

不完全交通データを活用した道路ネットワーク 交通状態のリアルタイム更新

川村 雄斗¹・峪 龍一²・内田 賢悦³

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

E-mail: ky12081023@eis.hokudai.ac.jp

² 正会員 北海道大学大学院 工学研究院 (〒060-8628 北海道札幌市北区北 13 条西 8 丁目)

E-mail: r-tani@eng.hokudai.ac.jp

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

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

本研究では、交通感知器とプローブカーから観測されたデータを用いて、複数の時間帯における道路ネットワーク全体のリンク交通量とリンク遅れ時間を確率変数として推定する手法を提案する。多変量対数正規分布に従うリンク交通量と遅れ時間を仮定し、観測リンク交通量とリンク遅れ時間に関する尤度最大化問題を解いて交通状態を推定する。本モデルでは交通観測データの観測箇所の不完全性を考慮して尤度を計算する。最尤推定によって推定された交通状態の確率分布を事前分布として、交通感知器から得られる交通量データおよびプローブデータから得られるリンク移動時間を尤度とするベイズ推定を行い、交通状態の事後分布を推定する手法を提案する。ここでは、リアルタイムに得られた欠測のあるデータの利用を想定する。最後にテストデータを用いた提案モデルの検証結果を示す。

Key Words: probe car data, traffic counter data, network uncertainty, traffic state estimation, Bayse estimation

1. はじめに

道路ネットワーク全体の交通状態を推定するために、プローブカーデータと交通感知器が利用されている。プローブカーデータは、安価で大量のデータを取得できることから、先進国・新興国問わず、広く利用されることが期待されており、各車両の挙動を軌跡として観測できる。しかし、同時時間帯におけるネットワーク規模での交通状態の網羅的な観測はプローブカーの密度に依存しており、プローブカーの密度が低い場合には、同時時間帯におけるネットワーク全体の交通状態を推定することは困難である。したがって、同時時間帯において、プローブカーで観測できていないリンクの交通状態を推定することが課題となっている。

プローブカーデータの利用が検討されている一方で、日本では交通量を観測するために交通感知器が一部のリンクに設置されている。交通感知器を設置して運用するための費用は高いことから、道路ネットワークを網羅するように交通感知器を設置することは財政的な都合上、現実的ではないのが実情である。しかし、固定的かつ連

続的に交通量を観測できるという特徴は、プローブカーデータにはない側面である。プローブカーデータと交通感知器データを補完的に利用することによって、効率的な交通状態観測が実現される可能性がある。

プローブカーデータを用いて、道路ネットワークにおける未観測な交通状態を推定する研究がいくつか行われている。原ら (2016)¹⁾、Hara et al. (2018)²⁾ はそれぞれプローブカーデータを用いて、道路ネットワーク全体における、リンク速度の補完推定を行っている。原ら (2016)¹⁾ はグラフィカルラッソを用いて、リンク間の共分散構造を推定するモデルを提案した。Hara et al. (2018)²⁾ は同様な問題に対して、混合ガウシアングラフィカルモデルを利用している。対象とする道路ネットワークの規模が大きくなると推定すべき共分散構造に対して、推定すべきパラメータは増大する。そのため、共分散構造のスパース性に着目して、交通状態推定を行った。

複数の交通状態観測手法を組み合わせる研究はすでに行われている。例えば、Seo et al. (2015)³⁾ は交通感知器とフローティングカーからそれぞれ観測されたデータを組み合わせることによって、道路リンクにおける動的な密

度推定を行う手法を提案している。峪・内田 (2020)⁴⁾は道路ネットワークの不確実性を考慮した交通量配分モデルに基づく均衡制約付きの最尤推定モデルを構築し、複数時間帯を対象としてネットワーク全体のリンク交通量とリンク移動時間の同時分布を推定する手法を提案した。

本研究では、峪・内田 (2020)⁴⁾によって構築された最尤推定モデルから推定された交通状態の確率分布を事前分布として、交通感知器から得られる交通量データおよびプローブデータから得られる移動時間を尤度とするベイズ推定を行い、交通状態の事後分布を推定する手法を提案する。さらにテストネットワークにおける数値計算により手法の検証を行う。

本論文の構成は以下の通りである。第2章では、本モデルを構築するために設けた仮定と、議論の際に使用した記号について説明する。第3章ではネットワークモデルと複数時間帯の問題を考慮できる尤度最大化問題を定義する。第4章では交通状態の事後分布を推定するベイズ推定を定義する。第5章ではテストネットワークを用いて、複数時間帯における交通状態の事後分布を推定する計算例を示す。最後に第6章では本研究のまとめと今後の展望を示す。

2. 本研究における仮定

(1) 仮定

以下に、本研究で設定した仮定を示す。

- 本研究で推定の対象とする交通状態は、リンク交通量とリンク遅れ時間である。リンク遅れ時間はリンク移動時間に対応しており、それぞれの同時分布は自由走行時間の分だけ平行移動したものとなっている。
- リンク交通量は交通感知器が設置されているリンクにおいてのみ観測されるものとする。リンク遅れ時間はすべてのリンクにおいて観測されるものの、各観測機会において観測される確率は、道路ネットワーク全体におけるプローブカー混入率によって決まるものとする。つまり、各観測機会において、すべてのリンクにおいて交通量と移動時間の双方が観測されている状況ではなく、時空間的に欠測のあるデータセットを利用することを前提としている。これはつまり、現実のデータの不完全性を考慮していることとなる。
- ネットワーク上の生成交通量は対数正規分布に従う確率変数である。生成交通量は、各時間帯の時間帯別生成交通量の総和として定義される。次章以降で説明するように、生成交通量を確率変数として、時間帯別生成交通量、OD交通量、経路交通

量、リンク交通量はそれぞれ、生成交通量の実数倍として定義される。

(2) 記号

本稿で用いる主な記号は以下に示す通りである。ここで、集合および確率変数は大文字で表すものとし、その集合の要素、あるいは確定的な変数は小文字で表すものとする。

A	ネットワーク上のリンク集合
A_0	ネットワーク上の交通感知器により、交通データが観測されているリンクの集合
I	O-D ペアの集合
J	経路集合
T	時間帯集合
δ_{aj}	リンク a が経路 j の一部であれば 1, それ以外は 0 をとる変数
Q	ネットワーク上の総交通需要 (生成交通量)
Q_t	各時間帯の時間帯別生成交通量
Q_{ti}	時間帯 t の OD ペア i 間の確率的交通需要 (OD 交通量)
F_{tij}	時間帯 t の OD ペア i 間の経路 j の確率的交通量 (経路交通量)
V_{ta}	時間帯 t のリンク a の確率的交通量 (リンク交通量)
p_{ti}	時間帯 t の OD ペア i 間の交通需要の配分率
p_{tij}	時間帯 t の OD ペア i の経路 j を選択する確率 (経路選択確率)
cv	生成交通量の変動係数
c_{tij}	時間帯 t の OD ペア i 間の経路 j の一般化費用
t_{ta}^0	時間帯 t のリンク a の自由走行時間
C_{ta}	時間帯 t のリンク a の確率的交通容量
γ_a, λ_a	リンク a の移動時間のパラメータ
Ξ_{tij}	時間帯 t の OD ペア i の経路 j の確率的移動時間 (経路移動時間)

3. モデル

(1) 交通流

確率的総交通需要 (生成交通量) Q は対数正規分布に従う。

$$Q \sim LN(\mu_Q, \sigma_Q) \quad (1)$$

μ_Q, σ_Q は対数正規分布のパラメータである。時間帯別生成交通量 Q_t は、割合 p_t を用いて以下のように表される。

$$Q_t = p_t \cdot Q \quad \forall t \in T \quad (2)$$

生成交通量 Q_t は対数正規分布に従い、そのパラメータは生成交通量 Q のパラメータで表現される。

$$Q_t \sim LN(\mu_{Q_t}, \sigma_{Q_t}^2) \quad \forall t \in T \quad (3)$$

ここで、以下の関係が成立する。

$$\mu_{Qt} = \mu_Q + \ln(p_t) \quad \forall t \in T \quad (4)$$

$$\sigma_{Qt}^2 = \sigma_Q^2 \quad \forall t \in T \quad (5)$$

OD 交通量 Q_{ti} は、OD 配分率 p_{ti} と時間帯別生成交通量 Q_t の積として表現される。

$$Q_{ti} = p_{ti} \cdot Q_t \quad \forall t \in T, \forall i \in I \quad (6)$$

OD 交通需要 Q_{ti} は対数正規分布に従い、そのパラメータは、生成交通量 Q のパラメータで表現される。

$$Q_{ti} \sim LN(\mu_{Q_{ti}}, \sigma_{Q_{ti}}^2) \quad \forall t \in T, \forall i \in I \quad (7)$$

ここで、以下の関係が成立する。

$$\mu_{Q_{ti}} = \mu_Q + \ln(p_t \cdot p_{ti}) \quad \forall t \in T, \forall i \in I \quad (8)$$

$$\sigma_{Q_{ti}}^2 = \sigma_Q^2 \quad \forall t \in T, \forall i \in I \quad (9)$$

経路交通量 F_{tij} は、OD 交通需要 Q_{ti} と、確率的利用者均衡配分モデルによって推定される経路選択確率 p_{tij} の積によって表現されるものとし、対数正規分布に従う。

$$F_{tij} = p_{tij} \cdot Q_{ti} = p_{tij} \cdot p_{ti} \cdot Q_t \quad \forall t \in T, \forall i \in I, \forall j \in J_t \quad (10)$$

$$F_{tij} \sim LN(\mu_{F_{tij}}, \sigma_{F_{tij}}^2) \quad \forall t \in T, \forall i \in I, \forall j \in J_t \quad (11)$$

ここで、以下の式が成立する。

$$\mu_{F_{tij}} = \mu_Q + \ln(p_t \cdot p_{tij} \cdot p_{ti}) \quad \forall i \in I, \forall j \in J_t \quad (12)$$

$$\sigma_{F_{tij}}^2 = \sigma_Q^2 \quad \forall i \in I, \forall j \in J_t \quad (13)$$

リンク交通量はそのリンクを通過する経路交通量の和として表される。

$$\begin{aligned} V_{ta} &= \sum_{i \in I} \sum_{j \in J_t} \delta_{aj} \cdot F_{tij} \\ &= \sum_{i \in I} \sum_{j \in J_t} \delta_{aj} \cdot p_{tij} \cdot Q_{ti} = \hat{p}_a \cdot Q_t \quad \forall a \in A, \forall t \in T \end{aligned} \quad (14)$$

ここで、以下の関係が成立する。

$$\begin{aligned} \hat{p}_a &= \sum_{i \in I} \sum_{j \in J_t} \delta_{aj} \cdot p_{tij} \cdot p_{ti} \quad \forall a \in A \\ &= \sum_{i \in I} \sum_{j \in J_t} \delta_{aj} \cdot p_{tij} \cdot Q_{ti} = \hat{p}_a \cdot Q_t \quad \forall a \in A, \forall t \in T \end{aligned} \quad (15)$$

上記までの定式化によって、リンク交通量はスカラー量 \hat{p}_a と生成交通量の積として定義される。したがって、リンク交通量もまた、対数正規分布に従い、そのパラメータは生成交通量 Q のパラメータで表現される。

$$V_{ta} \sim LN(\mu_{V_{ta}}, \sigma_{V_{ta}}^2) \quad \forall t \in T, \forall a \in A \quad (16)$$

ここで、以下の関係が成立する。

$$\mu_{V_{ta}} = \ln(\hat{p}_a) + \mu_Q \quad \forall t \in T, \forall a \in A \quad (17)$$

$$\sigma_{V_{ta}}^2 = \sigma_Q^2 \quad \forall a \in A \quad (18)$$

(2) 移動時間

リンク移動時間は以下に示す BPR 関数に基づいて定義されるものとする。

$$T_{ta}(V_{ta}, C_{ta}) = t_{ta}^0 \left(1 + \gamma_a \cdot \left(\frac{V_{ta}}{C_{ta}} \right)^{\lambda_a} \right) \quad \forall t \in T, \forall a \in A \quad (19)$$

上に示す、確率的なリンク移動時間は確定項と確率項に

分離することができる。ここで、リンク交通量とリンク交通容量の変動に影響される確率項を、「リンク遅れ時間」と定義する。以下ではリンク移動時間に着目する。リンク交通量とリンク交通容量はそれぞれ対数正規分布に従うことから、リンク遅れ時間もまた対数正規分布に従うことがわかる。ここで、リンク遅れ時間はリンク移動時間を確定項の分だけ平行移動したものであるため、それぞれの分散共分散は等しい。各リンクにおける自由走行時間が所与であるとする、リンク移動時間を推定することとリンク遅れ時間を推定することは等しくなる。しかし、移動時間と同時に推定する交通状態としてリンク交通量を含める場合、尤度を計算するために、分布形状を等しくする都合、推定する交通状態としてリンク遅れ時間を採用することが適当である。

$$T_{ta}(V_{ta}, C_{ta}) = t_{ta}^0 + \frac{t_{ta}^0 \cdot \gamma_a}{C_{ta}^{\lambda_a}} \cdot (V_{ta})^{\lambda_a} \quad (20)$$

経路移動時間は、その経路を通過するリンクの移動時間の和として表現される。

$$\Xi_{tij} = \sum_{a \in A} T_{ta} \cdot \delta_{aj} \quad \forall t \in T, \forall i \in I, \forall j \in J_t \quad (21)$$

(3) 均衡配分モデル

経路交通量は、一般化費用 c_{tij} を評価指標として、確率的利用者均衡配分モデルを解いて求められる経路選択確率 p_{tij} を用いて、以下に示す不動点問題を解くことで求められる。

$$E[F_{tij}] = p_{tij}(\mathbf{c}_t(\mathbf{F}_t)) \cdot E[Q_{ti}] \quad \forall t \in T, \forall i \in I, \forall j \in J_t \quad (22)$$

$$p_{tij} = \frac{\exp(-\phi \cdot c_{ij}(\mathbf{F}_t))}{\sum_{j=1}^{|J_t|} \exp(-\phi \cdot c_{ij}(\mathbf{F}_t))} \quad \forall t \in T, \forall i \in I, \forall j \in J_t \quad (23)$$

ここで、以下の関係が成立する。

$$\mathbf{c}_t = (c_{t11}(\mathbf{F}_t), \dots, c_{t|I||J_t|}(\mathbf{F}_t)) \quad (24)$$

$$\mathbf{F}_t = (F_{11}, \dots, F_{|I||J_t|}) \quad (25)$$

\mathbf{c}_t は時間帯 t における経路選択費用を表すベクトルである。 \mathbf{F}_t は時間帯 t における確率的経路交通量を表すベクトルである。一般化費用 $c_{tij}(\mathbf{F}_t)$ は、経路移動時間の平均と分散によって表現される。なお、 η はドライバーのリスク回避的な経路選択行動を示すためのパラメータである。

$$c_{tij}(\mathbf{F}_t) = E[\Xi_{tij}] + \eta \cdot \text{var}[\Xi_{tij}] \quad \forall t \in T, \forall i \in I, \forall j \in J_t \quad (26)$$

(4) 複数時間帯における最尤推定

本節では、峪・内田 (2020)⁴⁾ で提案された手法に基づいて、複数時間帯を対象として、尤度最大化によって、交通状態を推定する手法について述べる。

まず、時間帯 t において交通状態を推定する手法につい

て説明する。各リンクにおいて推定される交通状態を $\mathbf{X}_{at} = (V_{at} \ T_{at}')$ とすると、リンク交通量およびリンク遅れ時間はそれぞれ対数正規分布に従うことから、 $\mathbf{X}_{at} \sim MVLN(\boldsymbol{\mu}_{at}, \boldsymbol{\Sigma}_{aa,t})$ となり、2変量の対数正規分布に従う。ここで、 $\boldsymbol{\mu}_{at}$ と $\boldsymbol{\Sigma}_{aa,t}$ はそれぞれ V_{at} と T_{at}' に対応する多変量正規分布の平均と分散共分散である。道路ネットワーク全体を対象とする交通状態ベクトル、 $\mathbf{X}_t = (\mathbf{X}_{1t}, \dots, \mathbf{X}_{|A|t})$ も同様に、 $\mathbf{X}_t \sim MVLN(\boldsymbol{\mu}_t, \boldsymbol{\Sigma}_t)$ となり、多変量対数正規分布に従う。

次に、時間帯 t において、 G 日間の交通観測データが得られたとき、 g 日目の観測に対する観測状態ベクトル \mathbf{m}_{gt} と観測値ベクトル $\hat{\mathbf{d}}_{gt}$ を定義する。リンク a において観測された交通状態のパターンを K_a とする。ここで、 K_a の各パターンは、リンク交通量のみが観測される場合 $\{v\}$ 、リンク遅れ時間のみが観測される場合 $\{t\}$ 、リンク交通量とリンク遅れ時間の双方が観測される場合 $\{v, t\}$ 、いずれも観測されない場合の4種類からなる。

$g \in \{1, \dots, G\}$ 日目に観測されるデータ数を $n(g)$ 、時間帯 t とする。データの欠測を考慮した上で、多変量対数正規分布 $\mathbf{X} \sim LN(\boldsymbol{\mu}_{Qt}, \boldsymbol{\Sigma}_{Qt})$ に従う道路ネットワーク

全体の交通状態は、それに対応する以下の多変量正規分布の尤度関数を最大化することによって推定することができる。

$$L(\boldsymbol{\mu}_{Qt}, \boldsymbol{\Sigma}_{Qt}) \quad (27)$$

$$= \prod_{g=1}^G \frac{\exp\left(-\frac{1}{2}(\hat{\mathbf{d}}_{gt} - \boldsymbol{\mu}_{gt})^T \boldsymbol{\Sigma}_{gt}^{-1}(\hat{\mathbf{d}}_{gt} - \boldsymbol{\mu}_{gt})\right)}{2\pi^{n(g)/2} |\boldsymbol{\Sigma}_{gt}|^{1/2}}$$

ここで、以下の関係が成立する。

$$\boldsymbol{\Sigma}_{gt} = \mathbf{M}_{gt} \boldsymbol{\Sigma}_{Qt} \mathbf{M}_{gt}^T \quad (28)$$

$$\boldsymbol{\mu}_{gt} = \mathbf{M}_{gt} \boldsymbol{\mu}_{Qt} \quad (29)$$

$$\hat{\mathbf{d}}_{gt} = \mathbf{M}_{gt} \mathbf{d}_{gt}^T \quad (30)$$

\mathbf{d}_{gt} は、 $g \in \{1, \dots, G\}$ 日目に道路ネットワーク全体で観測される交通状態を表すベクトルで、以下のように定義する。

$$\mathbf{d}_{gt} = (\hat{v}_1 \ \hat{t}_1 \ \dots \ \hat{v}_a \ \hat{t}_a \ \dots \ \hat{v}_{|A|} \ \hat{t}_{|A|}) \quad (31)$$

ここで、 $\hat{v}_{|A|}$ 、 $\hat{t}_{|A|}$ はそれぞれ、観測されたリンク交通量、リンク遅れ時間の対数をとったものである。

\mathbf{M}_{gt} は、観測データ \mathbf{d}_{gt} を縮減する行列である。例えば、3本のリンクから構成される道路ネットワークにおいて、 $g \in \{1, \dots, G\}$ 日目にリンク 1 において交通量と遅れ時間が観測され、リンク 2 において遅れ時間のみが観測され、リンク 3 では何も観測されない場合、 \mathbf{d}_{gt} 及び \mathbf{M}_{gt} はそれぞれ以下のように表される。

$$\mathbf{d}_{gt} = (\hat{v}_1 \ \hat{t}_1 - \hat{t}_2 - \dots) \ \forall g \in \{1, \dots, G\} \quad (32)$$

$$\mathbf{M}_{gt} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix} \quad (33)$$

これら2つを乗算することで、観測データの行列が縮減

され、(30)で示す $\hat{\mathbf{d}}_{gt}$ として表現される。これを(27)で示す尤度関数に導入することによって、データの欠測を考慮した上で、各観測機会でスケールの異なる尤度関数が定義される。また、多変量分布のパラメータは(19)-(23)で定式化した確率的利用者均衡配分問題を解くことによって求まる。したがって、上記の問題は、確率的利用者均衡配分問題によって制約された尤度最大化問題となる。道路ネットワーク全体の交通状態の推定値は多変量対数正規分布に従い、 $\mathbf{X} \sim LN(\boldsymbol{\mu}_{Qt}, \boldsymbol{\Sigma}_{Qt})$ と表現される。こ

の交通状態の自然対数をとった確率変数は多変量正規分布に従い、 $\ln(\mathbf{X}) \sim N(\boldsymbol{\mu}_{Qt}, \boldsymbol{\Sigma}_{Qt})$ と表現される。ここで、

両者のパラメータは同一である。したがって、多変量対数正規分布に従う交通状態の観測値を尤度として最尤推定を行ったときの推定値と、その交通状態を対数変換して多変量正規分布について最尤推定を行ったときの推定値は一致する。本研究で用いる手法は、この性質を利用して、対数変換後に多変量正規分布に従う交通状態を用いて最尤推定を行い、パラメータを推定するものである。

本問題では、複数時間帯を対象として最尤推定を行うことから、推定の対象となる交通状態の次元が大きくなる。そこで、3章の(1)-(2)で定義するネットワーク交通流と移動時間の定式化に基づき、各時間帯のOD交通需要の平均のベクトルである \mathbf{q}_t と日生成交通需要の変動係数 cv を用いて、尤度関数を $L_t(\mathbf{q}_t, cv)$ と再定義する。

したがって、交通状態に関する尤度関数が、交通量配分問題を介して、OD交通需要の平均ベクトルと日生成交通需要の変動係数に関する尤度関数となった。(27)-(30)式に、前述で再定義した尤度関数 $L_t(\mathbf{q}_t, cv)$ を適用したものを以下に示す。

$$L_t(\mathbf{q}_t, cv) \quad (34)$$

$$= \prod_{g=1}^G \frac{\exp\left(-\frac{1}{2}(\hat{\mathbf{d}}_{gt} - \boldsymbol{\mu}_{gt})^T \boldsymbol{\Sigma}_{gt}^{-1}(\hat{\mathbf{d}}_{gt} - \boldsymbol{\mu}_{gt})\right)}{2\pi^{n(g)/2} |\boldsymbol{\Sigma}_{gt}|^{1/2}}$$

ここで、以下の関係が成立する。

$$\boldsymbol{\Sigma}_{gt} = \mathbf{M}_{gt} \boldsymbol{\Sigma}_{Qt}(cv) \mathbf{M}_{gt}^T \quad (35)$$

$$\boldsymbol{\mu}_{gt} = \mathbf{M}_{gt} \boldsymbol{\mu}_{Qt}(\mathbf{q}_t) \quad (36)$$

$$\hat{\mathbf{d}}_{gt} = \mathbf{M}_{gt} \mathbf{d}_{gt}^T \quad (37)$$

全時間帯全ODペアの交通需要を効率的に推定するため、 \mathbf{q}_t あるいは cv の片方を所与とした尤度関数 $L_t(\mathbf{q}_t|cv)$ と $L_t(cv|\mathbf{q}_t)$ をそれぞれ定義する。この2つの尤度関数に関する尤度最大化問題を交互に解くことによって、全時間帯全道路ネットワークの交通状態を推定するアルゴリズムを以下に示す。ここで、繰り返し回数 n 回目の日生成交通需要の変動係数を cv_n とし、時間帯 t の繰り返し回数 n 回目のOD交通量の平均ベクトルを \mathbf{q}_{tn}

とする。

- Step 1 繰り返し回数 $n = 1$ とし、日生成交通需要の変動係数の初期解を与える (cv_1)
- Step 2 日生成交通需要の変動係数 cv_n を所与として、各時間帯各OD交通需要の平均に関して尤度最大化問題を解き、各時間帯各OD交通需要の平均 q_{tn} を得る。
- Step 3 各時間帯各OD交通需要の平均 q_{tn} を所与とし、日生成交通需要の変動係数に関して尤度最大化問題を解き、日生成交通需要の変動係数 cv_{n+1} を得る。
- Step 4 対数尤度が収束していれば計算を終了し、収束していなければ、 $n = n + 1$ として、Step 2に戻る

Step 2において、時間帯 t についての尤度最大化問題を各時間帯についてそれぞれ独立に解いている。日生成交通需要の変動係数が固定されていることから、各時間帯における尤度最大化問題はそれぞれ独立となるためである。

4. 交通状態の事後分布の推定

前節では、不完全に観測されたデータに基づき、道路ネットワーク全体の交通状態を最尤法によって推定する手法を構築した。本節では、過去の観測データから推定された交通状態の確率分布を事前分布として、プローブカーから得られるリンク遅れ時間および交通感知器から得られる交通量データを尤度とするベイズ推定を行い、交通状態の事後分布を推定することを考える。ここで新たに得られる観測データは、リアルタイムに得られたものと想定する。

前述のように、プローブデータは普及率の低さから、すべてのリンクを観測できないため、欠損を持つデータとなる。Dominici et al.⁹⁾は、多変量正規分布を対象とする観測データの欠測を考慮可能なベイズ推定法を提案した。本研究では、対数正規分布のパラメータを用いて、多変量正規分布を記述し、これに同手法を適用する。つまり、多変量正規分布を推定することで、対応する対数正規分布のパラメータを推定し、対応する対数正規分布を得る。先述のように、パラメータを用いて、対数正規分布本来の平均及び分散を得ることは容易である。

前節の最尤推定法から推定される多変量対数正規分布に従う交通状態を事前分布とする。

$$\mu_{pri} | \Sigma_{pri} \sim MVN(\mu_{MLE}, q^{-1} \Sigma_{MLE}) \quad (38)$$

ここで、以下の関係が成立する。

$$\Sigma_{pri} = \Sigma_{MLE} \quad (39)$$

ここで、 q は、事前分布の確からしさを表す指標である。 μ_{MLE} と Σ_{MLE} はそれぞれ交通状態の平均ベクトルと分散共分散行列である。Dominici et al. (2000)⁹⁾は、多変量正規

分布に従うある確率変数の分散共分散行列が既知であり、その平均パラメータの事後分布 \tilde{m} は、以下に示す多変量正規分布に従うことを指摘している。

$$\mu_{post} | \Sigma_{post}, D \sim MVN(\tilde{m}, \tilde{B}) \quad (40)$$

ここで、以下の関係が成立する。

$$\tilde{B} = ((q + 1)M^T(M\Sigma_{pri}M^T)^{-1}M)^{-1} \quad (41)$$

$$\tilde{m} = \tilde{B}(q + 1)M^T(M\Sigma_{pri}M^T)^{-1}MM^T\tilde{m}_D \quad (42)$$

$$\tilde{m}_D = \frac{q\mu_{pri} + \hat{d}}{q + 1} \quad (43)$$

M は(33)の M_g と同様の考え方によって定義される行列である。 \hat{d} は、対数をとったリアルタイムに観測されたデータである。

Dominici et al. (2000)⁹⁾では、複数回の観測データを用いて事後分布を更新している。本研究においては、リアルタイムに観測されるデータセットは1つと仮定し、事後分布の推定を行う。

本研究のベイズ推定で扱う交通状態の事前分布の分散共分散行列は正則でない。ベイズ推定の過程で、逆行列を計算するためには、事前分布の分散共分散行列を正則化する必要がある。赤穂(2008)⁶⁾では、正則でない分散共分散行列に対して、重みづけした単位行列を足して行列を正則化する手法を紹介している。これを本研究のベイズ推定に適用し、事前分布の分散共分散行列を正則化する。

また、前節で示した最尤推定では、蓄積された観測データを活用して交通状態が推定されており、観測データ数が多いほど、推定される平均や分散の精度は向上する。しかしながら、本節で説明しているベイズ推定による交通状態の事後分布推定に関しては、前述で示したように新たに観測されるデータセットは1つである。データの観測が、ある特定のリンクに偏ることが前提とされている。そのため、ベイズ推定においては、データ計測時に発生し得る計測誤差が推定結果に与える影響が、最尤推定のとくと比べて、大きいと考えられる。

以上のことから、本研究では、事前分布の分散共分散行列を正則化するために、単位行列を微小なスカラーで重みづけしたものを誤差項として以下のように定義する。

$$\xi_{ti} \sim MVN(0, \Sigma_{Error}) \quad (44)$$

ここで、 Σ_{Error} は、分散、つまり対角成分のみ値を持つ単位行列である。計測機器間の相関、すなわち Σ_{Error} の共分散はないと考える。

よって、ベイズ推定に用いる事前分布の分散共分散行列 Σ_{pri} は、最尤推定によって算出された分散共分散行列に、誤差項 Σ_{Error} を付加し、次のように表される。

$$\Sigma_{pri} = \Sigma_{pri} + \Sigma_{Error} \quad (45)$$

全時間帯全道路ネットワークの交通状態の事後分布を推定するアルゴリズムを以下に示す。

- Step 0 観測された交通データから、最尤推定によって交通状態を推定する
- Step 1 STEP0で推定された交通状態を事前分布として、リアルタイムに観測されたデータを用いてベイズ推定を行い、交通状態の事後分

布 $MVN(\tilde{\mathbf{m}}, \tilde{\mathbf{B}})$ を推定する

Step2 STEP1で推定された交通状態の事後分布の分散共分散 $\tilde{\mathbf{B}}$ を次の時間帯の事前分布の分散共分散として共有し、再度ベイズ推定を用いて交通状態の事後分布 $MVN(\tilde{\mathbf{m}}, \tilde{\mathbf{B}})$ を推定する。

5. 数値計算

(1) 設定

使用するネットワークは、図-1に示すYang et al.⁷⁾のもので、それぞれODペア数8、経路数32、リンク数12である。各リンクの交通容量の平均は1000 [pcu/hour]、変動係数は0.2とする。観測データのサンプルを300組用意する。これらのサンプルは、事前に未知変数を所与として交通量配分問題を解き、均衡解として得られたリンク交通量とリンク遅れ時間を真値として、ランダムに発生させたものである。リンクコスト関数のパラメータ、 t_a^0 と γ_a はそれぞれ2と6とした。なお、プローブカー混入率は数値計算の便宜上、ネットワーク全体において一様であると仮定した。なお、今回の数値計算ではリンク1,2,5,7,9に交通感知器が設置されているものと仮定し、プローブカー混入率は0.3としている。すべてのリンクで、ある観測機会において、0.3の確率でリンク遅れ時間が観測されることを意味する。一方で、現実には、リンクごとにプローブカー混入率が異なることが知られている。本研究で提案する手法では、リンクごとに異なるプローブカー混入率を考慮することも可能である。

正午から午後6時にかけて、6の時間帯における各ODペアにおける交通需要の平均と分散・共分散を真値として設定した。このOD交通需要を所与として、交通量配分問題を解き、ネットワーク全体のリンク交通量とリンク移動時間の多変量分布を得る。続いて、この多変量分布の12時から17時までのリアルタイムデータをランダムにサンプリング、およびプローブデータに相当するデータを欠測させる作業を行った。そして、この作業の結果得られたサンプルを疑似データとして用いて最尤推定を行った。

本計算では、最尤推定によって得られた交通状態の確率分布を事前分布として、交通感知器から得られる交通量データを尤度とするベイズ推定を行い、交通状態の事後分布を推定する。図-2は、交通状態の事前分布の平均を示したもので、各時間帯のリンク交通量の平均がプロットされている。また、交通需要変動係数 cv は0.2024である。

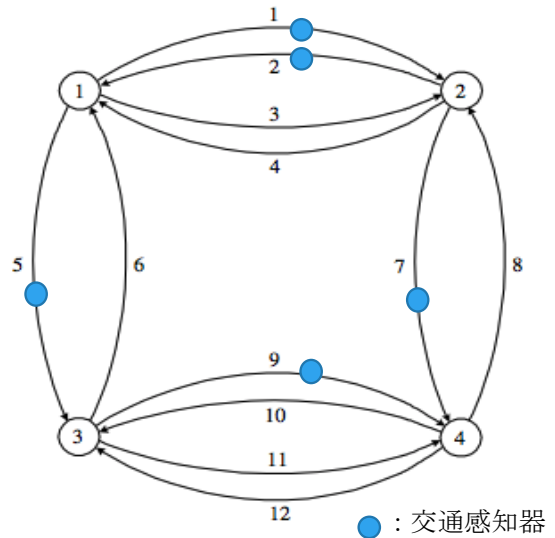


図-1 テストネットワーク

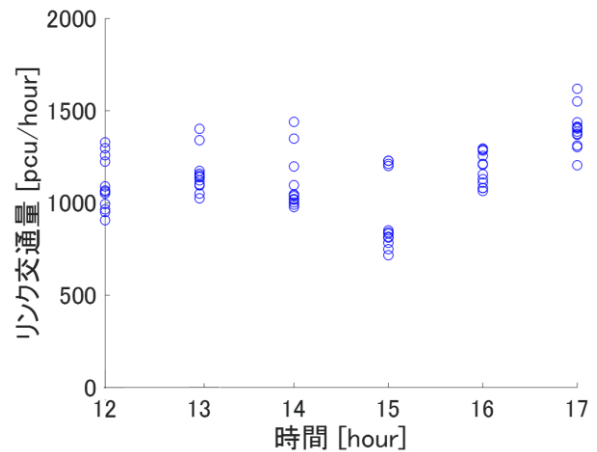


図-2 リンク交通量の事前分布の平均

(2) 結果

4章で示したアルゴリズムに従って、(1)でサンプリングした12時から17時までのデータを用いてベイズ更新を行った。図-3は、12時と17時におけるリンク1のリンク交通量に着目し、リンク交通量の事前分布と事後分布の確率分布の変化を示したものである。ベイズ更新を行ったことにより、リンク交通量の平均は変化し、分散は事前分布と比較して減少した。また、4章でのアルゴリズムで示したように、今回の計算では、前の時間帯の分散共分散を次の時間帯の分散共分散に共有している。よって、12時と比較して17時のリンク交通量の分散は小さくなった。このことから、複数時間帯での更新を行うと、分散は徐々に減少することが確認された。

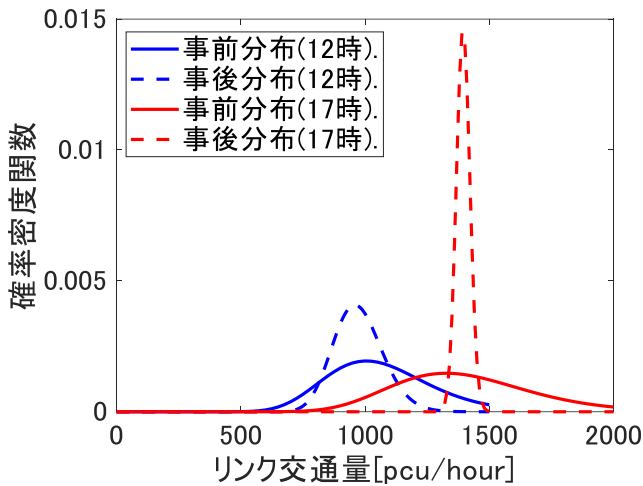


図3 リンク1の確率分布の推移

次に、交通感知器が設置されていないリンク4に注目する。図4は、12時と17時におけるリンク4のリンク交通量の事前分布と事後分布の確率分布の変化を示したものである。リンク1で示した結果と同様に、リンク交通量の平均が変化し、分散は事前分布と比較して減少し、複数時間帯で更新を行うほど更に減少した。前述のように、リンク4には交通感知器が設置されていないため、リンク交通量は、観測データが存在しない。しかしながら、リンクベースのモデルに対し、ベイズ推定を適用することで、直接観測されていないリンクに関しても、新たな確率分布を推定することができた。

以上の事から、新たな観測データが存在しないリンク事後分布は、他のリンクで観測されたデータによって更新され、計算結果より明らかなように、分散が減少し、更新の前後で分布形状が変化することが確認できる。また、直接観測されない変数に関して、変数間の相関関係を用いて、推定することが可能であることが示された。

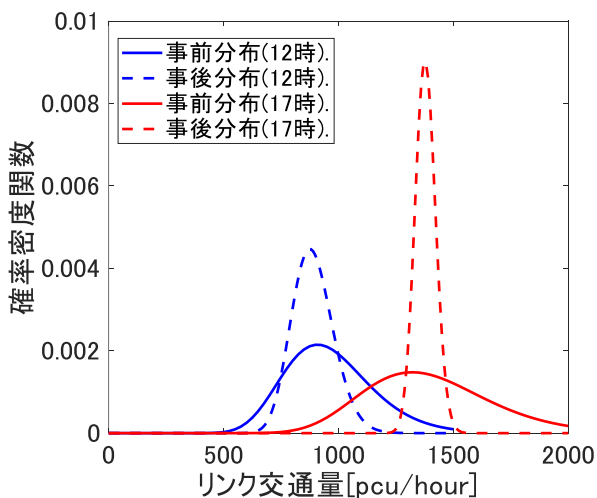


図4 リンク4の確率分布の推移

6. まとめ

本研究では、複数の時間帯にわたって観測された、プローブカーデータと交通感知器データを用いて、全時間帯全道路ネットワークにおける交通状態を確率的に推定し、推定された交通量データを事前分布とするベイズ推定を行い、交通状態の事後分布をリアルタイムに推定する手法を提案した。ここで、最尤推定によって得られた交通状態の事前分布は、部分的に観測された交通データの使用を前提として、交通状態の尤度を最大化することで求めた。同様に、事後分布の推定においても、リアルタイムに観測される交通データの不完全性を前提として、交通状態の事後分布を推定する手法を提案した。

数値計算として、提案するアルゴリズムに基づいて、6つの時間帯の最尤推定によって求めた交通状態を事前分布として、ベイズ推定によって事後分布を推定した。すべての時間帯において事前分布と比較して事後分布の分散が減少し、分布形状が変化した様子を確認した。

今後の課題として、同様の手法でリンク遅れ時間の事後分布を推定することが挙げられる。さらに、より複雑な道路ネットワークを対象とした数値計算による検証が必要である。

参考文献

- 1) 原祐輔, 花岡洋平, 桑原雅夫. 道路ネットワーク内の関係性に着目した長期観測プローブデータによるプローブ未観測リンクの交通状態補間, 交通工学論文集, Vol. 2, No. 1, pp. 1-10.
- 2) Hara, Y., Suzuki, J. and Kuwahara, M.: Network-wide traffic state estimation using a mixture Gaussian graphical model and graphical lasso. *Transportation Research Part C: Emerging Technologies*, Vol. 86, pp. 622-638, 2018.
- 3) Shao, H., Lam, W.H.K., Sumalee, A. and Hazelton, M.L. Estimation of mean and covariance of stochastic multi-class OD demands from classified traffic counts. *Transportation Research Part C*, Vol. 59, pp. 92-110, 2015.
- 4) 峪 龍一, 内田 賢悦: プローブデータと感知器データを組み合わせた交通状態の時空間推定, 第62回土木計画学研究発表会(2020)
- 5) Dominici, F., Parmigiani, G. and Clyde, M.: Con-jugate analysis of multivariate normal data with incomplete observations. *The Canadian Journal of Statistics*, Vol. 28, No. 3, pp. 533-550, 2000.
- 6) 赤穂 昭太郎: カーネル多変量解析—非線形データ解析の新しい展開(シリーズ確率と情報の科学)(2008)
- 7) Yang, H., Zhang, X. and Meng, Q. Stackelberg games and multiple equilibrium behaviors. *Transportation Research Part B*, Vol. 41, pp. 841-861, 2007.

(2021.10.1 受付)

Real-time update of road network traffic conditions using incomplete traffic data

Yuto KAWAMURA, Ryuichi TANI and Kenetsu UCHIDA

In this study, we propose a method for estimating the link traffic volume and link delay time of the entire road network at multiple time periods as random variables using data observed from traffic detectors and probe cars. We assume that the link traffic volume and the delay time follow a multivariate lognormal distribution, and estimate the traffic conditions by solving the likelihood maximization problem for the observed link traffic volume and the link delay time. In this model, the likelihood is calculated considering the incompleteness of the observation points of the traffic observation data. We propose a method to estimate the posterior distribution of the traffic state by using the probability distribution of the traffic state estimated by maximum likelihood estimation as the prior distribution and Bayesian estimation using the traffic volume data obtained from traffic detectors and the link travel time obtained from probe data as the likelihood. The proposed method is based on the Bayesian estimation of the posterior distribution of the traffic conditions. Finally, we show the validation results of the proposed model using test data.