレームワークを提案した．

双子実験によるモデルの検証では，リンクごとの誤差分散推定及びスイッチングアルゴリズムの導入により，経路の特定精度が向上することを確認した．一方で事前情報を導入する場合には，観測結果が行動パラメータに大きく依存することを確認した．構造推定の結果として，パラメータの初期値に依らず一定の値に

収束し、かつ推定結果が真値に近づくことを明らかにした。最後に愛媛県松山市中心市街地ネットワークにおける実歩行者データを用いた検証を行い、構造推定によって観測が難しいアーケードの選択メカニズムを捉えた。また推定された観測誤差分散の空間的なばらつきは直感に沿うものであった。これらの結果は本手法の有用性を示していると考えられよう。

(2) 適用上の注意と今後の研究課題

今後の課題として、サイクリック経路の観測や、リンク空間特性と σ との関係についての統計的検証、双子実験の精緻化などが挙げられる。また、逐次リンク観測を生かしたより一般的なフレームワークへの展開が求められる。本稿では計算を記憶容量及び計算速度の観点から効率的にするため、逐次的にリンクを確定させる枠組みを提案した。数値計算からは解の収束が確認され、有意な結果が導かれたものの、解の存在・一意性・収束性といった問題の理論的な性質についての妥当性が証明されていない点、提案手法をさまざまな場面に適用する際の注意点である。そのため、次の展開として、経路観測モデル内の各時点において計算されるリンクの観測確率を保存し、以下の尤度を最大化させることで経路選択モデルとの同時推定を行うことを考える。

$$p(\hat{\mathbf{m}}_{1:T}; \sigma, \theta) = \sum_{a_{1:T} \in \mathcal{P}} p(\hat{\mathbf{m}}_{1:T} | a_{1:T}; \sigma) p(a_{1:T}; \theta), \quad (27)$$

ここで $\mathcal{P}$ は観測モデルによって得られる経路集合である。式 (27) は観測点列 $\hat{\mathbf{m}}_{1:T}$ が経路選択モデル・経路観測モデルを通じて再生産される確率（尤度）を意味し、この尤度を最大化させることで σ, θ を同時推定することが可能と考えられる。しかし、逐次的にリンクの確率を保存していくことで、その組み合わせ数（ $|\mathcal{P}|$ ）は少ない時点数でも膨大となる。経路観測確率保存のアルゴリズム、式 (27) の最大化アルゴリズム（EM アルゴリズム等）の検討・開発、そして解の存在・一意性・収束性といった理論的な検証を次の研究課題として言及したい。

謝辞： 本研究は JSPS 科研費 JP14J10824, JP17J09979 の助成を受けたものです。ここに感謝の意を表します。

参考文献

- Hato, E.: Development of MoALs (Mobile Activity Loggers supported by GPS-phones) for travel behavior analysis, *Transportation Research Board 85th Annual Meeting*, 2006.
- Bierlaire, M. and Frejinger, E.: Route choice modeling with network-free data, *Transportation Research Part C: Emerging Technologies*, Vol.16, No.2, pp.187-198, 2008.
- Bierlaire, M., Chen, J. and Newman, J.: A probabilistic map matching method for smartphone GPS data, *Transportation Research Part C: Emerging Technologies*, Vol.26, pp.78-98, 2013.
- 布施孝志, 中西航: 歩行者挙動モデルを統合した人物追跡手法の構築, 土木学会論文集 D3 (土木計画学), Vol.68, No.2, pp.92-104, 2012.
- Danalet, A., Farooq, B. and Bierlaire, M.: A Bayesian approach to detect pedestrian destination-sequences from WiFi signatures, *Transportation Research Part C: Emerging Technologies*, Vol.44, pp.146-170, 2014.
- White, C. E., Bernstein, D. and Kornhauser, A. L.: Some map matching algorithms for personal navigation assistants, *Transportation Research Part C: Emerging Technologies*, Vol.8, No.1-6, pp.91-108, 2000.
- Greenfeld, J. S.: Matching GPS observations to locations on a digital map, *Transportation Research Board 81st Annual Meeting*, 2002.
- Quddus, M. A., Ochieng, W. Y., Zhao, L. and Noland, R. B.: A general map matching algorithm for transport telematics applications, *GPS Solutions*, Vol.7, No.3, pp.157-167, 2003.
- Velaga, N. R., Quddus, M. A. and Bristow, A. L.: Developing an enhanced weight-based topological map-matching algorithm for intelligent transport systems, *Transportation Research Part C: Emerging Technologies*, Vol.17, No.6, pp.672-683, 2009.
- Ochieng, W. Y., Quddus, M. A. and Noland, R. B.: Map-matching in complex urban road networks, *Brazilian Journal of Cartography*, Vol.55, No.2, pp.1-14, 2003.
- Quddus, M. A., Ochieng, W. Y. and Noland, R. B.: Integrity of map-matching algorithms, *Transportation Research Part C: Emerging Technologies*, Vol.14, No.4, pp.283-302, 2006.
- Hunter, T., Abbeel, P. and Bayen, A.: The path inference filter: model-based low-latency map matching of probe vehicle data, *IEEE Transactions on Intelligent Transportation Systems*, Vol.15, No.2, pp.507-529, 2014.
- Pyo, J.S., Shin, D.H. and Sung, T.K.: Development of a map matching method using the multiple hypothesis technique, *IEEE Intelligent Transportation Systems Conference Proceedings*, pp. 23-27, 2001.
- Chen, J. and Bierlaire, M.: Probabilistic multimodal map matching with rich smartphone data, *Journal of Intelligent Transportation Systems*, Vol.19, No.2, pp.134-148, 2015.
- Fosgerau, M., Frejinger, E. and Karlstrom, A.: A link based network route choice model with unrestricted choice set, *Transportation Research Part B: Methodological*, Vol.56, pp.70-80, 2013.
- 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.
- Mai, T.: A method of integrating correlation structures for a generalized recursive route choice model, *Transportation Research Part B: Methodological*, Vol.93, pp.146-161, 2016.
- McFadden, D.: Modelling the choice of residential location, 1978.
- 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, Vol.3, No.5, pp.1-10, 2017.
- 久保信明: GPS 測位におけるマルチパス誤差の低減化と高精度測位の可能性について, 東京大学博士論文, 2005.
- van Diggelen, F.: GNSS Accuracy: Lies, Damn Lies, and Statistics, *GPS World*, Vol.18, No.1, pp.26-33, 2007.
- Quddus, M. A., Noland, R. B. and Ochieng, W. Y.: Validation of map matching algorithms using high precision positioning with GPS, *Journal of Navigation*, Vol.58, No.2, pp.257-271, 2005.

- 23) Rust, J.: Optimal replacement of GMC bus engines: An empirical model of Harold Zurcher, *Econometrica: Journal of the Econometric Society*, pp. 999-1033, 1987.
- 24) Aguirregabiria, V. and Mira, P.: Swapping the nested fixed point algorithm: A class of estimators for discrete Markov decision models, *Econometrica*, Vol.70, No.4, pp.1519-1543, 2002.
- (2017. 2. 24 受付)

STRUCTURAL ESTIMATION FOR ROUTE CHOICE MODELS CONSIDERING LINK SPECIFICITY OF MEASUREMENT ERROR VARIANCES

Yuki OYAMA and Eiji HATO

For estimating parameters of discrete choice models, observations corresponding to the models are required. In the context of route choice models, we need the information of paths, which are sequences of links that connect between the origin-destination pairs. Passive monitoring with Global Positioning System (GPS) is more and more used to observe trip data, because it contributes to facilitating to observe trip data automatically. However, data from monitoring with GPS is not consistent, in formats, with network, which is the fundamental description of route choice models, and it has the heteroscedasticity of measurement errors dependently on devices and locations. These errors cause the biased observation of route choices, and as the result, the parameter estimation results of route choice models can be biased. In this study, we propose a sequential link measurement method, which is a Bayesian approach and incorporates a Markovian route choice model as the prior. It allows one to infer links based on both measurements and behavioral mechanisms, and at the same time, to estimate the variance of GPS measurement error on each link. Moreover, we propose a structural estimation method for a route choice model to remove biases regarding the initial parameter settings. The performances of these methods are examined through a numerical example and a case study of applying in a real pedestrian network. As a result, the estimated preferences of route choices using the structural estimation method are less biased and show the different trend from those using the biased route choice observations. Moreover, the estimated variances of GPS measurement errors are realistic.