時間推移する相関構造を考慮した混合ガウシアングラフィカルモデルによる交通状態補間

鈴木　惇平[1]・原　祐輔[2]・桑原　雅夫[3]・

[1]非会員　東北大学 大学院情報科学研究科（〒 980-8579 仙台市青葉区荒巻字青葉 6-6-06）
E-mail: j_suzuki@plan.civil.tohoku.ac.jp
[2]正会員　東北大学 大学院情報科学研究科（〒 980-8579 仙台市青葉区荒巻字青葉 6-6-06）
E-mail: hara@plan.civil.tohoku.ac.jp
[3]正会員　東北大学 大学院情報科学研究科（〒 980-8579 仙台市青葉区荒巻字青葉 6-6-06）
E-mail: kuwahara@plan.civil.tohoku.ac.jp

本研究では，リアルタイムに観測されるプローブデータの補間を目的とし，過去に観測されたプローブデータをもとに，ネットワーク全体の交通状態を表現する生成モデルの学習手法を提案する．モデルの学習では，ネットワーク上の道路リンクの交通状態は，複数のガウシアングラフィカルモデル (GGM) が混合した混合 GGM から生成されていると仮定をおき，プローブデータから混合 GGM を学習するアプローチを行った．また，パラメータ推定の逐次解法である EM アルゴリズムを適用することで，プローブデータに基づいたモデルの学習を可能にした．さらに，L1 正則化付き最適化問題の解法である Graphical Lasso を用いることで，わずかなサンプル数から膨大なパラメータの学習を可能にした．また，実データによる検証を行い，本モデルは観測データごとに異なる混合 GGM を表現することが可能であり，時間推移する相関構造を考慮したモデルであることを示した．

Key Words : *estimation of unobserved road link, Graphical Lasso,Gaussian Graphical Model*

1. はじめに

道路交通制御・管制のために，車両感知器やプローブデータ等と利用して，道路ネットワークのモニタリングが行われてきた．しかし，センサーデータはあくまでネットワーク全体の交通状態のサンプルデータであり，未観測領域が存在する．そのため，現状の観測データから未観測領域を補間する手法が必要である．

観測データから未観測領域の補間を行う手法として，花岡ら[1)]はネットワーク上の交通状態の生成モデルを推定する手法が提案した．この手法は，蓄積されたプローブデータからモデルの学習を行い，オンラインで得られるプローブデータを学習されたモデルを用いて，リアルタイムに補間する手法である．

しかし，花岡らのモデルは，ネットワーク上の交通状態はただ 1 つ多次元正規分布を確率分布とするガウシアングラフィカルモデル (GGM) によって生成されるという仮定を置いている．そのため，花岡らにより提案された生成モデルでは，時間帯に応じてリンク間の相関関係が変化するネットワーク上の交通状態を正確に表現できるとは言い難い．

以上を踏まえて，本研究では過去に観測されたプローブデータと，現状で観測されたプローブデータにより，ネットワーク全体の交通状態を把握することを目的とする．具体的には，花岡らによって提案された手法を拡張し，時間推移するネットワークの交通状態の相関関係を適切に表現したモデルの学習手法を提案する．

2. 混合 GGM のモデル化と対数尤度関数

(1) 本研究の仮定

本研究では，ネットワーク上の交通状態は複数の GGM が混合した混合 GGM によって生成されていると仮定する．また，混合 GGM の混合要素として，GGM を採用していることにより，パラメータ推定を行う際の計算を容易かつ厳密に行うことができる．そして，未観測を含む交通状態データが観測されたときに，事後確率の周辺分布を用いることで，未観測データの補間が可能である．

(2) 混合 GGM のモデル化

次に，混合 GGM のモデル化を行う．ネットワーク上の交通状態 $\boldsymbol{x} = (x_1, x_2, ..., x_{|V|})$ は N 個の GGM が混合した確率分布から生成されると仮定し，この確率分布に従う生成モデルを混合 GGM と定義する．ここで，i 番目の GGM の平均ベクトルは $\boldsymbol{\mu}_i = (\mu_{i1}, \mu_{i2}, ..., \mu_{i|V|})$ で，分散共分散行列は $\boldsymbol{\Sigma}_i$ であり，N 個の GGM は混合

比 $\boldsymbol{\pi} = (\pi_1, \pi_2, ..., \pi_N)$ の割合で混合しているものとする．また，観測データ $\boldsymbol{x}$ が生成された基となる GGM が i 番目の GGM であることを表すとき $\omega_i = 1$，それ以外であれば，$\omega_i = 0$ となるような離散変数ベクトル $\boldsymbol{\omega} = (\omega_1, \omega_2, ..., \omega_N)$ を定義する．

いま，$\omega_i = 1$ である時に，観測データ $\boldsymbol{x}$ が得られる確率を $p(\boldsymbol{x}|\omega_i) = p(\boldsymbol{x}|\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)$ とすると，N 個の混合 GGM による観測データ $\boldsymbol{x}$ の生成モデルは，以下となる．

$$p(\boldsymbol{x}) = \sum_{i=1}^{N} p(\boldsymbol{x}|\omega_i)p(\omega_i) = \sum_{i=1}^{N} \pi_i \cdot p(\boldsymbol{x}|\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i) \quad (1)$$

ここで，$\sum_{i=1}^{N} \pi_i = 1$ であり，$0 \leq \pi_i \leq 1, \forall i$ である．また，ω_i の確率分布は $p(\omega_i = 1) = \pi_i, \forall i$ で表わされる．

次に，ある観測データ $\boldsymbol{x}$ が得られたときに，観測データ $\boldsymbol{x}$ が i 番目の GGM から生成される負担率 $\gamma_i(\boldsymbol{x})$ を定義する．これは，観測データ $\boldsymbol{x}$ が得られた後の事後確率 $p(\omega_i|\boldsymbol{x})$ なので，$p(\omega_i) = p(\omega_i = 1)$ と定義すると

$$\gamma_i(\boldsymbol{x}) = \frac{p(\boldsymbol{x}|\omega_i)p(\omega_i)}{p(\boldsymbol{x})} = \frac{p(\boldsymbol{x}|\omega_i)p(\omega_i)}{\sum_{j=1}^{N} p(\omega_j)p(\boldsymbol{x}|\omega_j)} = \frac{\pi_i \cdot p(\boldsymbol{x}|\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)}{\sum_{j=1}^{N} \pi_j \cdot p(\boldsymbol{x}|\boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j)} \quad (2)$$

と求めることができる．また，π_i は $\omega_i = 1$ となる事象の事前確率を示し，$\gamma_i(\boldsymbol{x})$ は $\boldsymbol{x}$ が観測されたときに $\omega_i = 1$ となる事象の事後確率ととらえることができる．そして，この負担率は混合 GGM のパラメータを推定する上で重要な役割を担う．

(3) 混合 GGM の対数尤度関数の導出

いま，観測データ $\boldsymbol{x}^d$ とそれに対応する潜在変数ベクトル $\boldsymbol{\omega}^d$ が与えられたとすると，式 (1) の混合 GGM の尤度関数 $f(\boldsymbol{\pi}, \boldsymbol{\mu}, \boldsymbol{\Sigma}|\boldsymbol{x}^d, \boldsymbol{\omega}^d)$ は，

$$f(\boldsymbol{\pi}, \boldsymbol{\mu}, \boldsymbol{\Sigma}|\boldsymbol{x}^d, \boldsymbol{\omega}^d) = \sum_{i=1}^{N} \pi_i \cdot p(\boldsymbol{x}^d|\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i) \cdot \delta(\omega_i^d) \quad (3)$$

と表される．ここで，$\delta(\omega_i^d)$ は $\omega_i^d = 1$ のとき $\delta(\omega_i^d) = 1$ であり，$\omega_i^d = 0$ のとき $\delta(\omega_i^d) = 0$ である．ここで，観測データ $\boldsymbol{x}^d$ に対応する潜在変数ベクトル $\boldsymbol{\omega}^d$ はベクトル要素のうちどれか 1 つの要素のみ 1 となるため，式 (4) は以下のように式変形できる．

$$f(\boldsymbol{\pi}, \boldsymbol{\mu}, \boldsymbol{\Sigma}|\boldsymbol{x}^d, \boldsymbol{\omega}^d) = \prod_{i=1}^{N} \left[\pi_i \cdot p(\boldsymbol{x}^d|\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)\right]^{\omega_i^d} \quad (4)$$

したがって，各観測において，観測結果 $\boldsymbol{x}^d$ が互いに独立であると仮定すると，観測結果 $\boldsymbol{X} = (\boldsymbol{x}^1, ..., \boldsymbol{x}^d, ..., \boldsymbol{x}^D)$ と各観測結果に対応する潜在変数 $\boldsymbol{\Omega} = (\boldsymbol{\omega}^1, ..., \boldsymbol{\omega}^d, ..., \boldsymbol{\omega}^D)$ が与えられた時の混合 GGM の尤度関数 $f(\boldsymbol{\pi}, \boldsymbol{\mu}, \boldsymbol{\Sigma}|\boldsymbol{X}, \boldsymbol{\Omega})$ は，以下で表現できる．

$$f(\boldsymbol{\pi}, \boldsymbol{\mu}, \boldsymbol{\Sigma}|\boldsymbol{X}, \boldsymbol{\Omega}) = \prod_{d=1}^{D} \prod_{i=1}^{N} \left[\pi_i \cdot p(\boldsymbol{x}^d|\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)\right]^{\omega_i^d} \quad (5)$$

したがって，対数尤度関数は，以下となる．

$$\ln f(\boldsymbol{\pi}, \boldsymbol{\mu}, \Sigma|\boldsymbol{X}, \boldsymbol{\Omega}) = \sum_{d=1}^{D} \sum_{i=1}^{N} \omega_i^d \cdot \{\log \pi_i + \log p(\boldsymbol{x}^d|\boldsymbol{\mu}_i, \Sigma_i)\} \quad (6)$$

3. プローブデータに基づいた混合 GGM の学習手法

(1) プローブデータに基づくの対数尤度関数

いま，実際に得られる観測データは時間帯ごとに観測されるリンクが異なると仮定する．そして，得られる観測データを $\boldsymbol{x}^d = (\boldsymbol{x}_u^d, \boldsymbol{x}_o^d)$ とし，$\boldsymbol{x}_u$ と $\boldsymbol{x}_o$ をそれぞれ未観測の道路リンクの交通状態と観測された道路リンクの交通状態と定義し，これらの道路について，$p(\boldsymbol{x}|\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)$ の平均ベクトル $\boldsymbol{\mu}_i$，共分散行列 $\boldsymbol{\Sigma}_i$ をそれぞれ

$$\boldsymbol{\mu}_i = (\boldsymbol{\mu}_{iu}, \boldsymbol{\mu}_{io}), \qquad \boldsymbol{\Sigma}_i = \begin{pmatrix} \boldsymbol{\Sigma}_{iuu} & \boldsymbol{\Sigma}_{iuo} \\ \boldsymbol{\Sigma}_{iou} & \boldsymbol{\Sigma}_{ioo} \end{pmatrix} \quad (7)$$

と分割する．今部分的な観測データ $\boldsymbol{x}^d$ に基づく観測結果 $\boldsymbol{X} = (\boldsymbol{x}_u, \boldsymbol{x}_o)$ とそれに対応する $\boldsymbol{\Omega}$ が与えられた時の混合 GGM の対数尤度関数は，式 (8) より以下となる．

$$\ln f(\boldsymbol{\pi}, \boldsymbol{\mu}, \Sigma, |\boldsymbol{X}, \boldsymbol{\Omega}) = \sum_{d=1}^{D} \sum_{i=1}^{N} \omega_{io}^d \cdot \{\log \pi_i + \log p(\boldsymbol{x}_u^d, \boldsymbol{x}_o^d|\boldsymbol{\mu}_i, \Sigma_i)\} \quad (8)$$

ここで，ω_{io}^d は観測リンクの交通状態 $\boldsymbol{x}_o^d$ が生成された基となる GGM が i 番目の GGM であることを表すとき $\omega_{io}^d = 1$，それ以外であれば，$\omega_{io}^d = 0$ となるような離散変数ベクトル $\boldsymbol{\omega}_o^d = (\omega_{1o}^d, \omega_{2o}^d, ..., \omega_{No}^d)$ と定義する．

(2) パラメータ推定の枠組み

いま，式 (8) の対数尤度関数を最大化するようなパラメータを求めたい．しかしながら，この対数尤度関数はデータからは観測できない潜在変数ベクトル $\boldsymbol{\omega}^d$ と未観測の交通状態 $\boldsymbol{x}_u^d$ を含み，直接計算することができない．そこで，この問題に対するために，EM アルゴリズム[2] による手法を援用する．EM アルゴリズムでは，はじめに求めたいパラメータに適当な初期値を与える．そして，次の 2 段階の更新手続きを繰り返し行う．まず，現在のパラメータを用いて潜在変数の期待値を計算する．これを E ステップという．次に，これらの事後確率に基づき，平均，分散，混合パラメータを再計算する．これを M ステップという．すなわち，EM アルゴリズムでは，E ステップにて潜在変数の期待値を計算し，M ステップにて E ステップで求めた潜在変数に基づきパラメータを推定する．また，M ステップでは得られる分散

共分散行列に対して，正則が保障され，スパースな分散共分散行列を求めることができる Graphical Lasso(GL)[3] のアルゴリズムを用いて推定する．以下では，あらかじめ初期値が与えられたもとで，E ステップと M ステップをそれぞれ説明する．

(3) EM を用いたパラメータ推定

a) E ステップ

E ステップでは，既知のパラメータ $\theta^{(t)} = (\pi^{(t)}, \boldsymbol{\mu}^{(t)}, \boldsymbol{\Sigma}^{(t)})$ を用いて潜在変数ベクトル $\boldsymbol{\omega}^d$ と未観測の交通状態 $\boldsymbol{x}_u^d$ の期待値を求める．まず，潜在変数ベクトル $\boldsymbol{\omega}^d$ の期待値は，ある観測データ $\boldsymbol{x}$ が得られたときに，観測データ $\boldsymbol{x}^d$ のうち観測された道路リンクの交通状態 $\boldsymbol{x}_o^d$ が i 番目の GGM から生成される確率を表しており，未観測道路リンク $\boldsymbol{x}_u^d$ に対して周辺化した負担率 $\boldsymbol{\gamma}(\boldsymbol{x}_o^d)$ と対応している．そこで，$\boldsymbol{\omega}^d$ の期待値は，式 (2) より，以下で表される．

$$E[\omega_{io}^d] = \frac{\pi_i \cdot p(\boldsymbol{x}_o^d | \boldsymbol{\mu}_{io}^{(t)}, \boldsymbol{\Sigma}_{ioo}^{(t)})}{\sum_{j=1}^{N} \pi_j \cdot p(\boldsymbol{x}_o^d | \boldsymbol{\mu}_{jo}^{(t)}, \boldsymbol{\Sigma}_{joo}^{(t)})}$$

次に未観測の交通状態 $\boldsymbol{x}_u^d$ の期待値は，未観測道路リンクに対して周辺化した事後確率の最大化により求めることができるため，以下の最大化問題を解くことによって求められる．

$$\begin{aligned} E[\boldsymbol{x}_u^d] &= \arg\max_{\boldsymbol{x}_u^d} p(\boldsymbol{x}_u^d | \boldsymbol{x}_o^d, \boldsymbol{\pi}^{(t)}, \boldsymbol{\mu}^{(t)}, \boldsymbol{\Sigma}^{(t)}) \\ &= \arg\max_{\boldsymbol{x}_u^d} \sum_{i=1}^{N} E[\omega_{io}^d] p(\boldsymbol{x}_u^d | \boldsymbol{x}_o^d, \boldsymbol{\mu}_i^{(t)}, \boldsymbol{\Sigma}_i^{(t)}) \end{aligned}$$

b) M ステップ

M ステップでは，E ステップで求めた未知変数の期待値と $\boldsymbol{\theta}^{(t)}$ を用いて，パラメータの更新を行う．まず，式 (8) の対数尤度関数の期待値を Q 関数とすると，Q 関数は，以下で表される．

$$\begin{aligned} Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(t)}) &= E_{\boldsymbol{\omega}\boldsymbol{x}_u}[\log f(\boldsymbol{\theta}^{(t)} | \boldsymbol{x}_u, \boldsymbol{x}_o, \boldsymbol{\omega})] \\ &= \sum_{d=1}^{D} \sum_{i=1}^{N} E[\omega_{io}^d] \left\{ \log \pi_i^{(t)} + \log f(\boldsymbol{\mu}_i^{(t)}, \boldsymbol{\Sigma}_i^{(t)} | \boldsymbol{x}_o^d, E[\boldsymbol{x}_u^d]) \right\} \end{aligned} \tag{9}$$

いま，i 番目の GGM は確率分布として多次元正規分布をもつために，式 (9) は以下のように変形できる．

$$\begin{aligned} Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(t)}) = \sum_{d=1}^{D} \sum_{i=1}^{N} E[\omega_i^d] \cdot \Big\{ &\log \pi_i^{(t)} + \frac{D}{2} \log \det \boldsymbol{\Theta} \\ &- \frac{1}{2} \sum_{d=1}^{D} (\boldsymbol{x}^{d\prime} - \boldsymbol{\mu}_i^{(t+1)}) \boldsymbol{\Theta} (\boldsymbol{x}^{d\prime} - \boldsymbol{\mu}_i^{(t+1)})^{\mathrm{T}} \Big\} \end{aligned} \tag{10}$$

ただし，i 番目の GGM の分散共分散行列の逆行列を $\boldsymbol{\Theta} = \boldsymbol{\Sigma}^{-1}$ と定義し，$\boldsymbol{x}^{d\prime}$ は，以下のように定義する．

$$\boldsymbol{x}^{d\prime} = \begin{cases} \boldsymbol{x}_o^d, & if \quad i \in O^d \\ E[\boldsymbol{x}_u^d], & if \quad i \in U^d \end{cases}$$

ここで，ネットワーク全体の道路リンクのうち d 番目の観測において，観測された道路リンクの集合を O^d，未観測道路リンクの集合を U^d と表す．

いま，混合 GGM の各要素 i の平均ベクトル $\boldsymbol{\mu}_i$，分散共分散行列 $\boldsymbol{\Sigma}_i$，混合比 π_i は式 (10) の Q 関数を各パラメータに関して最大化を行うことで，求めることが可能である．そこで，まずは混合 GGM の各要素 i の平均ベクトル $\boldsymbol{\mu}_i$ の導出を行う．式 (10) の Q 関数を GGM 要素 i の平均 $\boldsymbol{\mu}_i$ に関して偏微分し 0 とおくと，以下を得る．

$$\begin{aligned} &\frac{\partial Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(t)})}{\partial \boldsymbol{\mu}_i} = \boldsymbol{\Sigma}_i^{-1} \sum_{d=1}^{D} E[\omega_i^d](\boldsymbol{x}^{d\prime} - \boldsymbol{\mu}_i^{(t+1)}) = 0 \\ &\therefore \boldsymbol{\mu}_i^{(t+1)} = \frac{\sum_{d=1}^{D} E[\omega_i^d](\boldsymbol{x}^{d\prime} - \boldsymbol{\mu}_i^{(t+1)})}{\sum_{d=1}^{D} E[\omega_i^d]} \end{aligned} \tag{11}$$

次に，分散共分散行列 $\boldsymbol{\Sigma}_i$ の更新を行う．式 (11) と同様にして，式 (9) の Q 関数を各要素 i の共分散の逆行列 $\boldsymbol{\Theta}_i$ に関して偏微分し 0 とおくと，各要素 i の共分散 $\boldsymbol{\Sigma}_i$ は以下のように与えられる．

$$\begin{aligned} \frac{\partial Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(t)})}{\partial \boldsymbol{\Theta}_i} = \sum_{d=1}^{D} E[\omega_i^d] \Bigg\{ &\frac{1}{2} \frac{\partial \det \boldsymbol{\Theta}_i}{\partial \boldsymbol{\Theta}_i} \\ &- \frac{1}{2} \frac{\partial (\boldsymbol{x}^{d\prime} - \boldsymbol{\mu}_i^{(t+1)}) \boldsymbol{\Theta}_i (\boldsymbol{x}^{d\prime} - \boldsymbol{\mu}_i^{(t+1)})}{\partial \boldsymbol{\Theta}_i} \Bigg\} \end{aligned}$$

$$\boldsymbol{\Sigma}_i^{(t+1)} = \frac{\sum_{d=1}^{D} E[\omega_i^d](\boldsymbol{x}^{d\prime} - \boldsymbol{\mu}_i^{(t+1)})(\boldsymbol{x}^{d\prime} - \boldsymbol{\mu}_i^{(t+1)})^{\mathrm{T}}}{\sum_{d=1}^{D} E[\omega_i^d]} \tag{12}$$

最後に，混合比 π_i の更新を行う．ここでも式 (11) と同様にして，混合比 π_i は式 (10) の Q 関数を混合比 π_i に関して最大化を行うことで，求めることが可能である．ただし，$\sum_{i=1}^{N} \pi_i = 1$ であるという制約条件を考慮しなくてはならないので，ラグランジュ未定乗数法を用いる．ここで，ラグランジュ関数 $L(\boldsymbol{\pi}, \boldsymbol{\mu}, \boldsymbol{\Sigma})$ は，

$$L(\boldsymbol{\pi}, \boldsymbol{\mu}, \boldsymbol{\Sigma}) = Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(t)}) + \lambda \Big(\sum_{i=1}^{N} \pi_i - 1 \Big) \tag{13}$$

で表される．ここで λ はラグランジュ未定乗数を表す．したがって，ラグランジュ関数 $L(\boldsymbol{\pi}, \boldsymbol{\mu}, \boldsymbol{\Sigma})$ を $\pi_i (i =$

$1,...,N$) で微分して 0 とおくと，

$$\frac{\partial Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(t)})}{\partial \pi_i} = \frac{1}{\pi_i^{(t+1)}} \sum_{d=1}^{D} E[\omega_{io}^d] - \lambda = 0$$

$$\pi_i^{(t+1)} \lambda = \sum_{d=1}^{D} E[\omega_{io}^d] \tag{14}$$

を得る．ここで，d 番目の観測における潜在変数 ω_{io}^d の総和は 1 となるため，上式の両辺に $\sum_{i=1}^{N}$ を施すと，$\lambda = D$ を得る．したがって，以下を得る．

$$\pi_i^{(t+1)} = \frac{1}{D} \sum_{d=1}^{D} E[\omega_{io}^d] = \frac{D_i}{D} \tag{15}$$

ここで，D_i は観測結果 $\boldsymbol{X}$ が得られた時に観測データ $\boldsymbol{x}^d$ が生成された基となる GGM が i 番目の GGM である回数の期待値を表す．

(4) GL による分散共分散行列の推定

M ステップで示したように理論的には，i 番目の GGM における 3 つのパラメータ $\boldsymbol{\mu}_i$，$\boldsymbol{\Sigma}_i$，π_i の推定値を求めることが可能である．しかし，このパラメータ推定には大きく 2 つの課題がある．まず，1 つ目の問題は過学習の可能性があることが挙げられる．これはパラメータ数に対して，推定のためのデータが十分に得られていない場合，推定されたパラメータは推定の際に用いたデータに強く依存してしまい，モデルの汎用性が低下してしまう問題である．2 つ目の問題は，標本分散共分散行列の正則性である．式 (12) の $\boldsymbol{\Sigma}_i$ は，標本分散共分散行列であり，正則になることは保障されておらず，逆行列を求めることはできない．そこで，これらの問題を解決するために，本研究では，多次元正規分布に正則化項を加えた分布を予測モデルとし，Freadman *et al* により提案された GL アルゴリズムによりパラメータを推定する方法を用いる．この GL のアルゴリズムを用いることで，標本分散共分散行列から正則化が保障された分散共分散行列を求めることができる．

いま，式 (12) より得られる標本分散共分散行列を $\boldsymbol{S}_i$ とすると，i 番目の GGM の対数尤度関数は以下のように整理できる．

$$\ln f(\boldsymbol{\mu}_i, \boldsymbol{\Theta}_i) = \frac{1}{2} D_i \cdot \log\det \boldsymbol{\Theta}_i - tr(\boldsymbol{S}_i \boldsymbol{\Theta}_i) + const \tag{16}$$

GL アプローチでは，対数尤度関数 (16) に正則化項を加えた以下の最適化問題を解く．

$$\boldsymbol{\Theta}_i = \arg\max_{\boldsymbol{\Theta}_i} \frac{D_i}{2} \log\det \boldsymbol{\Theta}_i - tr(\boldsymbol{S}_i \boldsymbol{\Theta}_i) - \rho \|\boldsymbol{\Theta}_i\|_1 \tag{17}$$

ここで，$\|\boldsymbol{\Theta}_i\|_1$ は，L1 ノルムであり，$\boldsymbol{\Theta}_i$ の絶対値の和 $\sum_{i,j=1}^{|V|\times|V|} |\boldsymbol{\Theta}_i|$ によって定義される．

式 (17) の最適化問題は，第三項に L1 ノルムが含まれており，$\Theta_{ij} = 0$ で微分不可能であるため勾配法を用いることができない．しかしながら，劣勾配法による解法である GL によるアルゴリズムを用いることで，この問題を解くことが可能である．

そこで次に，GL による完全観測データに基づく共分散構造推定について解説する．GL では，式 (17) の問題を，特定の変数に関して式変形を行う．そして，各変数ごとの L1 制約付きの回帰問題に帰着させることで，最適解を求める．式 (17) を $\boldsymbol{\Theta}_i$ について微分すると以下を得る．

$$\frac{\partial}{\partial \boldsymbol{\Theta}_i} \ln f(\boldsymbol{\Theta}_i) = \boldsymbol{\Theta}_i^{-1} - \boldsymbol{S}_i - \rho \cdot Sign(\boldsymbol{\Theta}_i) \tag{18}$$

ここでは，劣勾配記法を用いており，$\theta_{ij} \neq 0$ のとき $Sign(\theta_{ij}) = sign(\theta_{ij})$ であり，$\theta_{ij} = 0$ のとき $Sign(\theta_{ij}) \in [-1, 1]$ である．対数尤度関数 (17) を最大化するためには，式 (18) を用いることにより，以下の方程式の解を求めなければならない．

$$\boldsymbol{\Theta}_i^{-1} - \boldsymbol{S}_i - \rho \cdot Sign(\boldsymbol{\Theta}_i) = 0 \tag{19}$$

いま，分散共分散行列 $\boldsymbol{\Sigma}_i = \boldsymbol{\Theta}_i^{-1}$ の最適解を $\boldsymbol{W}_i$ とし，$\boldsymbol{\Theta}_i$，$\boldsymbol{W}_i$ および $\boldsymbol{S}_i$ を以下のように分割する．

$$\boldsymbol{\Theta}_i = \begin{pmatrix} \boldsymbol{\Theta}_{i11} & \boldsymbol{\theta}_{i12} \\ \boldsymbol{\theta}_{i12}^T & \theta_{i22} \end{pmatrix}, \quad \boldsymbol{W}_i = \begin{pmatrix} \boldsymbol{W}_{i11} & \boldsymbol{w}_{i12} \\ \boldsymbol{w}_{i12}^T & w_{i22} \end{pmatrix},$$
$$\boldsymbol{S}_i = \begin{pmatrix} \boldsymbol{S}_{i11} & \boldsymbol{s}_{i12} \\ \boldsymbol{s}_{i12}^T & s_{i22} \end{pmatrix} \tag{20}$$

ここで，$\boldsymbol{\theta}_{i12}$，$\boldsymbol{w}_{i12}$，$\boldsymbol{s}_{i12}$ はそれぞれ，各行列の一列を抜き出し，非対角項のみで構成されたベクトルであり，θ_{i22}，w_{i22}，s_{i22} はその対角項であり，スカラーで表わされる．

ここで，式 (19) の解法を求める．GL を用いると，非対角の要素に関しては，他の変数を全て固定したという条件の下で，以下の問題として表わされる．

$$\frac{\partial}{\partial \boldsymbol{\beta}_i} \left\{ \frac{1}{2} \left(\boldsymbol{W}_{i11}^{1/2} \boldsymbol{\beta}_i - \boldsymbol{b} \right)^2 + \rho \|\boldsymbol{\beta}_i\|_1 \right\} = 0 \tag{21}$$

これは，$\boldsymbol{\beta}_i$ についての L1 制約付き回帰問題である．ここで，$\boldsymbol{\beta}_i = \boldsymbol{W}_{i11}^{-1} \boldsymbol{w}_{i12}$，$\boldsymbol{\beta}_i \in \mathbb{R}^{|V|-1}$，$\boldsymbol{b}_i = \boldsymbol{W}_{i11}^{-1/2} \boldsymbol{s}_{i12}$ である．

いま分散共分散行列 $\boldsymbol{\Sigma}_i$ の最適解 $\boldsymbol{W}_i$ を求めたい．式 (21) より，最適解 $\boldsymbol{W}_i$ を分割した右上ブロック $\boldsymbol{w}_{i12}$ は，$\boldsymbol{\beta}_i$ についての L1 制約付き回帰問題であるため，次の手順で求めるられる．まず，$\boldsymbol{\beta}_i$ の最適解を求める．次に，$\boldsymbol{\beta}_i = \boldsymbol{W}_{i11}^{-1} \boldsymbol{w}_{i12}$ により $\boldsymbol{w}_{i12}$ を求める．さらに，$w_{ii} = s_{ii} + \rho$ によって対角項を得る．そこで，本研究では，式 (12) で求められる標本分散共分散に対して，GL アルゴリズムを適用することによって，スパースで正則が保障された分散共分散行列 $\boldsymbol{\Sigma}_i$ を推定する．

4. 実データによる精度検証

(1) 精度検証の設定条件

a) データ概要

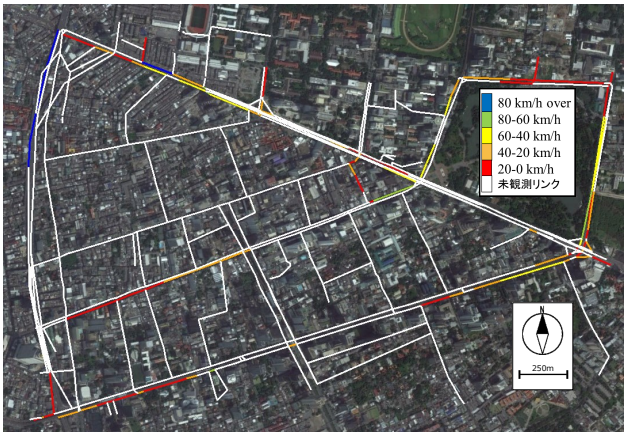
対象ネットワークをバンコク中心部とし，Open Street Map から 508 本のリンクにより構成されるネットワークを作成した．また本検証で用いるデータは，バンコク中心部において 2013/9/1-9/7，10/14-11/5 の延べ 30 日間で取得されたプローブカーデータである．このデータを作成したネットワークに対してマップマッチングを行い，プローブカーが通行した道路リンクにおいて，5 分間平均リンク速度を算出した．対象時間は午前 0 時から午後 24 時までの 24 時間 (288 単位) とし，1 データを 508 リンクの 5 分間平均速度とすると全データセットは全部で 8640 個 (288 時間帯× 30 日間) 存在する．

b) 精度検証方法

全観測期間 (8640 データ) のうち，データを学習用の訓練データ (7776) と精度検証を行うための検証データ (864) に分割する．ただし，検証用データは本データは未観測リンクの真値は得られていないため，各データごとに観測リンクをランダムに 10 個欠損させ，真値と予測値の比較を行う．

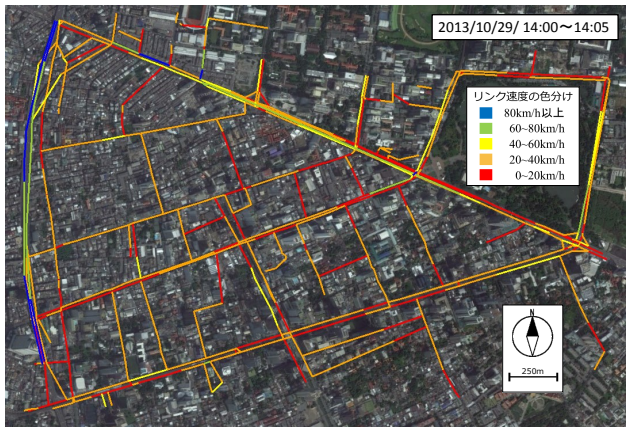
(2) 精度検証結果

以下では，正則化係数 $\rho = 5$，混合数 $N = 4$ と設定し，交通状態補間を行った結果を示す．図 1，2 はそれぞれ 2013 年 10 月 29 日の 14:00〜14:05 における実際の観測状態と提案手法による補間結果を示す．このように，本手法を用いて，空間的な観測率が低い観測データからネットワーク全体の交通状態の補間が可能である．



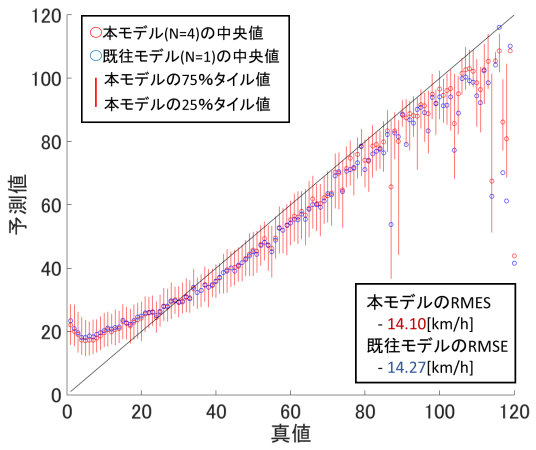
図–1 観測プローブデータのリンク速度

次に，真値と予測値の関係を示す．本モデル (赤) による補間結果における真値と予測値の関係を図 3 に示す．ここで，真値と予測値のばらつきを比較するため，



図–2 本手法が補間したネットワーク全体のリンク速度

真値に対する予測値の中央値，25 パーセントタイル値と 75 パーセントタイル値をプロットした．また，横軸に真値，縦軸に予測値をとっており，補間精度が高ければ，両者の関係は 45° 線に近い値を示す．また，図 3 の青の中央値は混合数 $N = 1$ である既往モデルの中央値を表しており，既往モデルよりも真値に近いことがわかる．また，既往モデルでは RMSE が 14.27$[km/h]$ であったのに対して，混合数 $N = 4$ としたときの本モデルでは RMSE は 14.10$[km/h]$ となり，既往モデルよりも若干精度が向上した．



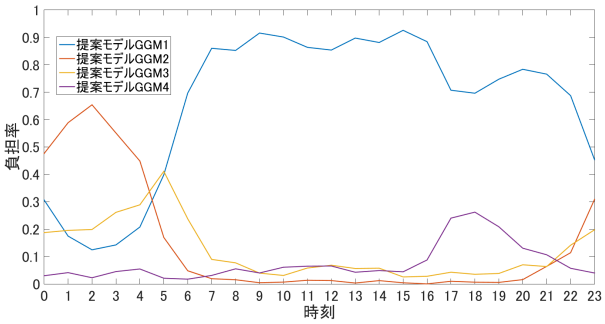
図–3 本モデル (N=4) における真値と予測値の関係

(3) 検証結果の考察

ここでは，本モデルが学習した 4 つの GGM を順に GGM1〜4 とし，各 GGM の考察を行う．図 4 は，学習された負担率のうち，平日の結果のみを抽出し，1 時間単位の平均値をとったもので，平日における負担率の時間推移を表す．また，各 GGM のパラメータ $\boldsymbol{\mu}_i$，$\boldsymbol{\Sigma}_i$ に関しては，紙面の都合上，分析結果のみを示す．

まず，学習された各 GGM の平均値 $\boldsymbol{\mu}_i$ は，GGM2 が

最も高い値を示し，その次に GGM3 が続いた．一方で，GGM4 は最も低い値をとり，GGM3 はその次に低い値をとった．次に，各 GGM の共分散構造の学習結果に関しては，GGM1，GGM3，GGM4 は主要幹線道路に強い正の相関を持ち，特に GGM4 では，非常に強い正の相関を示した．一方で，GGM2 は，他の GGM と同様に主要幹線道路に正の相関を持つが，ネットワーク全体として相関が小さい値を示した．



図–4 学習された負担率の平日平均の時間推移

以上の結果を踏まえて，図 4 より各 GGM は以下のように位置づけられる．まず，日中の長い時間帯において負担率が高い GGM1 は，低い平均値をもち，ネットワーク上の主要道路に強い正の共分散を持つ GGM であり，7 時から 22 時にかけての渋滞しやすい交通状態を捉えている．深夜から明け方にかけて負担率が高い GGM2 は，平均値として高い速度をとり，ネットワーク上の相関関係が比較的小さいため，同時間帯の交通量が少なく，各道路リンクは高い速度をとりやすい現象を捉えている．

次に，明け方 5 時に負担率が高い GGM3 は，平均値としてやや高い速度をとり，ネットワーク上の主要道路に比較的大きい共分散を持つため，GGM1 と混合することで，明け方から通勤時間帯にかけての渋滞直前の交通状態捉えられている．最後に夕方 18 時前後に負担率を持つ GGM4 は平均値として非常に低い速度をとり，ネットワーク上の主要道路に非常に強い正の共分散を持つ GGM であり，18 時前後の激しく渋滞しやすいの交通状態を捉えている．

5. まとめ

今回の検証では，GGM の混合数が 1 つである既往モデルと比べて混合 GGM を用いた手法の方が高精度な結果が得られたため，混合 GGM によって推定された予測モデルの方がより現実の交通現象を表現するモデルであったと考えられる．さらに，学習された混合 GGM の各要素はそれぞれ異なる相関構造を持ち，各 GGM の負担率が観測データごとに変化することでネットワークの相関構造の時間推移を捉えていることを示した．

ただし，今回対象としたネットワークは非常に小さく，ネットワーク上の各道路リンクは空間的に似た性質を持っていることが考えられ，複数の GGM が扱えるという本モデルの特徴を評価するためには，時空間に均一ではないエリアを対象として検証することが求められる．

参考文献

1) 花岡洋平, 原祐輔, 片岡駿, 桑原雅夫: Graphical Lasso を用いた長期観測プローブデータによるリンク交通状態補間, 土木計画学研究・講演集, Vol.51, CD-ROM, (2015).
2) Dempster, A. P., Laird, N. M., and Rubin, D. B., :Maximum Likelihood from Incomplete Data via the EM Algorithm, Journal of the Royal Statistical Society. Series B (Methodological), Vol.39, No.1, pp.1-38, (1977).
3) Freedman, J., Hastie, T. and Tibshirani, R., :Sparse inverse covariance estimation with the graphical lasso, Biostatistics , 9, 3, pp. 432―441, (2008).

(2016.4.22 受付)