

交通観測データの欠測を考慮した複数時間帯における 道路ネットワーク交通状態推定

Traffic state estimation in road networks with multiple time zones using incomplete observation data

北海道大学工学部

北海道大学大学院工学院

北海道大学大学院工学研究院

○学生員

学生員

正員

川村雄斗 (Yuto Kawamura)

峪 龍一 (Ryuichi Tani)

内田賢悦 (Kenetsu Uchida)

1. はじめに

ネットワーク上で観測される交通データは、基本的に不完全性を有する。交通感知器は、固定的かつ連続的に交通量を観測することができる。しかし、運用費用が高いことから、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。ETC2.0 に代表されるプローブカーデータは、各車両の挙動を軌跡として遅れ時間を観測できるが、車載器の普及率が低いデータ数が限られている。また、直接的に交通量を観測することはできない。以上のことから、現在の道路交通において、道路ネットワーク全体の交通状態を直接観測することは困難である。

本研究では、日々観測される交通データから、全時間帯のすべてのリンクの交通量と遅れ時間を確率的に推定する手法を提案する。提案する手法は、道路ネットワークの全リンクの交通量と遅れ時間を多変量対数正規分布で表現し、観測されたリンク交通量と遅れ時間について、尤度を最大化することによって、交通状態を推定するのである。

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おり、峪・内田 (2020)¹⁾は、交通量配分問題の定式化及び特殊なネットワークの単一時間帯における数値計算の結果を示した。本研究では、峪・内田(2020)¹⁾によって定式化された交通量配分問題に基づき、任意の道路ネットワークに対応した、複数時間帯を対象とする交通状態推定手法を構築し、数値計算による手法の検証を行う。

2. モデル

2.1 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、その確率的な変数については小文字で表すこととする。

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
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_i	OD ペア i 間の確率的交通需要(OD 交通量)

F_{ij}	OD ペア i 間の経路 j の確率的交通量(経路交通量)
V_a	リンク a の確率的交通量 (リンク交通量)
V_{ab}	リンク a 、 b の両方を通過する確率的交通量
p_i	OD ペア i 間の交通需要の配分率
p_{ij}	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
$c_{ij}(\mathbf{F})$	OD ペア i 間の経路 j の一般化費用
t_a^0	リンク a の自由走行時間
C_a	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_{ij}	OD ペア i の経路 j の確率的移動時間(経路移動時間)

2.2 交通状態のモデル化

(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。OD 交通量 Q_i は、OD 配分率と総交通需要 Q の積として表現される。

$$Q_i = p_i \cdot Q \quad \forall i \in I \quad (2)$$

OD 交通需要 Q_i は対数正規分布に従い、そのパラメータは、総交通需要 Q のパラメータを用いて表現される。

$$Q_i \sim LN(\mu_{Q_i}, \sigma_{Q_i}^2) \quad \forall i \in I \quad (3)$$

ここで、

$$\mu_{Q_i} = \mu + \ln(p_i) \quad \forall i \in I \quad (4)$$

$$\sigma_{Q_i}^2 = \sigma^2 \quad \forall i \in I \quad (5)$$

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

$$F_{ij} = p_{ij} \cdot Q_i = p_{ij} \cdot p_i \cdot Q \quad \forall i \in I, \forall j \in J_i \quad (6)$$

$$F_{ij} \sim LN(\mu_{F_{ij}}, \sigma_{F_{ij}}^2) \quad \forall i \in I, \forall j \in J_i \quad (7)$$

ここで、

$$\mu_{F_{ij}} = \mu + \ln(p_{ij} \cdot p_i) \quad \forall i \in I, \forall j \in J_i \quad (8)$$

$$\sigma_{F_{ij}}^2 = \sigma^2 \forall i \in I, \forall j \in J_i \quad (9)$$

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。

$$\begin{aligned} V_a &= \sum_{i \in I} \sum_{j \in J_i} \delta_{aj} \cdot F_{ij} \\ &= \sum_{i \in I} \sum_{j \in J_i} \delta_{aj} \cdot p_{ij} \cdot Q_i = \hat{p}_a \cdot Q \quad \forall a \in A \end{aligned} \quad (10)$$

ここで、

$$\begin{aligned} \hat{p}_a &= \sum_{i \in I} \sum_{j \in J_i} \delta_{aj} \cdot p_{ij} \cdot p_i \quad \forall a \in A \\ &= \sum_{i \in I} \sum_{j \in J_i} \delta_{aj} \cdot p_{ij} \cdot Q_i = \hat{p}_a \cdot Q \quad \forall a \in A \end{aligned} \quad (11)$$

以上の定式化によって、リンク交通量はスカラー量 $\hat{p}_a$ と総交通需要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、そのパラメータは総交通需要 Q のパラメータを用いて表現される。

$$V_a \sim LN(\mu_{V_a}, \sigma_{V_a}^2) \quad \forall a \in A \quad (12)$$

ここで、

$$\mu_{V_a} = \ln(\hat{p}_a) + \mu \quad \forall a \in A \quad (13)$$

$$\sigma_{V_a}^2 = \sigma^2 \quad \forall a \in A \quad (14)$$

(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。

$$T_a(V_a, C_a) = t_a^0 \left(1 + \gamma_a \cdot \left(\frac{V_a}{C_a} \right)^{\lambda_a} \right) \quad \forall a \in A \quad (15)$$

式(15)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。

$$T_a(V_a, C_a) = t_a^0 + \frac{t_a^0 \cdot \gamma_a}{C_a^{\lambda_a}} \cdot (V_a)^{\lambda_a} \quad (16)$$

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。

$$\Xi_{ij} = \sum_{a \in A} T_a \cdot \delta_{aj} \quad \forall i \in I, \forall j \in J_i \quad (17)$$

(3) 均衡配分モデル

経路交通量は、一般化費用 $c_{ij}(\mathbf{F})$ を評価指標として、確率的利用者均衡配分モデルから推計される経路選択確率 p_{ij} を用いて、以下に示す均衡条件を満たすものとする。

$$E[F_{ij}] = p_{ij} \cdot c(\mathbf{F}) \cdot E[Q_i] \quad \forall i \in I, \forall j \in J_i \quad (18)$$

$$p_{ij} = \frac{\exp(-\phi \cdot c_{ij}(\mathbf{F}))}{\sum_{j=1}^{|J_i|} \exp(-\phi \cdot c_{ij}(\mathbf{F}))} \quad \forall i \in I, \forall j \in J_i \quad (19)$$

ここで、

$$\mathbf{c} = (c_{11}(\mathbf{F}), \dots, c_{|I| \cdot |J_{|I|}|}(\mathbf{F})) \quad (20)$$

$$\mathbf{F} = (F_{11}, \dots, F_{|I| \cdot |J_{|I|}|}) \quad (21)$$

一般化費用 $c_{ij}(\mathbf{F})$ 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。

$$c_{ij}(\mathbf{F}) = E[\Xi_{ij}] + \eta \cdot \text{var}[\Xi_{ij}] \quad \forall i \in I, \forall j \in J_i \quad (22)$$

2.3 交通状態の推定手法

本研究で推定の対象とする交通状態はリンク交通量 [pcu/hour] とリンク遅れ時間 [hour] である。

本節では、ある特定の時間帯に道路ネットワーク全体のうち、交通感知器とプローブカーによって、過去 G 日間、部分的に交通状態が観測されている場合を想定して、道路ネットワーク全体の交通状態を推定する手法を述べる。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 $n(g)$ とする

と、道路ネットワーク全体における交通状態に対応する多変量分布は、以下に示すデータの欠測を考慮した、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計することができる。

$$L(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \quad (23)$$

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

ここで、

$$\boldsymbol{\Sigma}_g = \mathbf{M}_g \boldsymbol{\Sigma} \mathbf{M}_g^T \quad (24)$$

$$\boldsymbol{\mu}_g = \mathbf{M}_g \boldsymbol{\mu}^T \quad (25)$$

$$\hat{\mathbf{d}}_g = \mathbf{M}_g \mathbf{d}_g^T \quad (26)$$

$\mathbf{d}_g$ 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。

$$\mathbf{d}_g = (\hat{v}_1 \ \hat{t}_1 \ \dots \ \hat{v}_a \ \hat{t}_a \ \dots \ \hat{v}_{|A|} \ \hat{t}_{|A|}) \quad (27)$$

$\mathbf{M}_g$ は、観測データ $\mathbf{d}_g$ 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、 $\mathbf{d}_g$ 及び $\mathbf{M}_g$ は以下のように表される。

$$\mathbf{d}_g = (\hat{v}_1 \ \hat{t}_1 \ - \ \hat{t}_2 \ - \ -) \quad \forall g \in \{1, \dots, G\} \quad (28)$$

$$\mathbf{M}_g = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix} \quad (29)$$

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る。これを上記の尤度関数に $\hat{\mathbf{d}}_g$ として導入することで、データの欠測を考慮した、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2.2 (1)-(3)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計値は多変量対数正規分布に従い、 $\mathbf{X} \sim LN(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ 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、 $\ln(\mathbf{X}) \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ 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したときの推定値と、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し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、多変量正規分布に従う、対数変換した交通状態を用いて最尤推定を実施し、パラメータを推定するものである。

2.4 複数時間帯における交通状態推定への拡張

上記までの議論では、単一の時間帯における交通状態を推計する問題を扱ってきた。本節では、複数の時間帯、即ち、1日全体の交通状態を推定する方法を構築する。

対数正規分布に従う日生成交通量を定義し、時間変動係数によって時間帯別生成交通量 Q_h が求まることを仮定する。

$$Q_h \sim LN(\mu_h, \sigma_h^2) \quad \forall h \in H \quad (30)$$

ここで、日生成交通量のパラメータ σ^2 を所与としたときに、各時間帯別の尤度 L_h を最大化すると、すべての時間帯における尤度の積 $\prod_h L_h$ も最大化される。上記の内容は、各時間帯における各 OD 交通需要の平均ベクトル μ_{Q_i} に関する尤度最大化問題となる。

上記では、パラメータ σ^2 が所与である場合を考えた。次は反対に、尤度関数 $\prod_h L_h$ について、各時間帯における OD 交通需要の平均ベクトル μ_{Q_i} を所与としたときに、 σ^2 を求める問題を考える。尤度を最大化する σ^2 を σ_*^2 と表すと $\frac{\partial \prod_h L_h}{\partial \sigma^2} \Big|_{\sigma^2=\sigma_*^2} = 0$ という関係を得る。

このように、尤度関数 $\prod_h L_h$ に対して、2つの問題を考えることができる。これらの関係を利用して、日生成交通需要あるいは各時間帯別の生成交通量の各種パラメータを推定するためのアルゴリズムを説明する。概要は以下の通りである。

STEP 1: (各時間帯の各 OD 交通需要の平均 μ_{Q_i} を求める)

生成交通需要のパラメータ σ^2 を所与とし、各時間帯 h における $\prod_h L_h(\mu_{Q_i}|\sigma^2)$ の尤度最大化問題を解き、 μ_{Q_i} を推定する。

STEP 2: (日生成交通量の σ^2 を求める)

STEP1 で求めた各時間帯 h における μ_{Q_i} を所与とし、全時間帯における $\prod_h L_h(\sigma^2|\mu_{Q_i})$ の尤度最大化問題を解き、 σ^2 を推定する。

STEP1 では、1日全体の全時間帯の交通状態を同時に解く必要がなく、各時間帯別の問題を個別に解くことを繰り返すことによって、1日全体の交通状態を推計することが可能となる。

STEP2 では、全時間帯の交通状態を同時に解くことで、全時間帯全ネットワークに対して1つの変数だけを推定する問題となる。

このように、全時間帯全ネットワークの問題を解きやすい小さな問題に分割して、これを順番に解くのがこのアルゴリズムの特徴である。

3. 数値計算

3.1 条件設定

使用するネットワークは、図-1に示す Yang et al.²⁾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 3000 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。また、交通感知器の設置箇所をリンク 1、2、5、7、9、とし、プローブカーの混入率を全時間帯全ネットワークにおいて 30% と設定する。

以下に示す 3 種類の数値計算を行った。それぞれで未知変数として設定する内容が異なるので、以下にまとめて記載する

① 単一の時間帯における交通状態推定

ここでの未知変数は OD 交通需要の平均 (1000 [pcu/hour]) である。最尤推定では時間帯別生成交通量から各 OD ペアに配分する交通需要の配分率は全 OD ペアにおいて固定とする。

② 複数時間帯における交通状態推定 (STEP1)

ここでの未知変数は OD 交通需要の平均 (各時間帯で個別に設定) である。各時間帯における全 OD ペアの交通需要の平均を足し合わせて、時間帯別生成交通量として表現したものを図-2に示す。

③ 複数時間帯における交通状態推定 (STEP2)

ここでの未知変数は生成交通需要の変動係数である。

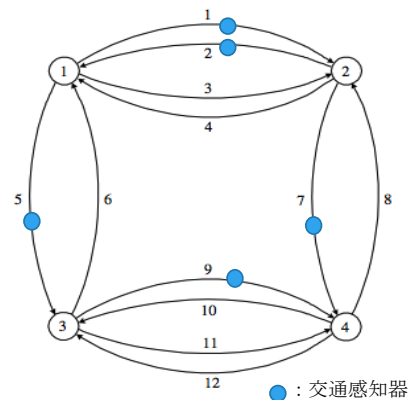


図-1 テストネットワーク (Yang et al)

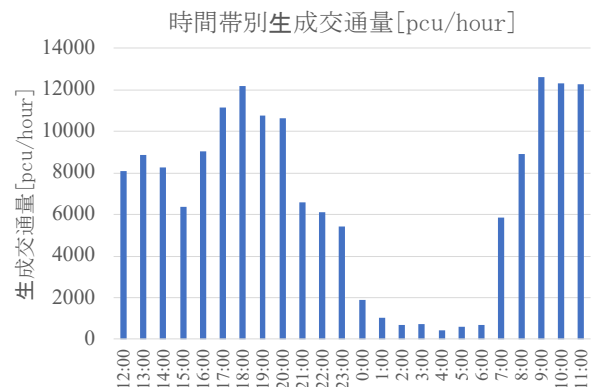


図-2 時間帯別生成交通量

ここでは、上記の未知変数を基に、交通量配分問題を解き、均衡解としてのリンク交通量とリンク遅れ時間を得る。そして、サンプリングされたリンク交通量とリンク遅れ時間について最尤推定を行う。

未知変数の初期設定値（真値）と、上記の交通量配分問題を制約として最尤推定される未知変数の推定値を比較したものを、次節で計算結果としてそれぞれ示す。

3.2 結果

(1) 単一の時間帯における交通状態推定

図-3 に、交通需要の平均と変動係数をそれぞれ変化させた時の交通状態（リンク交通量とリンク遅れ時間）についての対数尤度の変化を示す。真値 [1000,0.2] 付近で尤度が最大化される様子がわかる。このことから、交通観測データの空間的な欠測を考慮した場合でも、高い推定精度が得られることが確認できた。

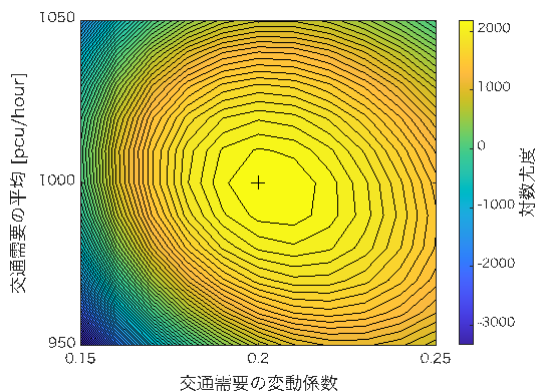


図-3 交通需要の平均と変動係数をそれぞれ変化させたときの尤度の変化

(2) 複数時間帯における交通状態推定 (STEP1)

図-4 は全時間帯全 OD ペアの交通需要の真値と推定値の関係を表す散布図である。決定係数は 0.9964、相関係数が 0.9982 となり、高い相関が得られた。

図-5 は各時間帯の OD 交通需要の二乗平均平方根誤差を示している。各時間帯により誤差にばらつきが生じた。これは、各時間帯で、異なるサンプルを用いて最尤推定を行ったためである。つまり、各データサンプルセットに由来する精度のばらつきが生じていると考えられる。

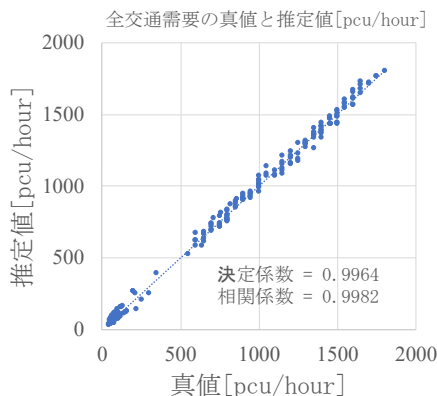


図-4 全交通需要の真値と推定値

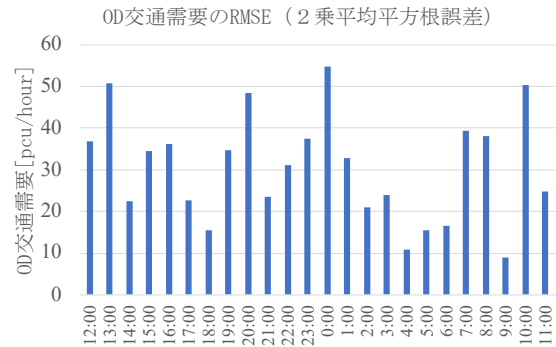


図-5 OD 交通需要の2乗平均平方根誤差

(3) 複数時間帯における交通状態推定(STEP2)

図-6 に、交通需要の変動係数を変化させた時の交通状態についての対数尤度の変化を示す。尤度最大化問題を解いて得られた最尤推定値は 0.203 となり、図-6 から、真値 0.2 付近で尤度が最大化される様子がわかる。

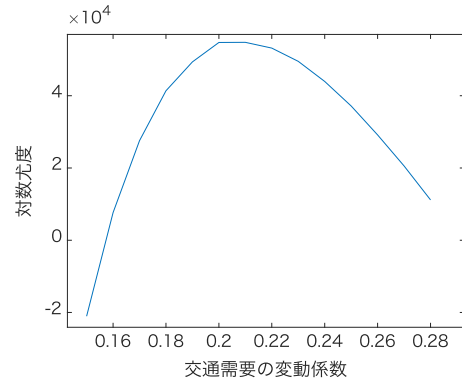


図-6 交通需要の変動係数を変化させたときの尤度の変化

4. まとめ

本研究では、プローブカーデータと交通感知器データを用いて、道路ネットワーク全体の移動時間と交通量について、異なるリンク同士の交通状態を推定する手法を提案した。数値計算は、提案したモデルの挙動を確かめることを主目的として、簡易なネットワークを対象にサンプル数を多く用意したため、良好な結果が得られた。

今後は、プローブカーの混入率を更に低く設定し、より少ないサンプル数における推定精度を検証し、より現実問題に近いネットワークでの計算を行う必要がある。一方、ベイズ推定を用いて、過去の交通観測から推定された確率的な交通状態を用いることで、現在の交通状態をリアルタイムに更新することができる。これらは今後の課題としたい。

参考文献

- 1) 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)
- 2) Yang et al, and Meng, Q. : Stackelberg games and multiple equilibrium behaviors on networks. Transportation Research Part B : Methodological, 41 (8), 841-861 (2007).