

ロジスティックタッカー分解に基づく 時空間上で連続な渋滞のパターン抽出

木曾 俊輔¹・井上 亮²

¹学生会員 東北大学博士前期課程 大学院情報科学研究科人間社会情報科学専攻
(〒980-8579 仙台市青葉区荒巻字青葉 6-6-06)
E-mail: shunsuke.kiso.q8@dc.tohoku.ac.jp

²正会員 東北大学准教授 大学院情報科学研究科人間社会情報科学専攻
(〒980-8579 仙台市青葉区荒巻字青葉 6-6-06)
E-mail: rinoue@tohoku.ac.jp

交通渋滞を緩和するためには、道路ネットワークにおける交通渋滞の時空間的なパターンを把握することが重要である。近年、複雑な関係性を有するデータから特徴抽出を行う手法としてテンソル分解が注目されている。テンソル分解を用いた分析では、高次のテンソルで表されるデータを、低次のテンソルの積和で表現する操作を通してデータの特徴を抽出する。この方法は、日・時刻・地点などに関してパターンを有する交通データの分析にも適用されており、これまで、交通状態予測や異常検知などの交通データ解析に適用されている。しかし、分析対象のテンソルを構成する要素が従う確率分布に対応した分解手法や、渋滞の時空間的な連続性を考慮した分解手法に関する議論は不十分である。そこで本研究では、テンソルデータの分布と渋滞の時空間的連続性を考慮できるグラフラプラシアン正則化を用いたロジスティック非負値タッカー分解を提案し、首都高速道路網の渋滞観測データを用いて提案手法の有効性を確認した。

Key Words: *traffic congestion patterns, logistic tensor decomposition, Graph Laplacian regularization, Spatio-temporal data,*

1. はじめに

交通渋滞は、所要時間の増加による経済損失や、発進・停車に伴う騒音・排ガスといった環境への悪影響を及ぼしており、世界中の都市で解決すべき問題として認識されている。これまで、渋滞緩和を目指して、交通状態の特徴を捉えた将来予測や、典型的なパターン抽出、非典型的な渋滞発生の検出等を目的とした多くの研究が行われ、近年では、収集された大量の交通観測データを用いた実証研究が盛んに行われている。

渋滞特性を捉える上で注目すべきは、交通状態の時間的特徴と空間的特徴である。時間的特徴は、日間変動・日内推移が存在するという性質であり、空間的特徴は、場所によって混雑状況が異なるという性質である。さらに実際には、日・時刻・地点の特徴が組み合わされた時空間的特徴が存在しており、渋滞特性把握においては、この時空間パターンの解明が重要であると言える。

これまで、交通観測データから生成された渋滞の有無や頻度・割合を示すデータをもとに、典型的な混雑パタ

ーンを解析する研究が数多く行われてきた。例えば、頻出パターンマイニングアルゴリズムや、パターンの類似性に基づくクラスタリング手法、テンソル分解手法など様々な手法^{1,3)}で、渋滞の時空間パターン分析が行われている。

道路ネットワーク上の交通状態は、時間的・空間的特徴が相互に影響しあい、複雑な多項関係を有する。このようなデータから特徴を抽出する分析手法の一つは、行列分解を多次元配列(テンソル)に拡張したテンソル分解で、道路交通データ分析にも活用されている。Han and Moutarde (2016)⁴⁾は、グラフラプラシアン正則化を導入した非負値CP分解を、混雑割合を示すテンソルデータに適用し、道路リンクの接続関係を考慮して渋滞が伝播する空間的範囲の抽出に成功した。馬場ら (2022)⁵⁾は、首都高速道路の渋滞時空間のパターンの解明を目的に、非負値タッカー分解と、抽出結果に対するクラスタ分析を組み合わせた分析を提案した。時間的・空間的な渋滞パターンに加えて、日・時刻・地点の相互作用による時空間パターンを抽出できることを確認し、テンソ

ル分解による渋滞分析の有効性を確認した。

しかし、これらの研究は、分析対象のテンソルを構成する要素が従う確率分布を考慮したモデルの検討や、渋滞現象の時間的・空間的な連続性を同時に考慮した分析に関する議論が不十分である。

以上を踏まえ、本研究は、データの従う確率分布、および、渋滞の時空間の連続性を考慮可能なテンソル分解手法として、グラフラプラシアン正則化を導入したロジスティック非負値タッカー分解モデルを提案する。この手法の有効性を確認するため、首都高速道路の観測データから生成した渋滞の有無を表すデータから特徴を抽出・分析し、渋滞の時空間パターン分析における提案手法の有効性を考察する。

2. 提案手法

テンソル分解は、行列分解をテンソルに拡張した手法である。テンソルデータの各要素軸（モード）の次元を圧縮し、多要素によって構成される複雑なデータの特徴を抽出できる。

テンソル分解は主にタッカー分解とCP分解に分類される。

タッカー分解は、テンソルを各モードの特徴を表す因子行列と低次元配列であるコアテンソルに分解する。得られた因子行列の列ベクトル（以下、「因子ベクトル」）は各モードにおけるパターンを表す。以下、3階テンソル $X \in \mathcal{R}^{I_1 \times I_2 \times I_3}$ を例に説明する。タッカー分解モデルは、

$$X' = \Lambda \times U^{(1)} \times U^{(2)} \times U^{(3)} \quad (1)$$

で表され、 $\Lambda \in \mathcal{R}^{R_1 \times R_2 \times R_3}$ はコアテンソル、 $U^{(i)} \in \mathcal{R}^{I_i \times R_i}$ ($i = 1, 2, 3$) は各モードの因子行列、 (R_1, R_2, R_3) は各モードの因子数を表す。

Hong et al. (2019)⁶⁾は、テンソルの要素が従う分布に対応したテンソル分解モデルの提案を行っている。この中で、二項分布に対応したロジスティックCP分解をはじめ、ポアソン分布、ガンマ分布などのいくつかの確率分布に対応したCP分解モデルを提案している。本研究の分析対象は渋滞の有無を示す $[0, 1]$ の二値データで、二項分布に従うことを仮定するのが適切である。そこで、二項分布を仮定したテンソル分解モデルを設定することで、より適切に渋滞の時空間パターンを抽出できることが期待できる。

CP分解は、式(1)のコアテンソルが対角との制約を加えたタッカー分解の特殊形である。全てのモードに同じ次元数を設定し、因子ベクトルは一対一に対応する。一方、タッカー分解は各モードの次元数を自由に設定でき、

また各モードの各因子ベクトルを組み合わせることで入力テンソルを表現できるため、複雑な多項関係を含むデータの分析に適する。また、タッカー分解に非負制約を課した非負値タッカー分解は、モデルが和算のみで表現され、データに内在する複雑な関係性をパターンの重なりとして表現することから、解釈性が高い。

そこで本研究は、分析対象のテンソルが従う分布に対応したモデルの表現力や解釈性の高いモデルとして、ロジスティックCP分解を拡張した、ロジスティック非負値タッカー分解を提案する。本手法は、入力テンソルの平均値（分析対象のネットワークで発生する渋滞の平均発生頻度を表す）をオフセット項として設定し、ロジスティックCP分解を非負値タッカー分解へと拡張するため、オフセット項よりも小さな値を除外した分解を行う。

また、交通渋滞は、ボトルネックを起点に発生した後、上流へ時間を掛けて伝播する。また、最終的に解消するまで一定の時間にわたり継続する。そのため、各地点では連続した時間帯で継続して渋滞が観測され、隣接地点では同時刻帯に渋滞が観測される場合が多い。そのため、渋滞が時空間的に連続的に発生することを明示的に踏まえた特徴抽出が有効だと考えられる。こうしたデータ間の関係性を考慮可能な手法として、グラフラプラシアン正則化を導入した非負値タッカー分解⁷⁾が提案されている。この手法は、時刻の連続性や感知器間の接続関係を隣接行列(グラフ)として定義し、隣接行列とそこから計算される次数行列を基に求められるグラフラプラシアンを用いた正則化を行う。グラフラプラシアン正則化を用いることで、グラフの枝が存在するペア(例：連続する時間帯の観測の隣接する感知器ペア)のパターンは類似するものとして特徴抽出が可能となるため、渋滞の時空間的な連続性を考慮すると、より解釈性の高いパターン抽出が期待できる。

以上より、二項分布に従う渋滞の有無を表すテンソルデータを、渋滞の時空間上の連続性を明示的に考慮して分解する手法として、グラフラプラシアン正則化を導入したロジスティック非負値タッカー分解を提案する。提案手法の最適化問題を式(2)に示す。

$$\min_{\mu, \Lambda, U^{(1)}, U^{(2)}, U^{(3)}} -\langle X, \Theta \rangle + \langle \mathbf{1}_{I_1 I_2 I_3}, \log(\mathbf{1}_{I_1 I_2 I_3} + e^\Theta) \rangle \quad (2)$$

$$+ \frac{\alpha}{2} \text{Tr}(U^{(2)T} L^{(2)} U^{(2)} + U^{(3)T} L^{(3)} U^{(3)}),$$

$$\text{s. t. } \Theta = \mu \mathbf{1}_{I_1 I_2 I_3}$$

$$+ \sum_{r_1}^{R_1} \sum_{r_2}^{R_2} \sum_{r_3}^{R_3} \lambda_{r_1 r_2 r_3} (\mathbf{u}_{r_1}^{(1)} \circ \mathbf{u}_{r_2}^{(2)} \circ \mathbf{u}_{r_3}^{(3)}),$$

$$\Lambda > 0,$$

$$U^{(1)}, U^{(2)}, U^{(3)} > 0.$$

Θ はロジットパラメータテンソル, α はグラフラプラシアン正則化パラメータ, μ は入力テンソルの平均値を設定したオフセット項, $L^{(i)} \in \mathcal{R}^{l_i \times l_i}$ ($i = 2,3$)はモード*i*のラプラシアン行列を表す. また本研究は, 渋滞発生の時刻の連続性・地点の接続性に着目するため, 時刻と地点のモードに正則化を導入する.

Zhang et al. (2021)⁸⁾は, ロジスティックCP分解の最適化手法としてMajorization-Minimization (MM) アルゴリズムを適用した. これにより, ロジスティックCP分解の最適化問題をMMアルゴリズムで補助関数を導入することで, 既存のCP分解手法と同じ分解アルゴリズムによる計算が可能となった. 本研究は, Zhang et al. (2019)⁹⁾を参考に, 式(2)の最適化手法としてMMアルゴリズムを採用する. MMアルゴリズムは, 元の損失関数を補助関数に置き換え, 置き換えた補助関数を反復的に最小化することで, 元の損失関数を間接的に最小化するアルゴリズムである. 本手法では, テンソル分解の各反復で, 式(2)に対して補助関数に置き換え, その後, Qu et al. (2019)⁷⁾で示されたグラフラプラシアン正則化を導入した非負値タッカー分解アルゴリズムを用いて, 補助関数の反復的な最小化を行う.

3. 分析対象・分析テンソルデータの生成

(1) 分析対象

表-1に示す首都高速道路15路線の, 原則として東京都内の区間を対象とする. 図-1は対象範囲の道路網を簡略化した模式図で, 以降の分析結果の表現に用いる. 2014年1月1日から12月31日までの1年間に, 本線上1,441箇所の感知器で観測された1分間隔の時間平均速度データを, 15分間隔で集計したデータを分析する.

(2) 渋滞テンソルデータの生成

渋滞現象分析にあたり, 各感知器の地点における 15分間毎の渋滞の有無を表すデータを生成する. 具体的には, (1)で述べた交通観測データに対して, a) 渋滞状態の判定, b) 軸(モード)の設定の手順で生成する.

a) 渋滞状態の判定

渋滞現象の分析にあたり, 各感知器の地点における時々刻々の渋滞状態の判定基準は, 明確な基準がなく研究によって手法が異なる. 本研究は, 観測された交通速度データの分布に着目した. これまでの研究^{9,10)}で, 渋滞・非渋滞状態が存在する場合, 観測された速度データの分布は多峰性を有し, その分布は混合正規分布として近似できることが確認されている. 以上を踏まえて本研究は, 観測された速度データの分布に着目した地点単位

表-1 対象路線一覧

記号	路線名	記号	路線名
1	1号羽田線	7	7号小松川線
18	1号上野線	9	9号深川線
2	2号目黒線	11	11号台場線
3	3号渋谷線	C1	都心環状線
4	4号新宿線	C2	中央環状線
5	5号池袋線	B	湾岸線
6	6号向島線	S1	川口線
16	6号三郷線		

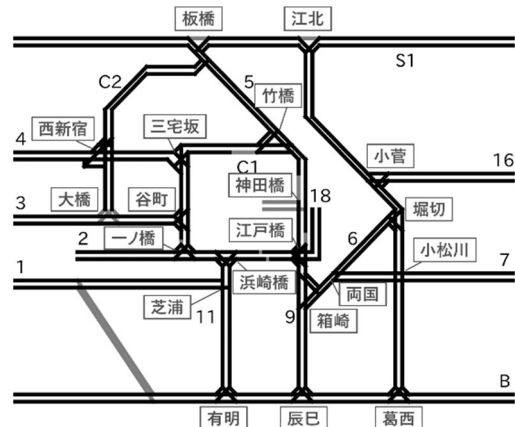


図-1 首都高速道路の模式図

(図上に主な JCT 名を示す)

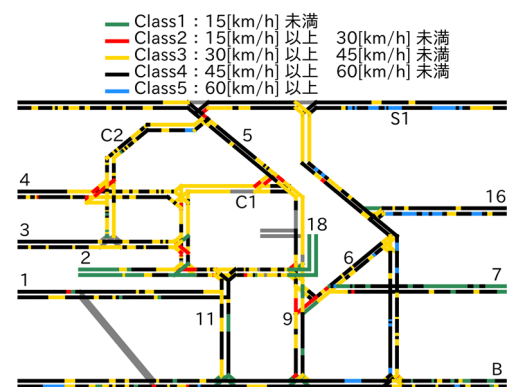


図-2 渋滞判定速度の空間分布

での渋滞判定手法として, 画像の二値分類で用いられる Kittler 法¹¹⁾を適用する. Kittler 法とは, 画像処理の分野で開発された手法で, 画像内の対象領域と背景領域の濃淡値が共に正規分布に従うという仮定の下で, 2つの分布の境界となる閾値の設定に用いられる. Kittler 法による渋滞状態判定は, 白石ら(2013)¹²⁾や坪田ら(2018)¹³⁾で行われている.

本分析では, 分析対象路線内に設置された本線上の車両感知器 1,441 箇所ごとに集計された速度データに対して, Kittler 法を適用し, 感知器毎の渋滞判定速度を推定

する。推定された渋滞判定速度を用いて、日 d 、時刻 t 、地点 i の渋滞状態変数 $\delta_{d,t,i}$ を、式(3)の二値変数（渋滞/非渋滞）で定義した。

$$\delta_{d,t,i} = \begin{cases} 1 & \text{if } (v_{d,t,i} < V^*) \\ 0 & \text{if } (v_{d,t,i} \geq V^*) \end{cases} \quad (3)$$

Kittler 法により推定された渋滞判定速度の空間分布を図-2に示す。

b) 軸（モード）の設定

a) で生成した、渋滞の有無を示すデータに対して、日・時刻・地点モードの3軸を設定し、渋滞の有無を表すテンソルを生成した。テンソルのサイズは、365日×96時刻×1441地点で、総要素数は50,492,640である。

4. 提案手法を用いたパターン分析・考察

本項では、生成した3階テンソルに対し、グラフラプラシアン正則化を導入したロジスティック非負値タックカー分解を適用した分析結果を示す。圧縮次元数は、モード毎に因子ベクトルの寄与度を指標として、寄与度が最大寄与度の2.5%未満となる因子ベクトルが抽出されない限界の次元数(8, 8, 15)に設定した。また、正則化パラメータは、モデルの再構成誤差と、解釈性の観点から、 $\alpha = 100$ の分析結果を示す。なお、モード i ($i = 1, 2, 3$)における因子行列の r_i 番目の列ベクトルを第 r_i 因子と記載する。

(1) 各モードにおける渋滞パターンの分析

渋滞の時空間パターンの分析にあたり、まず、各モードに関して、日間変動・日内推移・渋滞発生地点の周期性や規則性の存在を検証する。

a) 日モード

日モードで得られた因子ベクトルを図-3に示す。各因子ベクトルの要素の最大値と最小値を基準としたカラースケールで表し、縦軸は週、横軸は曜日（月曜始まり）である。

図-3から、それぞれ季節や曜日に特徴をもつ因子ベクトルが抽出された。例えば、第1因子は繁忙期の平日を表しているのに対し、第2因子は、第1因子と色の濃さが反転しており、休日及び閑散期の平日を表す因子ベクトルであるといえる。また、第3因子は日曜日や祝日・大型連休といった、「休日」を表す因子であり、第4因子は「土曜日」を表す因子であることが見て取れる。また、第5因子は、2月の特定の日を表すパターンであり、これは記録的な降雪に見舞われた日である。

以上のことから、日モードでは、交通渋滞の季節性や

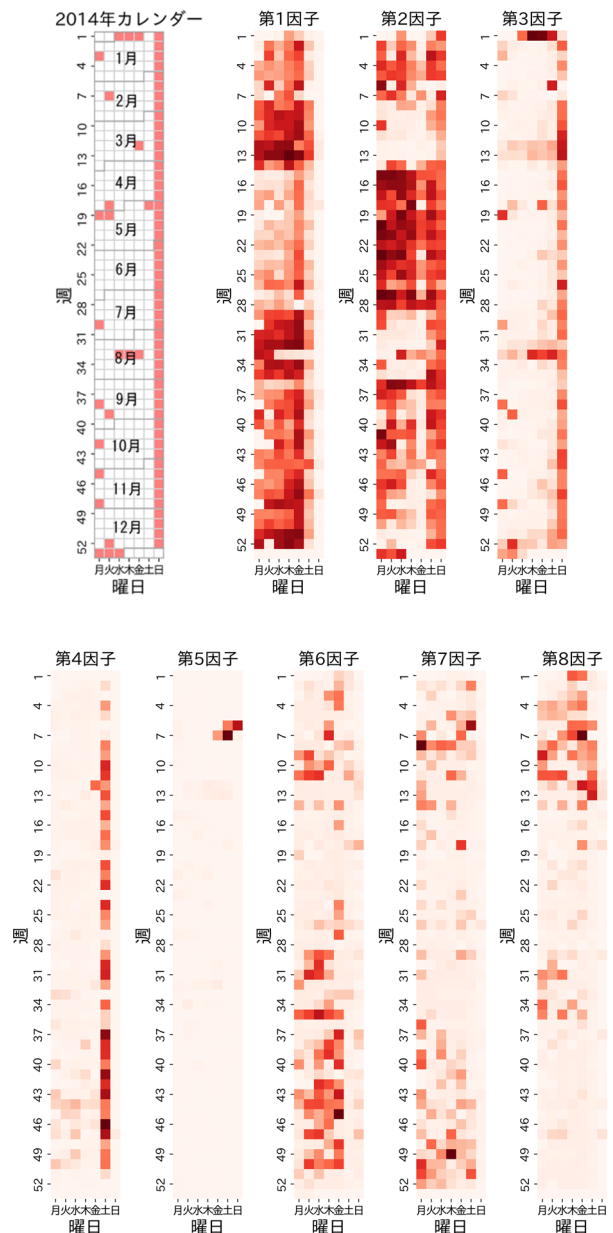


図-3 日モードの因子ベクトル

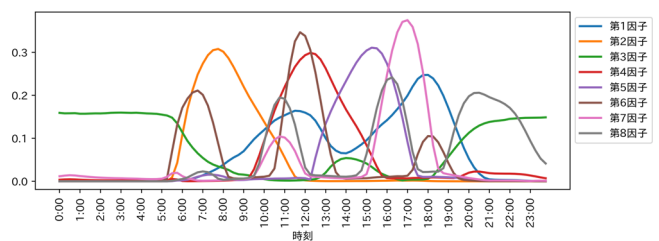


図-4 時刻モードの因子ベクトル

曜日の周期性に加えて、特異な渋滞を示すパターンが個別で抽出されたことが確認できる。

b) 時刻モード

図-4に示す時刻モードの各因子ベクトルは、それぞれ異なる時間帯に特徴を持つパターンが得られた。

例えば、第1因子は、朝から夕方にかけての長い時間

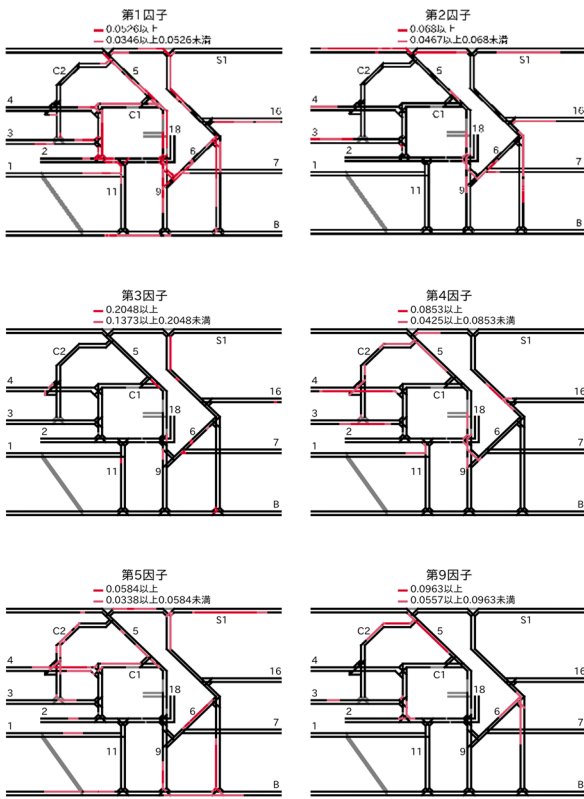


図-5 地点モードの因子ベクトル (一部)

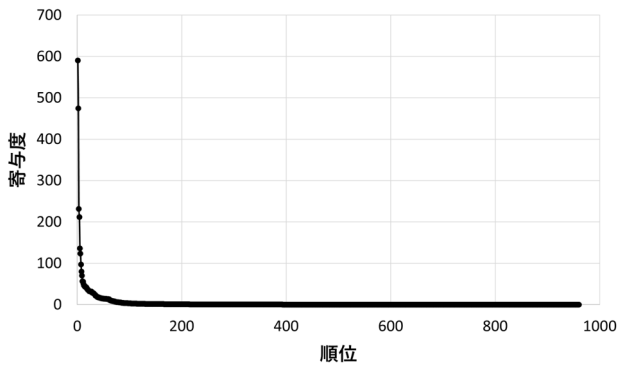


図-6 寄与度の分布傾向

表-2 寄与度の上位 5 パターン

順位	因子ベクトル組み合わせ			寄与度
	日	時刻	地点	
1	1	1	1	590.3
2	1	2	2	474.2
3	2	3	3	231.6
4	1	3	3	211.8
5	2	2	2	136.1

帯に分布し、継続的に交通量が多い日中に発生する渋滞パターンを表す。対して、第3因子は夜間に継続的に発生する渋滞パターンを示している。

一方で、第2因子や第4因子、第7因子は、一部の時間

帯に鋭いピークを持っており、朝や夕方のピークが抽出された。

このように、時刻モードでは、それぞれ異なる時間帯に発生する渋滞を表す日内推移パターンが抽出された。

c) 地点モード (渋滞空間分布)

地点モード (図-5) は、各因子ベクトルの値に基づき k-means法で地点を5分割し、値が大きい2分類に対応するリンクを赤陶酔赤で着色した。

得られた渋滞空間分布の傾向として、JCT付近を先頭に発生する渋滞のパターンが抽出され、第3因子の各JCT付近の局所的な渋滞を除くと、多くはある程度長い区間単位で渋滞が発生していることが見て取れる。

また、各因子ベクトルに注目すると、都心環状線に接続するJCTを先頭にした各路線の上り方面や箱崎JCT付近の空間分布 (第1因子) や、都心部へ向かう上り方面の空間分布 (第2因子) などが抽出された。また、第9因子などで見られる板橋JCTや堀切JCTを先頭にした渋滞は複数因子で見られ、渋滞が発生しやすいボトルネックが存在していることが確認できる。

以上より、地点モードでは、渋滞が同時に発生しやすい区間を表す渋滞空間分布のパターンが抽出されたことが確認できる。

(2) 寄与度を考慮した時空間パターン分析

コアテンソルの各要素は、3モードの因子ベクトルの組み合わせに対する重みで、テンソルデータを復元する際に各因子ベクトルがどの程度寄与するかを表す指標である。そこで、本稿ではこれを「寄与度」と呼ぶ。寄与度が大きい各モードの因子の組み合わせは、元のテンソルに含まれる重要な情報、つまり、顕著な渋滞時空間パターンであると言える。

a) コアテンソル値の分布傾向

寄与度の最大値は590.3、最小値は0.0、平均値は4.0であった。寄与度の分布傾向を示す図-6から、少数の限られた寄与度が大きな値を持つ一方、そのほかの多数の寄与度は0に近い値であることが読み取れる。すなわち、因子ベクトルの全組み合わせ $8 \times 8 \times 15 = 960$ のうち、少数の組み合わせが渋滞の有無を示すテンソルを表現する上で重要なパターンであることが言える。圧縮次元数の設定基準と同様に、最大寄与度の2.5%異常となる組み合わせのみを抽出すると、48組のみとなった。この結果から、限られた顕著な時空間パターンが存在することが示唆される。

b) 寄与度が大きい顕著な時空間パターン

本項では、寄与度が大きい値を持つ因子ベクトルの組み合わせ、つまり、顕著な渋滞時空間パターンの解釈を紹介する。

表-2に上位5組の寄与度の値を示す。日モードの因子

ベクトルに注目すると、上位は5つは第1因子と第2因子である。対して、第6・7・8因子は上位20組には入っていない。時刻・地点モードに関しても同様に、上位に出現する因子ベクトルの番号に偏りが生じている。

特に寄与度が大きい上位の日・時刻・地点パターンとの組み合わせが示す、渋滞の時空間パターンは以下の通りである。

・寄与度最大の組み合わせ (1, 1, 1)

「繁忙期平日」「日中」「都心環状線に接続する JCT を先頭とした」複数路線での渋滞を表す。

・寄与度上位 2 位の組み合わせ (1, 2, 2)

「繁忙期平日」「朝」「各 JCT を先頭に上り線」の渋滞を表す。

・寄与度上位 3 位の組み合わせ (1, 3, 3)

「繁忙期平日」「夜間」「各 JCT 付近」の渋滞を表す。

また、日・時刻・感知器の組み合わせに着目すると、5 位までに時刻モードの第 2 因子・感知器モードの第 2 因子のパターン (2 位 (1, 2, 2), 5 位 (2, 2, 2)) と、時刻モードの第 3 因子・感知器モードの第 3 因子のパターン (3 位 (1, 3, 3), 4 位 (2, 3, 3)) が見られる。前者の組み合わせは、繁忙期の平日と、休日・閑散期の平日で寄与度の値に差があるのに対し、後者の組み合わせは差が小さい。このことから、夜間の各 JCT 付近の渋滞はきせ津や曜日にかかわらず恒常的に発生するのに対し、朝の都心部に向かう上り線の渋滞は繁忙期の平日に発生頻度が高いことが読み取れる。

このように、寄与度を考慮して、因子ベクトルの組み合わせが表す渋滞を解釈することで、顕著な渋滞時空間パターンを捉えることが可能である。

(3) 抽出結果を用いた日・地点の分類と渋滞時空間パターンの考察

提案手法で抽出された因子行列に基づいてクラスター分析を行い、日・地点を分類した結果を示す。

日の分類結果を図-7に示す。重心ベクトルのノルムが大きいクラスターを順に分類1~5とし、赤から青に着色した。各分類は、分類1に、記録的な降雪に見舞われた日が分類され、分類2には土曜日、分類3には日曜・祝日・大型連休のように、平日・休日による違いなどの特徴が分類された。また、分類4は繁忙期の平日、分類5は閑散期の平日と一部の日曜日が分類された。ここで、式(4)で分類 k の時刻・地点モードに対する寄与度 $\mu_{kr_2r_3}$ を算出し、各分類の渋滞の時空間パターンを考察する。

$$\mu_{kr_2r_3} = \sum_{r_1}^{R_1} \lambda'_{r_1r_2r_3} v_{kr_1} \quad (4)$$

最も渋滞レベルの高い分類1は、夜間から朝にかけて継続的に、ネットワークの広い範囲でJCTを先頭とした比較的長い渋滞が発生している日を示しており、記録的な降雪によって首都高全域で渋滞が発生したことが示唆される。分類2や5などの比較的交通量の少ない日は、時刻モードの第1・2因子(図-4)、地点モードの第1・5因子(図-5)の寄与度が大きく、朝や日中にかけてネットワークの広範囲で発生する渋滞パターンが存在することを確認した。一方で、分類4の繁忙期の平日は、寄与度の上位3組が、時刻モードの第1・2・3因子で、地点モードの第1因子であることから、朝から晩にかけて恒常的に、都心環状線とそこに接続する他路線の上り線で渋滞が発生する日であることが確認できた。

次に、地点分類結果を図-8に示す。重心ベクトルのノルムが大きいクラスターを順に分類1~5とし着色した。渋滞発生地点(分類1~4)と、非渋滞発生地点(分類5)に分類された。前者について、式(4)と同様に日・時刻モードに対する寄与度を算出し、それぞれの渋滞発生日時の違いを確認した。例えば、分類2は、日モードの第1因子(図-3)と時刻モードの第1・5因子(図-4)の寄与度が大きく、繁忙期の平日の日中に渋滞が発生しやすい地点を表す。分類4は日モードの第1因子と時刻モードの第1因子と第6因子の寄与度が大きいことから、繁忙期の日中に加えて、6時・12時前後の時間帯に渋滞が発生しやすい地点である。これらの結果は、繁忙期の平日は分類2の地点で日中渋滞が発生し、朝や正午の短時間には分類4の地点まで渋滞が広がることを示している。

以上のように、提案手法のパターン抽出結果を用いた日・地点の分類により、渋滞特性が類似する日・地点が明らかとなった。提案したテンソル分解とその結果を用いたクラスター分析に基づく分析アプローチによって、分解結果の因子の組み合わせと寄与度の解析による渋滞特性の考察に加えて、日や地点によって渋滞パターンが異なることが明らかとなり、より詳細な渋滞の時空間パターンを解明することが可能となった。

(4) グラフラプラシアン正則化の効果の検証

本項では、グラフラプラシアン正則化の効果を検証するため、正則化パラメータ $\alpha = 0$ モデルとの比較分析を行う。分析では、正則化を課した時刻・地点モードの因子ベクトルの結果を比較することで、グラフラプラシアン正則化によるパターン抽出への効果を検証する。圧縮次元数は、本節冒頭で示した基準をもとに(6, 12, 26)に設定した。

時刻モード(図-9)の第1・6因子を見ると、正則化を課さないモデルではパターンがいびつな形状をしているのに対し、正則化を課したモデルは、滑らかで連続的な日内推移パターンとして抽出したことが確認できる。

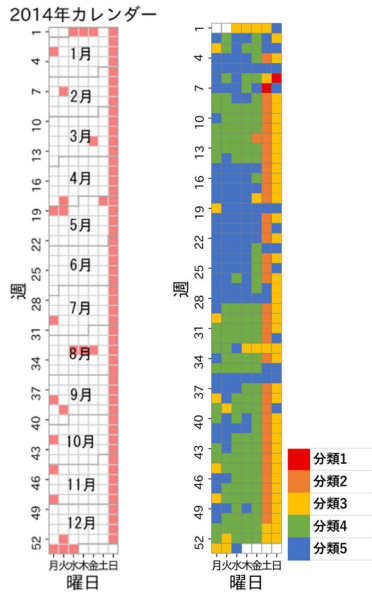


図-7 日の分類結果

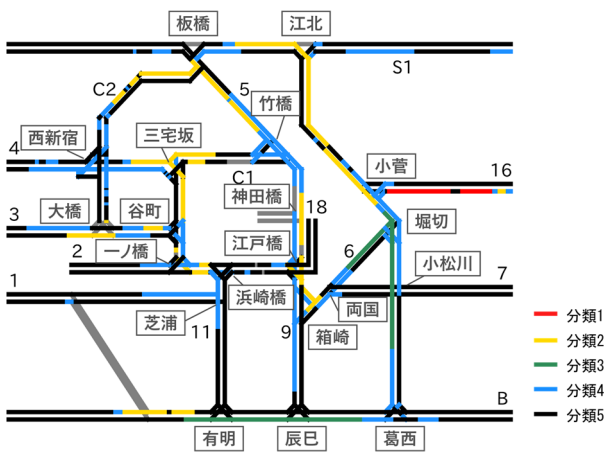


図-8 地点の分類結果

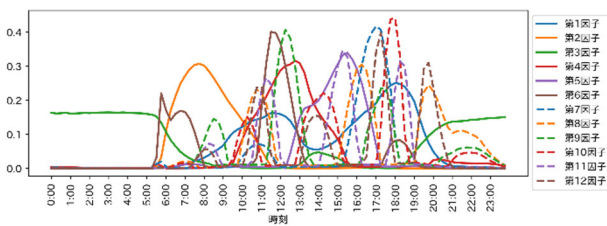


図-9 正則化なし：時刻モードの因子ベクトル

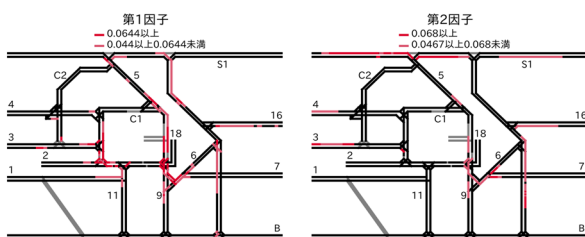


図-10 正則化なし：地点モードの因子ベクトル（一部）

また、地点モード(図-10)の第1因子でも、正則化を課さない場合は比較的短い細切れの区間として抽出された場所でも、正則化を課した場合は1つの連続した渋滞区間として抽出された。これらは、連続する時刻や接続する地点間の因子ベクトルの値を類似の値にする正則化の効果によって、1つの大きな渋滞パターンとして抽出できたと考えられる。

このように時刻と地点に関するグラフ正則化により、連続的な時間帯や区間を1つの大きな渋滞パターンとして抽出できる。結果として、正則化を行わない場合と比較して、より少ない次元数で再現精度が同等の分析が可能になったと考えられる。正則化を設定した場合の方が解釈がしやすく、提案手法が一定の有効性を有することが確認できた。

5. おわりに

本研究は、渋滞の有無を表すデータが二項分布に従い、渋滞が時空間上で連続に発生することを考慮したテンソル分解手法として、グラフラプラス正則化を導入したロジスティック非負値タッカー分解モデルを提案し、首都高速道路での交通観測データへの適用を通して、提案手法の渋滞パターン分析における有効性を確認した。

分析の結果、日・時刻・地点のモード毎に、季節・曜日等の特徴や、時間帯・道路区間に異なる特徴をもつ因子が抽出された。加えて、日・時刻・地点の特徴が組み合わせられた時空間パターンが抽出され、寄与度の上位の組み合わせから、特定の顕著な渋滞時空間パターンが存在することが示唆された。また、抽出結果を用いた日・地点のクラスター分析によって、類似の渋滞特性を持つ日・地点を明らかにし、それぞれの時空間パターンの違いを明らかにした。さらに、グラフ正則化を課さないモデルとのパターン比較分析を通して、渋滞の時空間的な連続性を捉えたパターン抽出の有効性を確認した。

本分析では、グラフ正則化を導入したモデルを、抽出されたパターンの解釈可能性の違いから評価した。しかし、この分析はモデルの近似性能を考慮しておらず、データに対するモデルの説明力に関する議論が不十分である。解釈性とモデル性能の両方を考慮した定量的な正則化パラメータ決定方法の考案を今後の課題とする。

謝辞：本研究を進めるにあたり、首都高速道路株式会社、日本道路交通情報センター(JARTIC)より貴重なデータをご提供いただいた。また、東北大学の酒井高良氏にはデータ整理にご尽力いただいた。ここに記し、感謝を表します。

参考文献

- 1) Inoue, R., Miyashita, A., and Sugita, M.: Mining spatio-temporal patterns of congested traffic in urban areas from traffic sensor data. In: *Proceedings of 2016 IEEE 19th International Conference on Intelligent Transportation Systems*, 731-736, 2016.
- 2) Ji, Y., Oeda, S., and Geroliminis, N.: Empirical observations of congestion propagation and dynamic partitioning with probe data for large-scale systems. *Transportation Research Record*, (2422), 1-11, 2014.
- 3) Yang, S., Wu, J., and Yang, T.: Revealing heterogeneous spatiotemporal traffic flow patterns of urban road network via tensor decomposition-based clustering approach. *Physica A: Statistical Mechanics and its Applications*, 526, No. 120688, 2019.
- 4) Han, Y and Moutarde, F.: Analysis of large-scale traffic dynamics in an urban transportation network using non-negative tensor factorization. *International Journal of intelligent Transportation systems Research*, 14(1), 36-49, 2016.
- 5) 馬場 静羽, 井上 亮, 酒井 高良, 赤松 隆: テンソル分解に基づく首都高速道路の交通渋滞時空間パターンの解明テンソル. 土木計画学研究・講演集, **65**:46-2, 2022.
- 6) Hong, D., Kolda, G., and Duersch, A.: Generalized canonical polyadic tensor decomposition. ArXiv: 1808.07452v2, 2019.
- 7) Qiu, Y., Zhou, G., and Xie, S.: Graph regularized nonnegative decomposition for tensor data representation, In: *Proceedings of 2019 IEEE International Conference on Acoustics Speech and Signal Processing*, 8613-8617, 2019.
- 8) Zhang, J., Lee, Y.: Sparse Logistic Tensor Decomposition for Binary Data. *arXiv: 2106.14258*, 2021.
- 9) Ko, J.: Characterization of Congestion Based on Speed Distribution: A Statistical Approach Using Gaussian Mixture Model. *TRB 2005 Annual Meeting CD-ROM*, 2005.
- 10) Zhao, W., McCormack, E., Dailey, D.J., Scharnhorst, E.: Using Truck Probe GPS Data to Identify and Rank Roadway Bottlenecks. *JOURNAL OF TRANSPORTATION ENGINEERING* , 139(1), 1-7, 2-13.
- 11) Kittler, J and Illingworth, J.: Minimum error thresholding, *Pattern Recognition*, 19, 41-47, 1986.
- 12) 白石智良, 赤羽弘和, 小根山裕之, 田中伸治.: 都市高速道路における臨界速度の自動設定法の開発とボトルネック容量分析への適用. 交通工学研究発表会論文集, Vol.32, Np.11, 2013.
- 13) 坪田隆宏, 吉井稔雄, 原田日郎, 神野裕昭.: 一般道路における路線 QV 式を用いた速度推定モデル. 交通工学論文集, 第 4 巻, 第 1 号(特集号 A), pp.A_216-A_222, 2018.

(2023. 3. 6 受付)

PATTERN EXTRACTION OF CONTINUOUS CONGESTION IN SPATIO-TEMPORAL SPACE BASED ON LOGISTIC TUCKER DECOMPOSITION

Shunsuke KISO and Ryo INOUE

In order to mitigate traffic congestion, it is important to understand the spatio-temporal patterns of congestion on the road network. Recently, tensor decomposition has attracted attention as a method for extracting features from data with complex relationships. Tensor decomposition, which is the representation of higher-order tensor data as the sum of products of lower-order tensors, has been applied to traffic data analysis, such as traffic condition prediction and anomaly detection. However, there is little discussion on decomposition methods that correspond to the probability distribution of the components of the tensor to be analyzed, or decomposition methods that take into account the spatio-temporal continuity of traffic congestion. This study proposes a logistic non-negative Tucker decomposition with graph Laplacian regularization, which can take into account the distribution of tensor data and the spatio-temporal continuity of traffic congestion, and confirms the effectiveness of the proposed method using observed congestion data from the Metropolitan Expressway network.