プローブ・トラカンデータを用いた最尤法による交通状態推計

中山晶一朗 [1]

[1] 正会員 博（工） 金沢大学環境デザイン学系（〒920-1192 金沢市角間町）

近年，プローブデータの蓄積が進んでおり，その活用が期待されている．プローブデータの利用で多い事例の一つにリンク速度の推定がある．リンク速度の推定であれば，対象となる時間でのサンプル数が比較的少なくともある程度の推定は可能となると思われる．サンプル率が低いため，あるリンクを通過したプローブ車両数からそのリンクの交通量を推定すると大きな誤差が生じるが，リンク速度の情報も加味すると，その誤差をある程度補正することが可能と期待できる．さらに，トラカンデータも併用することで，さらに交通量の推定精度を上げることができる．そこで，本研究では，プローブデータとトラカンデータの両方を用いて交通量や速度の推定を行う手法を開発することが本研究の目的である．

Key Words: *probe data, traffic counter data , maximum likelihood*

1. はじめに

近年，プローブデータの蓄積が進んでおり，その活用が期待されている．プローブデータは，データの総量は膨大であるものの，サンプル率が低く，偏っていることがあることや費用や利用制限等で利用できるプローブデータは限られることなどの課題がある．研究者が利用可能なETC2.0 などのプローブデータはサンプル率が低く，また，高速道路や国道などの走行に偏っている．サンプル率が低いため，あるリンクを通過したプローブ車両数からそのリンクの交通量を推定すると大きな誤差が生じると考えられる．プローブデータの利用で多い事例の一つにリンク速度の推定がある．リンク速度の推定であれば，対象となる時間でのサンプル数が比較的少なくともある程度の推定は可能となると思われる．サンプル率が低いため，あるリンクを通過したプローブ車両数からそのリンクの交通量を推定すると大きな誤差が生じると上述したが，リンク速度の情報も加味すると，その誤差をある程度補正することが可能と期待できる．

さらに，トラカンデータも併用することで，さらに交通量の推定精度を上げることができる．トラカンの観測地点は限られるが，観測地点の交通量は比較的高い精度で計測することができる．プローブデータとトラカンデータの両方を用いて交通量や速度の推定を行う手法を開発することが本研究の目的である．

2. 確率的な交通量

交通量は日々，時々刻々と変動しており，それを確率的に扱うことは自然である．プローブデータを用いる場合，上述の通り，旅行速度を分析する事例が多い．プローブから得られる旅行速度から，KV 曲線等を仮定して，交通量や交通密度を推定することができる．しかし，旅行速度の観測自体などに誤差を含むため，これも確率的に扱う必要がある．このような観測誤差のみを確率的に扱う場合，リンク間の相関は独立となる．つまり，観測誤差は独立と仮定することが自然であり，観測誤差だけを確率的に扱うと，（観測）リンク交通量は独立となる．しかしながら，実際のリンク交通量は直近の上流下流などでは相関が高いと考えられ，そのような相関関係を考える方が自然であるとともに，それを考えることで推定誤差を低減することが期待できる．本研究では，リンク交通量の相関を考えるが，交通需要などが確率的変動するため，その確率変動からリンク交通量間の分散共分散もしくは相関行列を仮定する．

（1） 確率的経路選択

まず，経路選択のみ確率的に扱うことを考える．本研究では，経路選択はロジットモデルによって与える．OD ペア i (= 1, 2,…, I) の各道路利用者が経路 j (= 1, 2,…, J_i) を選択する確率 p_{ij} は以下の式の通りである．

$$p_{ij} = \frac{\exp(-\theta c_{ij})}{\sum_{j'=1}^{J_i} \exp(-\theta c_{ij'})} \tag{1}$$

ここで，c_{ij} は旅行時間（平均旅行時間）であり，θ は正のパラメータである．このように本研究では，ロジットモデ

ルによる経路選択に基づく交通均衡の成立を前提とする．

上記のロジットモデルによって与えられる経路選択確率 $\mathbf{p}_i = (p_{i1},.., p_{ij},.., p_{iJ_i})^{\mathrm{T}}$ の確率の通りに道路利用者の経路選択を確率的に行わせる．ここで，T は転置を表し，$\sum_j p_{ij} = 1$ である．本節では，OD 交通量が与えられた場合の経路交通量の確率分布について考察する．OD ペア i の OD 交通量が n_i として固定的に与えられた場合の（OD ペア i の）経路交通量の（条件付）同時生起確率は以下の多項分布として与えられる．

$$f_{\mathbf{X}_i|n_i}(\mathbf{x}_i) = \frac{n_i!}{\prod_{j=1}^{J_i} x_{ij}!}\prod_{j=1}^{J_i} p_{ij}^{\ x_{ij}} \tag{2}$$

ここで，$f_{\mathbf{X}_i|n_i}(\mathbf{x}_i)$ は OD ペア i の OD 交通量が n_i の場合の経路交通量の同時生起確率，$\mathbf{X}_i|n_i$ は OD ペア i の J_i 次元の経路交通量の確率変数ベクトルで，その要素は $X_{ij}|n_i\ (j = 1, 2,\ldots, J_i)$ である．また，$\mathbf{x}_i$ は OD ペア i の J_i 次元の経路交通量の実現値のベクトルで，その要素は $x_{ij}\ (j = 1, 2,\ldots, J_i)$ である．

ある一つの経路の経路交通量は，式 (2) の同時生起確率の周辺確率として与えることが可能であり，以下の通り，それは次式の二項分布 $\mathrm{Bn}(n_i,p_{ij})$ となる．

$$\begin{aligned} f_{X_{ij}|n_i}(\mathbf{x}_i) &= \sum_{x_{i1}}\cdots\sum_{x_{ij-1}}\sum_{x_{ij+1}}\cdots\sum_{x_{iJ_i}} f_{\mathbf{X}_i|n_i}(\mathbf{x}_i) \\ &= \frac{n_i!}{x_{ij}!(n_i - x_{ij})!}\, p_{ij}^{\ x_{ij}}\left(1-p_{ij}\right)^{n_i - x_{ij}} \end{aligned} \tag{3}$$

なお，上の式は，ある着目する経路が j の時の交通量の確率分布である．同じ OD ペア内の経路 j 以外の経路交通量の和は，n_i から経路 j の交通量を引いたものであり，それも二項分布 $\mathrm{Bn}(n_i, 1-p_{ij})$ となる．式 (3) では，経路交通量間の共分散を0と考えているのではなく，経路 j のみに着目した場合の経路 j の交通量の分布を見たものである．

多項分布の性質から，経路 j と経路 j' の交通量の共分散は $-n_i\, p_{ij}\, p_{ij'}$ となる．しかし，上で述べたようにポアソン分布を仮定する場合，経路選択確率 p_{ij} 及び $p_{ij'}$ は小さいと仮定しているため，微小項の二乗は無視できるとし，経路交通量の共分散は 0 で近似できる．つまり，ポアソン分布を仮定することは経路交通量の共分散は考えないことを意味する．ただし，それは，リンク交通量の共分散（相関）を考慮しないことを意味しているものではなく，配分の際には，リンク交通量の共分散（相関）は考慮されなければならない．

次節で述べるが，経路選択確率がそれほど小さくなく，平均交通量が大きい場合，経路交通量は正規分布により近似することが出来る．この時，共分散は 0 で近似できるとは限らないため，全経路の交通量は多変量正規分布に従うことになる．

（2）確率的交通需要

OD ペア i の n_i 人の利用者が経路選択確率 $\mathbf{p}_i$ で確率的に経路を選択する場合は，既に述べたように経路交通量は多項分布 $\mathrm{Mn}(n_i, \mathbf{p}_i)$ に従う．このように（交通需要が所与の場合の）条件付経路交通量が多項分布に従う場合についても，交通需要（OD 交通量）の確率変動を同時に考慮することも可能である．OD 交通量が負の二項分布に従うと仮定すると，以下に示すように経路交通量は負の多項分布となる．つまり，負の二項分布と多項分布の複合を考えることを意味する．

OD交通量は以下のような確率関数を持つ負の二項分布と仮定する．

$$g_{N_i}(n_i) = \frac{\Gamma(\alpha_i + n_i)}{\Gamma(\alpha_i)\, n_i!}\left(\frac{1}{1+\beta_i}\right)^{\alpha_i}\left(\frac{\beta_i}{1+\beta_i}\right)^{n_i} \tag{4}$$

ここで，N_i は負の二項分布に従う OD ペア i の OD 交通量の確率変数，$g_{N_i}(n_i)$ は N_i が従う確率関数である．なお，この負の二項分布の平均及び分散はそれぞれ $\alpha_i\,\beta_i$ 及び $\alpha_i\,\beta_i(1+\beta_i)$ となる．ポアソン分布や二項分布では，分散を自由に設定することができないが，この負の二項分布では，α_i 及び β_i を変えることによってそれを自由に設定することができる．

ここで，負の多項分布の確率関数を取り上げる．負の多項分布の確率関数に $n_i = \sum_j x_{ij}$ を代入し，整理すると，次式を導くことができる．

$$\begin{aligned} f_{\mathbf{X}_i}(\mathbf{x}_i) &= \frac{\Gamma(\alpha_i + n_i)}{\Gamma(\alpha_i)\, n_i!}\left(\frac{1}{1+\beta_i}\right)^{\alpha_i}\left(\frac{\beta_i}{1+\beta_i}\right)^{n_i}\frac{n_i!}{\prod_{j=1}^{J_i} x_{ij}!}\prod_{j=1}^{J_i} p_{ij}^{\ x_{ij}} \\ &= f_{\mathbf{X}_i|n_i}(\mathbf{x}_i)\ g_{N_i}(n_i) \end{aligned} \tag{5}$$

以上のように，OD ペア i の OD 交通量が負の二項分布に従って確率変動する場合，負の多項分布となる．

（3）多変量正規分布への近似

多変量中心極限定理として，多数の多変量確率変数の和の極限は多変量正規分布に従うことが知られており，経路交通量を多変量正規分布によって近似することが可能である．ただし，本研究では，経路交通量を多変量正規分布で近似し，多変量正規分布とした経路交通量を取り扱うものではない．本節の目的は，前項までの議論では，経路交通量は，一見，負の多項分布という特殊な分布を使用しているように見えたが，一般に交通量は

小さい値ではないため，それは多変量正規分布ともみなせること，つまり，多変量正規分布とみなせるほどの（比較的）自然な確率分布であることを示すためである．

以上のことを確認する前に，まず，負の多項分布の確率母関数を考えよう．確率母関数は離散確率分布に定義される母関数である．OD ペア i の経路交通量の確率変数ベクトル $\mathbf{X}_i$ の確率母関数 $\phi_{\mathbf{X}_i}(\mathbf{u}_i)$ は $\mathrm{E}[\prod_j u_{ij}^{x_{ij}}]$ すなわち $\sum_{x_{i1}}\cdots\sum_{x_{iJ_i}} f_{\mathbf{X}_i}(\mathbf{x}_i)\prod_j u_{ij}^{x_{ij}}$ と定義され，負の多項分布の確率母関数 $\phi_{\mathbf{X}_i}(\mathbf{u}_i)$ は以下の通りである [18]．

$$\phi_{\mathbf{X}_i}(\mathbf{u}_i) = \left[1+\beta_i\left(1-\sum_{j=1}^{J_i} p_{ij}\,u_{ij}\right)\right]^{-\alpha_i} \tag{6}$$

独立な確率変数の和の（確率変数の）確率母関数はそれぞれの確率変数の確率母関数の積である．ここで，$\alpha_i = 1$ の負の多項分布を考えよう．なお，この $\alpha_i = 1$ の負の多項分布の確率母関数 $\phi^1_{\mathbf{X}_i}(\mathbf{u}_i)$ は式 (6) に α_i = 1 を代入することにより得られる．よって，式 (6) の負の多項分布の確率母関数は $\phi_{\mathbf{X}_i}(\mathbf{u}_i) = \left\{\phi^1_{\mathbf{X}_i}(\mathbf{u}_i)\right\}^{\alpha_i}$ となる．独立な確率変数の和の積率母関数は各変数の積率母関数の積であるため，$\phi_{\mathbf{X}_i}(\mathbf{u}_i)$ は独立な確率変数の和の確率母関数とみなせる．OD 交通量は一般に十分に大きな数であるため，α_i も十分に大きな数となる．したがって，上で述べた中心極限定理により，（負の多項分布に従う）経路交通量は極限近似としての多変量正規分布とみなすことができる．このように多変量正規分布に近似した場合の OD ペア i の経路交通量の確率変数ベクトルを $\tilde{\mathbf{X}}_i$ とする．その時の OD ペア i の経路交通量の同時確率密度関数は以下の通りとなる．

$$f_{\tilde{\mathbf{X}}_i}(\mathbf{x}_i) = \frac{1}{\sqrt{(2\pi)^{J_i}|\boldsymbol{\Sigma}_i|}}\exp\left\{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_i)^{\mathrm{T}}\boldsymbol{\Sigma}_i^{-1}(\mathbf{x}-\boldsymbol{\mu}_i)\right\} \tag{7}$$

ここで，$f_{\tilde{\mathbf{X}}_i}(\mathbf{x}_i)$ は OD ペア i の経路交通量が多変量正規分布に従う場合の同時確率密度関数，$\boldsymbol{\mu}_i$ は OD ペア i の経路交通量の平均のベクトル，即ち，$(\mu_{i1},\ldots,\mu_{iJ_i})^{\mathrm{T}}$，$\boldsymbol{\Sigma}_i$ は OD ペア i の経路交通量の分散・共分散行列であり，上で述べた σ_{ij}^2, $\sigma_{ij,j'}$ を要素に持った次式の行列である．

$$\boldsymbol{\Sigma}_i = \begin{bmatrix} \sigma_{ij}^2 & & \cdots & & \sigma_{i1,J_i} \\ & \ddots & & & \\ \vdots & & \sigma_{ij}^2 \quad \sigma_{ij,j'} & & \vdots \\ & & \sigma_{ij,j'} \quad \sigma_{ij'}^2 & \ddots & \\ \sigma_{i1,J_i} & & \cdots & & \sigma_{iJ_i}^2 \end{bmatrix} \tag{8}$$

OD ペア間では，経路交通量は独立であり，独立な多変量正規確率変数の和は多変量正規分布に従うため，リンク交通量の同時確率密度関数は次式となる．

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^{A}|\boldsymbol{\Sigma}_{\mathrm{A}}|}}\exp\left\{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_{\mathrm{A}})^{\mathrm{T}}\boldsymbol{\Sigma}_{\mathrm{A}}^{-1}(\mathbf{x}-\boldsymbol{\mu}_{\mathrm{A}})\right\} \tag{9}$$

ここで，$\boldsymbol{\mu}_{\mathrm{A}}$ はリンク交通量の平均値ベクトルであり，$\boldsymbol{\Sigma}_{\mathrm{A}}$ は，次式のように σ_a^2 及び $\sigma_{aa'}$ を要素に持つリンク交通量の分散・共分散行列である．

$$\boldsymbol{\Sigma}_{\mathrm{A}} = \begin{bmatrix} \sigma_1^2 & & \cdots & & \sigma_{1A} \\ & \ddots & & & \\ \vdots & & \sigma_a^2 \quad \sigma_{aa'} & & \vdots \\ & & \sigma_{aa'} \quad \sigma_{a'}^2 & \ddots & \\ \sigma_{1A} & & \cdots & & \sigma_A^2 \end{bmatrix} \tag{10}$$

ただし，$\boldsymbol{\mu}_{\mathrm{A}}$ と $\boldsymbol{\Sigma}_{\mathrm{A}}$ は以下のように与えられる．

$$\boldsymbol{\mu}_{\mathrm{A}} = \Delta\boldsymbol{\mu} \tag{11}$$

$$\boldsymbol{\Sigma}_{\mathrm{A}} = \Delta\boldsymbol{\Sigma}\Delta^{\mathrm{T}} \tag{12}$$

ここで，Δ はリンク・経路接続行列で，次章で説明する $\delta_{a,ij}$ から構成され，T は転置を意味する．$\boldsymbol{\Sigma}$ は（全ての）経路交通量の分散共分散行列であり，$\boldsymbol{\Sigma}_i$ と同様に σ_{ij}^2, $\sigma_{ij,j'}$ から構成される．このように経路交通量を多変量正規分布で近似すると，リンク交通量も多変量正規分布として与えられる．なお，本研究では，経路交通量は負の多項分布と仮定しており，この場合，その和とも言えるリンク交通量は必ずしも負の多項分布となるとは限らない．しかし，その場合でも，リンク交通量の平均や分散・共分散を計算することが出来る．繰り返しになるが，近似的にはそれも多変量正規分布とみなせる．

3. 最尤法によるパラメータ推定

交通量や交通密度と速度には関数関係を仮定することができる．その関数はグリーンシールズなど様々なものが提案されてきている．どの関数を用いるのかも重要ではあるが，関数のパラメータの推定も重要である．特定の関数を用いるとしても，車線数や制限速度等により状況は大きく異なるため，適切なパラメータの設定が関数系の選択以上に重要ともいえる．例えば，グリーンシールズの式を用いる場合，自由走行速度と飽和密度を与える必要があり，自由走行速度と飽和密度を決めるために制限速度や車線数等からそれらを与える式にはパラメータが必要となろう．このようなパラメータを $\boldsymbol{\theta} = (\theta_1, \theta_2,\ldots)^{\mathrm{T}}$ とおくこととする．また，リンク a の制限速度や車線数など自由走行速度や飽和密度の要因ベクトルを $\mathbf{z}_a = (z_{a1}, z_{a2},\ldots)^{\mathrm{T}}$ とする．交通密度と速度の関係式を v_a

$= h(k_a, \mathbf{z}_a|\boldsymbol{\theta})$ とする．パラメータ $\boldsymbol{\theta}$ が推定できれば，h の逆関数から，プローブデータなどから得られたリンク旅行速度から交通密度 k_a を推定することができる．また，この関係から交通量も計算することができる．

トラカンなどでリンク交通量の観測が行われ，観測リンク交通量ベクトルを $\tilde{\mathbf{x}}$ とする．観測されたリンクの集合を $\tilde{A}$ とする．また，複数回それらのリンクを観測している場合，第 r 回目の観測値を $\tilde{\mathbf{x}}_r$ とする $(r \in R)$．ただし，簡単のために，異なった回の観測値は独立と仮定する．また，プローブデータから得られるリンクの旅行速度ベクトルを $\tilde{\mathbf{v}} = (\ldots, \tilde{v}_a, \ldots)^T$ とする．ここで，$\tilde{v}_a$ はリンク a の旅行速度である．

観測リンク交通量は式 (2) の分布の周辺確率として以下の確率密度関数を持つ多変量正規分布に従う．

$$f_{\tilde{\mathbf{X}}}(\tilde{\mathbf{x}}) = \frac{1}{\sqrt{(2\pi)^{\tilde{n}}\left|\tilde{\boldsymbol{\Sigma}}\right|}} \exp\left\{-\frac{1}{2}(\tilde{\mathbf{x}}-\tilde{\boldsymbol{\mu}})^{\mathrm{T}}\tilde{\boldsymbol{\Sigma}}^{-1}(\tilde{\mathbf{x}}-\tilde{\boldsymbol{\mu}})\right\} \quad (13)$$

ここで，$\tilde{\mathbf{X}}$ は観測交通量の確率変数ベクトル，$\tilde{\boldsymbol{\mu}}$ は観測交通量の平均値ベクトル，$\tilde{\boldsymbol{\Sigma}}$ は観測交通量の分散共分散行列，$\tilde{n}$ は観測リンクの総数である．$\tilde{\boldsymbol{\mu}}$ と $\tilde{\boldsymbol{\Sigma}}$ は，式 (9) で用いられている $\boldsymbol{\mu}$ と $\boldsymbol{\Sigma}$ について，観測しているリンクに関する要素を抜き出して構成することができる．

リンク交通量の実現値，つまり，リンク交通量の観測値 $\tilde{\mathbf{x}}$ が与えられた場合，プローブからのリンク旅行速度データ $\tilde{\mathbf{v}}$ からパラメータを推定するための以下の対数尤度関数 $L(\boldsymbol{\theta}|\,\tilde{\mathbf{x}}, \tilde{\mathbf{v}})$ を定義することができる．

$$L\left(\boldsymbol{\theta}\middle|\tilde{\mathbf{x}}, \tilde{\mathbf{v}}\right) = \ln f_{\tilde{\mathbf{X}}}(\tilde{\mathbf{x}}) = \sum_{a \in \tilde{A}} \ln f_{X_a}(x_a) \quad (14)$$

ただし，$f_{X_a}(x_a)$ はリンク a の交通量の確率密度関数である．複数回の観測値がある場合，

$$L\left(\boldsymbol{\theta}\middle|\tilde{\mathbf{x}}_r, \tilde{\mathbf{v}}_r; \forall r \in R\right) = \sum_{r \in R} \ln f_{\tilde{\mathbf{X}}_r}(\tilde{\mathbf{x}}_r) \quad (15)$$

ここで，$f_{\tilde{\mathbf{X}}_r}(\tilde{\mathbf{x}}_r)$ は r 回目の観測交通量が生起する確率である．そして，尤度関数を l $(\ln l = L)$ で表すことにする．

最尤推定法によるパラメータ推定を以下のように定式化することができる．

$$\max_{\boldsymbol{\theta}} L\left(\boldsymbol{\theta}\middle|\tilde{\mathbf{x}}, \tilde{\mathbf{v}}\right) \quad (16)$$

ここで，$\boldsymbol{\theta}$ はパラメータベクトルで，θ_k $(k \in K)$ から構成される．

このように推定したパラメータを用いて交通量・交通密度と速度の関数関係から，プローブの旅行速度データから交通量を推定することができる．