

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

馬場 静羽¹・井上 亮²・酒井 高良³・赤松 隆⁴

¹ 学生会員 東北大学 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: shizuha.baba.s5@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

³ 学生会員 東北大学 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: takara.sakai.t1@dc.tohoku.ac.jp

⁴ 正会員 東北大学教授 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: akamatsu@plan.civil.tohoku.ac.jp

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。テンソル分解は、階数の多いテンソルで表されるデータを、より階数の少ない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方法であり、日・時刻・場所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いる。しかし、渋滞特性把握を目的とした例は限定的な上、抽出された時間的特徴と空間的特徴の個別解釈に留まり、その間の関係性の解釈や、抽出結果を用いた渋滞特性の議論は十分に行われていない。そこで本研究では、首都高速道路網の渋滞時空間パターンの解明を目指して、テンソル分解の抽出結果を用いた分析により、日・時刻・地点に関する渋滞パターンの解析を行った。その結果、顕著な時空間パターンや、渋滞パターンが類似する日や地点を明らかにした。

Key Words: *congestion pattern, tensor decomposition, non-negative tensor factorization, expressway network, long-term observation data*

1. はじめに

交通渋滞は、世界中の都市で解決すべき問題として認識されている。その緩和を目指し、短期的将来予測¹⁾や典型的なパターン抽出²⁾、突発事象の検出³⁾等を目的に、都市内道路ネットワーク上の交通状態の特徴を分析する研究が行われてきた。近年、情報処理技術の発展に伴い膨大なデータを用いた分析が可能になったことで、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⁴⁾。

交通状態の特徴を捉える上で注目すべきは、交通現象が有する時間的特徴と空間的特徴である。時間的特徴とは、季節毎や週毎、日毎に周期的に交通状態が推移する性質であり、空間的特徴とは、渋滞多発地点が存在するように、場所によって混雑状況が異なる性質である。さらに、渋滞多発地点毎に混雑が発生する条件や時期が異なるように、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、交通状態の特徴を把握する上では、この時空間パターンの解明が重要である。

酒井・赤松 (2019)⁵⁾は、首都圏高速道路網で観測された車両感知器データに基づき、渋滞空間分布とその推移過程の特性を分析した。クラスター分析手法であるk-means法による分析から、年間を通して渋滞空間分布の推移に定常性が存在することを明らかにした。しかし、一般的なクラスター分析に基づいた時空間パターン解析では、時間方向・空間方向に関して段階を踏んで分析を行う必要があり、分析手順によって結果が異なることが懸念される。

本研究は、多項関係を有するデータから特徴抽出を行なうテンソル分解に注目する。テンソル分解は、行列分解を多次元配列 (テンソル) に拡張した手法で、複雑な多項関係を有するデータの特徴を抽出できる。交通観測データが時間的特徴・空間的特徴を組み合わせられた複雑な関係性を含む多次元データであることに着目すると、交通データ分析においてテ

ンソル分解が有用であると期待できる。

ここで、交通観測データにテンソル分解を適用しパターン抽出を行なった先行研究を紹介する。Sun and Axhausen (2016)⁶⁾ は、シンガポールの交通ICカードデータから集計したトリップ数を、時間帯・乗客属性・出発ゾーン・到着ゾーンの4モードに分解した。その結果、シンガポールのマスタープランの地域区分と一致する発着地の空間パターンや、通勤帰宅時のCBD発着の移動パターンが抽出された。テンソル分解による交通データの特徴抽出に有用であることが確認されたが、これらの抽出結果は自明であり、移動特性に関する新たな知見が得られたとは言いがたい。Yang et al. (2019)⁷⁾ は、大規模道路ネットワーク上の車両感知器で観測された速度と密度のそれぞれにテンソル分解を適用し、日・時間・道路区間方向に分解した。得られた低次元潜在空間を基に道路区間を分類し、巨視的な交通状態指標により各サブネットワークの交通状態特性を明らかにした。しかし、分解結果に関する議論は、時間的特徴と空間的特徴を個別に解釈するに留まり、それらを組み合わせた時空間パターンについて十分に議論されていない。

以上を踏まえ、本研究は、都市内高速道路網における交通渋滞の時空間パターンの解明を目的とし、テンソル分解を用いた渋滞特性把握を試みる。具体的には、首都高速道路を対象に、車両感知器データから渋滞時空間分布を表すデータを生成し、そのテンソル分解からまず、日・時刻・地点の特徴を抽出する。さらに抽出された日・時刻・地点の特徴の組み合わせで表される時空間パターンの解明を目指し、日・地点をクラスター分析で分類し、日の分類に対応した渋滞時刻・渋滞地点の特徴、および、地点の分類に対応した渋滞日・渋滞時刻の特徴を明らかにする。

本論文の構成は以下の通りである。2. で分析手法を示し、3. で使用データとパラメータである圧縮次元数の設定方法を示す。4. で分析結果を示し、渋滞の時空間パターンを考察する。最後に、5. で本研究のまとめと課題を述べる。

2. 分析手法

(1) テンソル分解による特徴抽出

テンソル分解は、行列分解をテンソルに拡張した手法で、テンソルデータの各要素軸（モード）の次元を圧縮する際に用いられる。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⁸⁾とCP分解⁹⁾に分類される。タッカー分解は、テンソルを各モード

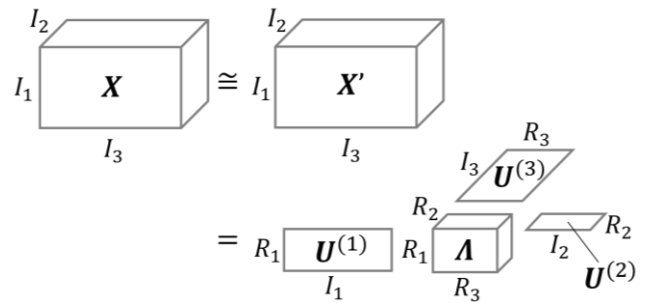


図-1 テンソル分解の概念図（3階テンソルの場合）

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（図-1）。以下、3階テンソル $\mathbf{X} \in \mathcal{R}^{I_1 \times I_2 \times I_3}$ を例に説明する。タッカー分解モデル \mathbf{X}' は、

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

で表され、 $\mathbf{A} \in \mathcal{R}^{R_1 \times R_2 \times R_3}$ はコアテンソル、 $\mathbf{U}^{(i)} \in \mathcal{R}^{I_i \times R_i}$ ($i = 1, 2, 3$) 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。式(1)のパラメータは \mathbf{A} 、 $\mathbf{U}^{(i)}$ で、式(2)の最適化問題で推定する。

$$\min_{\mathbf{A}, \mathbf{U}^{(1)}, \mathbf{U}^{(2)}, \mathbf{U}^{(3)}} \|\mathbf{X} - \mathbf{X}'\|_F^2 \quad (2)$$

ここで、 $\|\cdot\|_F$ はフロベニウスノルムを表す。式(3)のように因子行列をベクトルで表記すると、タッカー分解モデルは式(4)で書け、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。

$$\mathbf{U}^{(i)} = [\mathbf{u}_1^{(i)} \mathbf{u}_2^{(i)} \dots \mathbf{u}_{R_i}^{(i)}] \quad (3)$$

$$\mathbf{X}' = \sum_{r_1=1}^{R_1} \sum_{r_2=1}^{R_2} \sum_{r_3=1}^{R_3} \lambda_{r_1 r_2 r_3} (\mathbf{u}_{r_1}^{(1)} \circ \mathbf{u}_{r_2}^{(2)} \circ \mathbf{u}_{r_3}^{(3)}) \quad (4)$$

CP分解は、タッカー分解のコアテンソルが対角との制約が加わるため、得られた因子ベクトルは一対一に対応する。一方、タッカー分解は、モード毎に異なる次元数を設定でき、圧縮次元数の設定に関して自由度が高い上、より様々なパターンでモデルを表現するため、複雑な多項関係を含むデータ分析に適している。

さらに、式(2)の最適化問題に非負制約¹⁰⁾やスパース制約¹¹⁾など、様々な制約を課した手法が提案されている。その中で、モデルが和算のみで表現される非負制約付きのテンソル分解は、複雑な関係性が内在するデータ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く、適用例が多く見られる^{12),13),14)}。

以上より、本研究では、式(2)にコアテンソルと因子行列の要素すべてが非負であるという制約を加えた、非負値タッカー分解¹⁵⁾を用いて、日・時刻・地

点の渋滞パターンを抽出する. なお, 本分析では, 結果の解釈を容易にするため, 各因子ベクトルをノルムが1になるよう正規化する(式(5)). 以降は, 正規化後の因子ベクトル, コアテンソルを用いる.

$$X' = \sum_{r_1}^{R_1} \sum_{r_2}^{R_2} \sum_{r_3}^{R_3} \lambda'_{r_1 r_2 r_3} (\mathbf{u}'_{r_1}^{(1)} \circ \mathbf{u}'_{r_2}^{(2)} \circ \mathbf{u}'_{r_3}^{(3)}) \quad (5)$$

$$\mathbf{u}'_{r_i}^{(i)} = \frac{\mathbf{u}_{r_i}^{(i)}}{\|\mathbf{u}_{r_i}^{(i)}\|} \quad (i = 1, 2, 3)$$

$$\lambda'_{r_1 r_2 r_3} = \|\mathbf{u}_{r_1}^{(1)}\| \|\mathbf{u}_{r_2}^{(2)}\| \|\mathbf{u}_{r_3}^{(3)}\| \lambda_{r_1 r_2 r_3}$$

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あり, 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.

(2) テンソル分解結果を用いたクラスター分析

渋滞の時空間パターンを解釈するためには, テンソル分解を通して抽出された, 日・時刻・地点の特徴を表す因子ベクトル間の関係の強弱, すなわち寄与度の大小に注目する必要がある. しかし, 寄与度はコアテンソルの要素数 $R_1 \times R_2 \times R_3$ のデータがあり, そこから特徴を捉えることは容易ではない.

そこで, 本研究では, テンソル分解による分析結果をより縮約するため, 非負値タッカー分解で得られた因子行列を用いて, 各モードについてクラスターリングを行なう分析方法を提案する.

非負値タッカー分解で得られた因子行列 $U^{(i)}$ の l 番目の行ベクトルを, モード i の l 番目の「特徴ベクトル」と呼ぶ. この特徴ベクトルはモード i における l 番目の特徴量を表し, これを用いてクラスター分析を行なうことで, データに内在する特徴を考慮した分類が可能である.

クラスター分析には k-means 法¹⁶⁾を用いる. k-means 法は非階層クラスターリング手法の一つで, 式(6)の評価関数 ϕ を最小化するクラスター重心を見つけ, データ X を任意の k 個のクラスターに分割する.

$$\phi = \sum_{x_m \in X} \min_{l \in k} \|x_m - c_l\|^2 \quad (6)$$

ここで, x_m ($m \in \{1, 2, \dots, n\}$) は各データ, c_l はクラスター $l \in \{1, 2, \dots, k\}$ の重心を表す. つまり, k-means 法は, 各データ点から最も距離が近いクラスター重心との距離の総和が, 最小となるようなクラスター重心を求めて分類する.

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 | | |

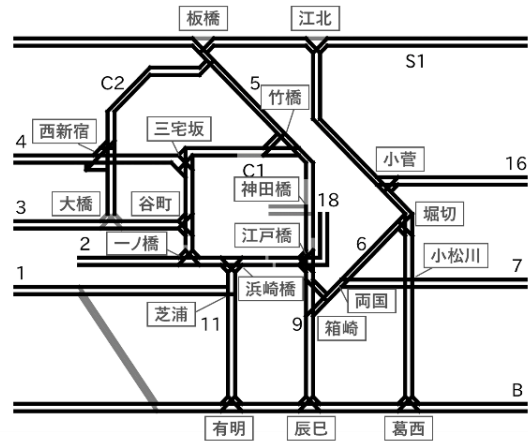


図-2 首都高速道路の模式図 (図上に主な JCT 名を示す)

さらに, 例えばモード 1 について, k-means 法による分類後, モード 1 のクラスター重心と他モードの因子行列を用いて, コアテンソル Λ' を用いてテンソルを再構成することで, クラスター単位のテンソル表現を得る (式(7)).

$$Y = \Lambda' \times V \times U^{(2)} \times U^{(3)} \quad (7)$$

$$V = [v_1 v_2 \dots v_k]^T$$

ここで, Y は再構成テンソル, v_k は k 番目のクラスターにおける重心ベクトルである.

以上をまとめると, 本分析では, 非負値タッカー分解により交通渋滞のパターンを抽出し, その後, k-means 法によるクラスター分析で, 類似の渋滞が発生する日や, 同時に渋滞が発生しやすい地点を特定し, それぞれの渋滞時空間パターンを考察する.

3. 使用データ・圧縮次元数の設定

(1) 使用データ

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-2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分析する。

渋滞現象分析にあたり、各感知器の地点における時々刻々の渋滞状態を表すデータを生成する。酒井・赤松 (2019)⁵⁾ は「信号のない高速道路網においては、速度の低下をそのまま渋滞と判定しても差し支えない」とし、日 d 、時刻 t 、地点 i の渋滞状態変数 $\delta_{d,t,i}$ を、式(8)の二値変数（渋滞/非渋滞）で定義し、渋滞判定速度 v^* を20km/hに設定した分析結果を示した。本分析でも同様に、時間平均速度から渋滞判定を行い、

$$\delta_{d,t,i} = \begin{cases} 1 & \text{if } (v_{d,t,i} < v^*) \\ 0 & \text{if } (v_{d,t,i} \geq v^*) \end{cases} \quad (8)$$

で渋滞状態変数を生成した。

次に、渋滞状態変数を集計し、15分間隔の渋滞回数データを生成する。幅120分のガウシアンカーネルを用い、15分間隔で集計した渋滞状態変数を時間方向に平滑化した値を用いる。各要素の最小値は0、最大値は1である。なお、データが欠損している時間帯・地点に関しては、渋滞回数を0回とした。このデータは「15分間で渋滞と判定された時間の割合」を意味するが、本稿では「渋滞回数」と呼ぶ。

最後に、渋滞回数データに対して、日モード・時刻モード・地点モードの3軸を設定し、365日×96時刻×1441地点の渋滞回数テンソルを生成した。

(2) 圧縮次元数の設定

本節では、本分析における圧縮次元数、すなわちコアテンソルの次元数の設定を説明する。テンソル分解手法における圧縮次元数の設定は、各モードのパターン数を決める重要な問題だが、その設定方法は未だ確立されていない。例えば、桑野ら(2017)¹⁴⁾は、4階テンソルに非負値タッカー分解を適用する際、データの特徴が最も表現されるよう、試行錯誤で次元数を設定している。しかし、この次元数が分析上適切な設定である保証はない。また、Han and Moutarde (2014)¹⁷⁾は、3階テンソルに正規化非負値CP分解を適用する際、テンソルデータとモデルの誤差指標を用いて、モデルの精度を担保しつつ少ない要素数でモデルを表現できる次元数に設定している。しかし、Han and Moutarde (2014)¹⁷⁾が用いたCP分解のハイパーパラメータは1つだが、本分析が用いるタッカー分解では複数のハイパーパラメータが存在するため、この方法を適用することは困難である。

本分析では、各因子ベクトルの組み合わせに関して、元データに対する重要度を表す寄与度に着目した、定量的な次元数の設定方法を提案する。提案方

法では、テンソルデータに内在する因子ベクトルを可能な限り多く表現しながら、寄与度が十分に小さく、意味解釈上重要ではない因子ベクトルが抽出されない次元数の設定を目指す。

次元数設定において指標とする、次元数 $\mathbf{R} = (R_1, R_2, R_3)$ における寄与度の代表値 $\lambda_{\mathbf{R}}$ を式(9)で算出する。

$$\begin{aligned} \lambda_{\mathbf{R}} &= \min_i (\lambda_{r_i \max}) & (9) \\ \lambda_{r_1 \max} &= \max_{r_2, r_3} (\lambda_{r_1 r_2 r_3}) \\ \lambda_{r_2 \max} &= \max_{r_1, r_3} (\lambda_{r_1 r_2 r_3}) \\ \lambda_{r_3 \max} &= \max_{r_1, r_2} (\lambda_{r_1 r_2 r_3}) \end{aligned}$$

$\lambda_{r_i \max}$ は、因子ベクトル $\mathbf{u}_{r_i}^{(i)}$ に関連する寄与度の最大値を表す。 $\lambda_{\mathbf{R}}$ は、ある次元数設定において抽出された全因子ベクトルのうち、 $\lambda_{r_i \max}$ が最小となる値を指す。すなわち、渋滞回数テンソルデータを表現する際に、最も寄与度の低い因子ベクトルが含まれる組み合わせが、最大でもモデルの表現にどの程度寄与しているかを意味する値である。これを「最小因子寄与度」と呼ぶ。最小因子寄与度の小さい次元数では、元データ内のあまり意味を持たないノイズがパターンとして抽出されていると考えられる一方で、最小因子寄与度の大きい次元数では、十分にパターンを抽出できていない可能性がある。そこで本分析では、最小因子寄与度が最大寄与度 $\max_{r_1, r_2, r_3} (\lambda_{r_1 r_2 r_3})$ に対して1%より大きいという条件下で、最小因子寄与度が最小となる次元数に設定する。

4. 渋滞時空間パターンの分析

(1) 非負値タッカー分解による特徴抽出結果

3. (1)で生成した3階の渋滞回数テンソルに対し、非負値タッカー分解を適用した。圧縮次元数は、6~16の計1331通りの中から、3. (2)に示した次元数設定方法により、 $\mathbf{R} = (13, 7, 11)$ に設定した。非負値タッカー分解によって、日・時刻・地点の各モードに関する因子行列(365×13, 96×7, 1441×11)とコアテンソル(13×7×11)を得た。モード i ($i = 1, 2, 3$)における r_i 番目の因子ベクトルを第 r_i 因子と記載する。

日モード(図-3)は、各因子ベクトルの最大値と最小値を基準としたカラースケールで表す。また、2014年カレンダーの着色日は、日曜・祝日・大型連休等の休日を表す。前半の因子ベクトルには、季節性や曜日の周期性が現われた。例えば、第1因子の繁忙期

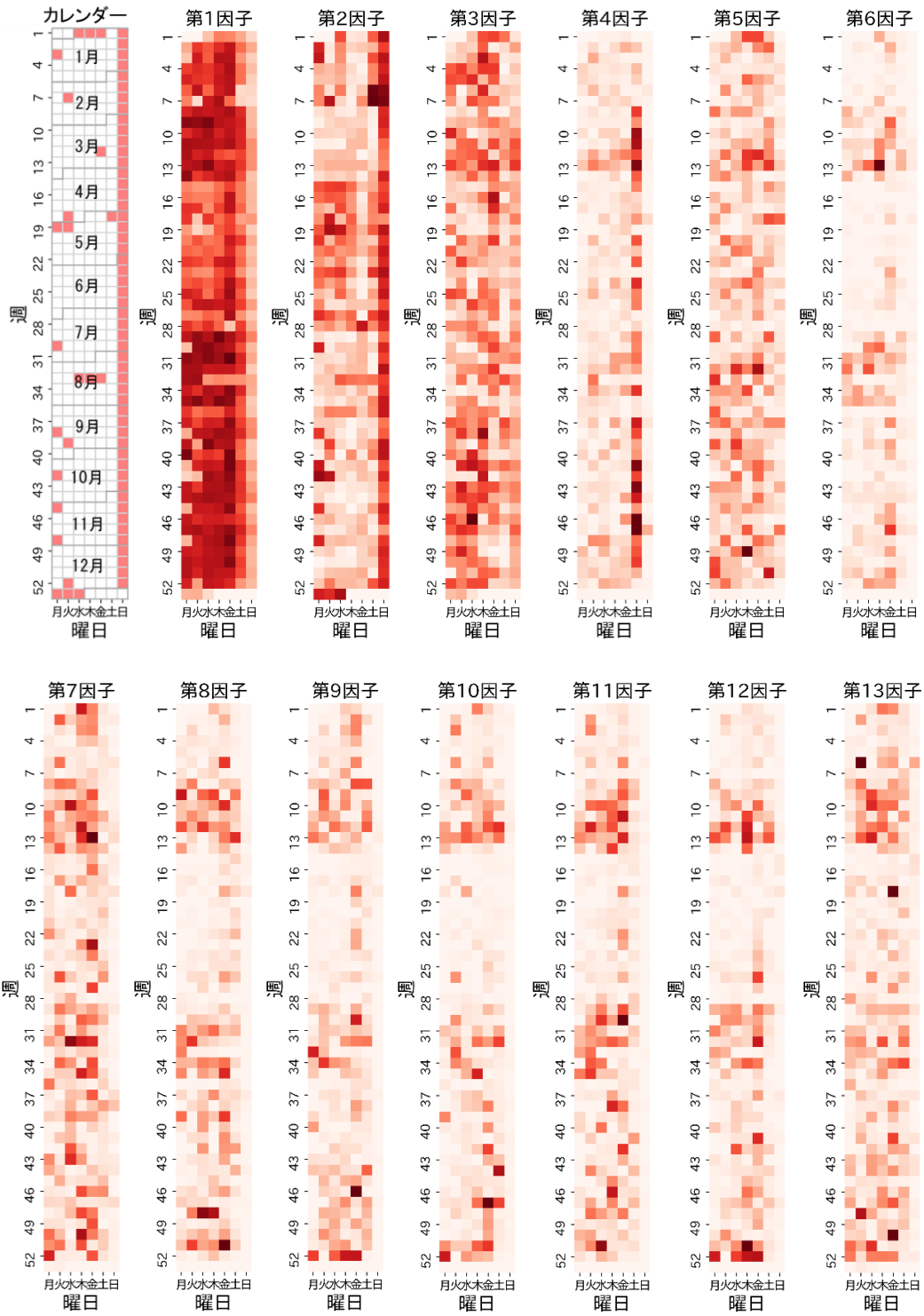


図-3 日モードの因子ベクトル

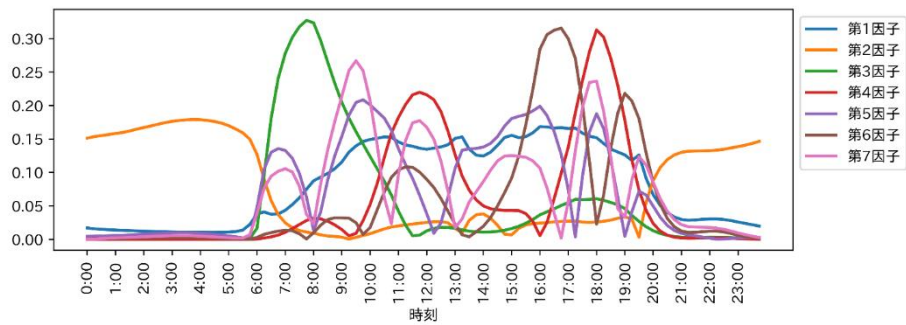


図-4 時刻モードの因子ベクトル

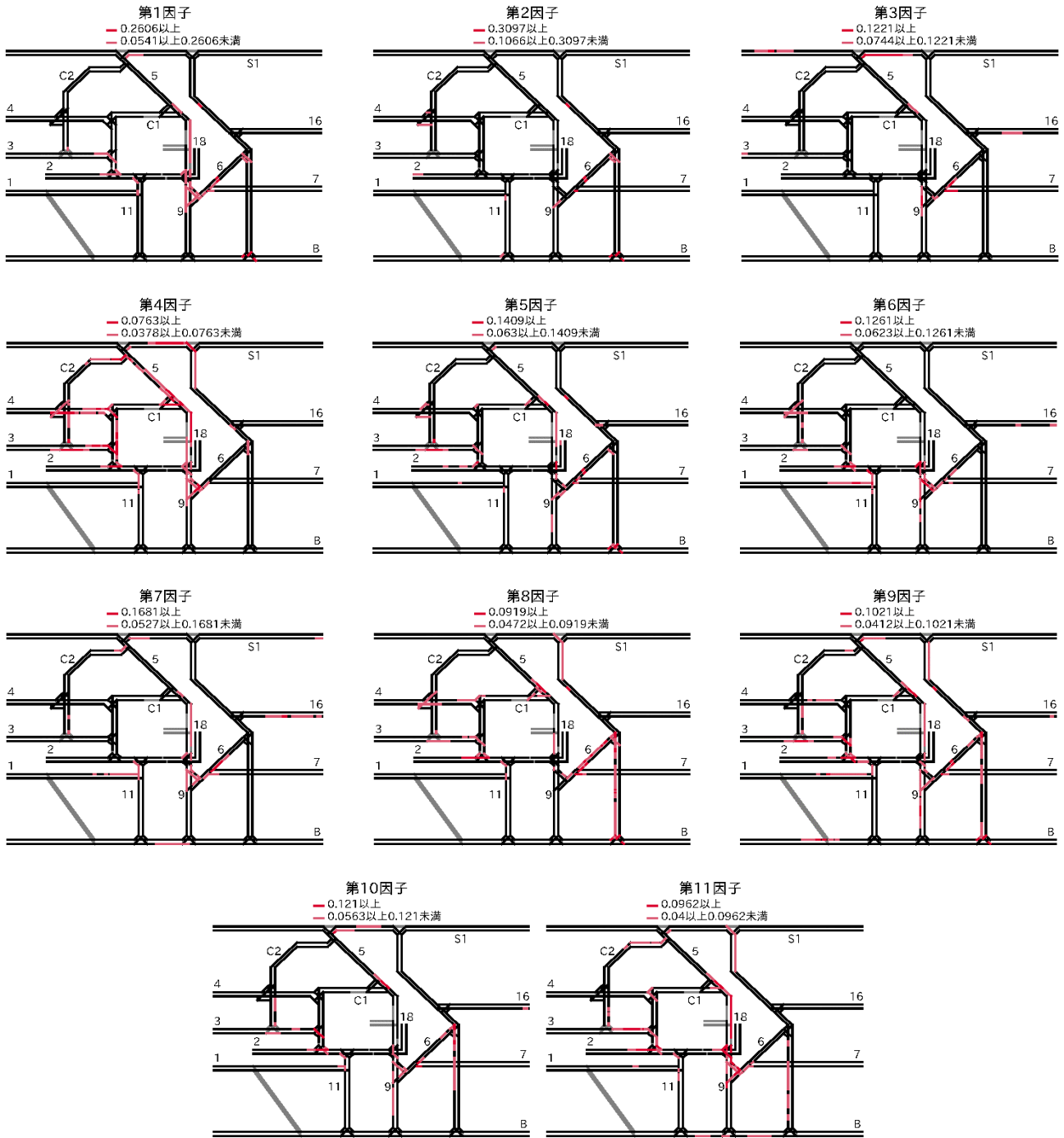


図-5 地点モードの因子ベクトル

平日、第 2 因子の休日や閑散期平日、第 4 因子の土曜日が該当する。一方で後半の因子ベクトルには、休日前後や降雨日、通行止め等の交通規制が実施された日など、非常日が特徴として現れた。

時刻モード (図-4) は、朝・夕ピークや日中・深夜など、それぞれ異なる時間帯を表す日内推移パターンが抽出された。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 4 分類し、平均値が大きい 2 分類に対応するリンクを赤、薄い赤で着色した。箱崎 JCT 付近や上り方面の空間分布 (第 1 因子)、上下線の広範囲に広がる空間分布 (第 4 因子)

などが抽出され、同時に渋滞が生じる箇所が明らかとなった。また、箱崎 JCT 付近は複数の因子ベクトルに共通して出現する一方で、第 3 因子の 5 号池袋線 (5) 上りや、第 9 因子の湾岸線 (B) 東行きは、限定的な因子ベクトルで出現するといった、地点による違いが見られた。

次に、モード間の関係性を表す寄与度から、渋滞時空間パターンについて考察する。寄与度の最大値は 251.4、最小値は 0.0、平均値は 2.0 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ち、多数の寄与度は 0 に近いことが読

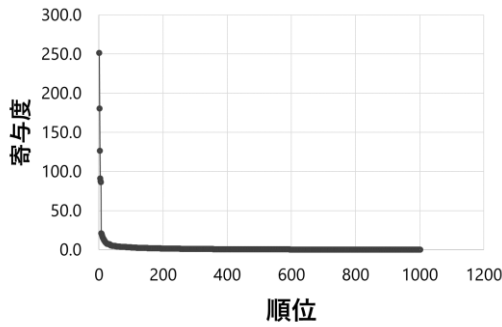


図-6 寄与度の分布傾向

表-2 寄与度（上位 20）

| 順位 | r_1 | r_2 | r_3 | 寄与度 |
|----|-------|-------|-------|-------|
| 1 | 1 | 2 | 2 | 251.4 |
| 2 | 1 | 3 | 3 | 180.6 |
| 3 | 1 | 1 | 1 | 126.5 |
| 4 | 1 | 4 | 4 | 91.0 |
| 5 | 2 | 1 | 2 | 87.7 |
| 6 | 2 | 2 | 2 | 86.5 |
| 7 | 2 | 3 | 2 | 20.9 |
| 8 | 9 | 4 | 4 | 20.8 |
| 9 | 2 | 2 | 1 | 19.3 |
| 10 | 11 | 4 | 4 | 17.6 |
| 11 | 3 | 3 | 3 | 16.0 |
| 12 | 8 | 3 | 3 | 14.9 |
| 13 | 1 | 3 | 2 | 14.7 |
| 14 | 3 | 2 | 2 | 13.7 |
| 15 | 1 | 4 | 11 | 12.4 |
| 16 | 2 | 2 | 5 | 12.2 |
| 17 | 2 | 3 | 3 | 11.4 |
| 18 | 4 | 3 | 3 | 10.8 |
| 19 | 5 | 2 | 2 | 10.7 |
| 20 | 1 | 6 | 6 | 9.7 |

み取れる。この結果は、限られた顕著な時空間パターンが存在することを示唆している。

ここで、寄与度が大きい組み合わせ、つまり顕著な渋滞時空間パターンに注目する。表-2に上位20の寄与度を示す。日モードの因子ベクトルの番号に注目すると、上位4位を「平日」のパターンの第1因子が占めて、次いで3つは第2因子が表す「休日」のパターンが続いている。第6因子、第7因子、第10因子は上位20番目までには入っていない。時刻・地点モードについても同様に、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3位までの日・時刻・地点のパターン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, 2, 2)は、「平日」「深夜」「各JCT付近」の渋滞を表す。次に、寄与度2位の組み合わせ(1, 3, 3)は、「平日」「朝」「板橋JCTを先頭にした中央環状線内回り、5号

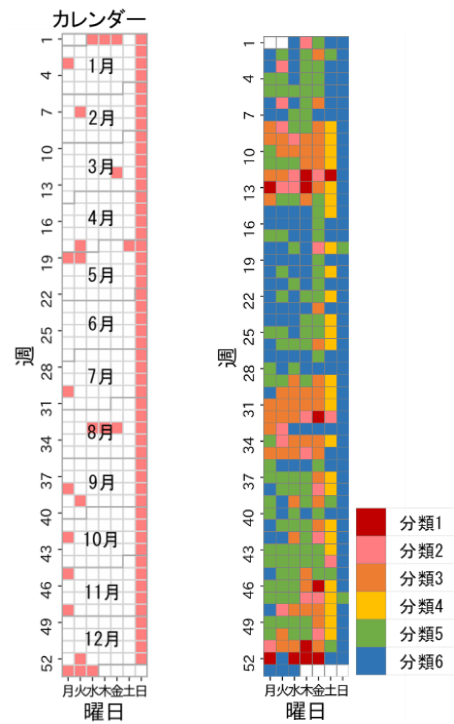


図-7 日の分類結果

池袋線・6号三郷線・9号深川線の上り」の渋滞を表す。そして、寄与度3位の組み合わせ(1, 1, 1)は、「平日」「日中」「主に、箱崎JCTを先頭にした都心環状線外回り、江戸橋JCTを先頭にした9号深川線の上り」の渋滞を表す。このように上位3位は、日の第1因子は共通しているが、時刻・地点モードの因子は異なる組み合わせであり、各時間帯で顕著に渋滞が発生する場所は異なることが読み取れる。

以上のように、非負値タッカー分解により、渋滞回数テンソルに内在する日・時刻・地点のパターンが抽出され、日・時刻・地点における渋滞発生規則性や、顕著な渋滞時空間パターンが明らかとなった。

(2) 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

非負値タッカー分解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1~6とし、赤から青に着色した。各分類は繁忙期の平日、閑散期の平日、連休前後の平日、土曜日、休日などの季節性や周期性などの特徴を持つ。また、式(10)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。

$$\mu_{kr_2r_3} = \sum_{r_1}^{R_1} \lambda'_{r_1 r_2 r_3} v_{k, r_1} \quad (10)$$

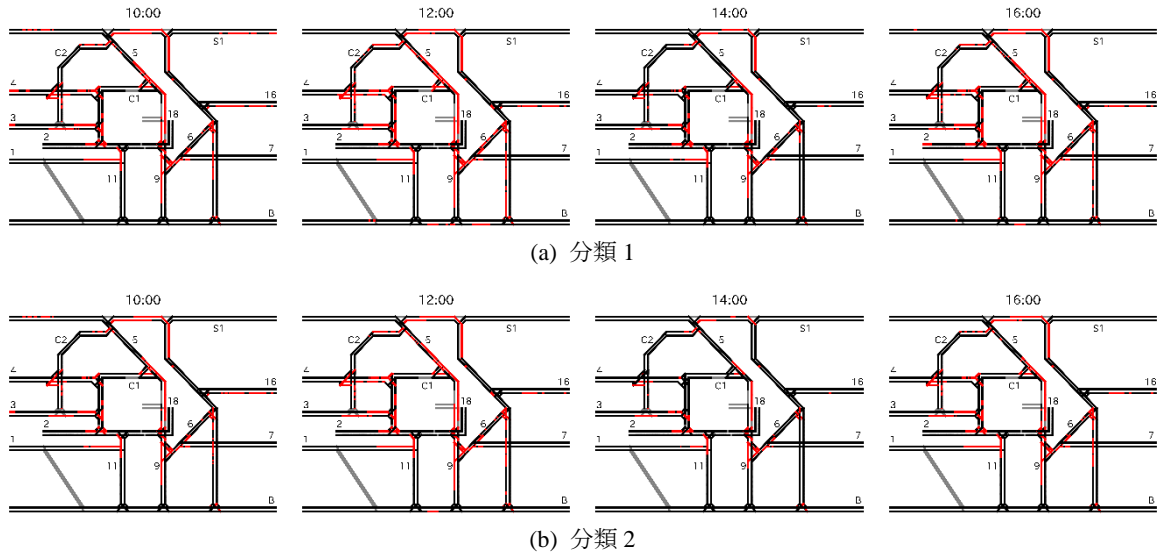


図-8 渋滞空間分布の推移

最も渋滞レベルの高い分類1は、朝から晩まで恒常的に、地点モードの第4因子が示す広範囲の渋滞が発生する日を示す。一方、分類2は、分類1に比べて夕方と日中に広範囲の渋滞の発生が弱い。また、分類3は、12時・18時をピークに、1号新宿線上りや、6号三郷線上り、湾岸線西行き方面で渋滞が発生する日であることが確認できた。

ここで、それぞれの日分類で生起する渋滞の日内推移と空間分布を確認するために、各分類の渋滞回数行列 Y_k を生成した (式 (11))。

$$Y_k = \sum_{r_2}^{R_2} \sum_{r_3}^{R_3} \mu_{kr_2r_3} (\mathbf{u}'_{r_2}{}^{(2)} \circ \mathbf{u}'_{r_3}{}^{(3)}) \quad (11)$$

渋滞回数行列は、渋滞空間分布の推移として表現できる。図-8 に分類 1 と分類 2 の渋滞空間分布の推移を示す。各分類において、時々刻々の渋滞空間分布と、渋滞の延伸・解消の様子を確認できる。また、分類間で、同時刻における渋滞空間分布の比較や、推移の比較が可能である。例えば、分類 1 と分類 2 では、10 時の川口線 (S1) 上りや、14 時頃の渋滞解消の様子に違いが見られ、渋滞レベルの高い分類 1 では、特に渋滞区間が長く、夕ピーク前の解消傾向が弱いことが分かった。

次に、地点分類結果を図-9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、非発生地点 (分類5) に分類された。前者について、式 (10) と同様に算出した日・時刻モードに対する寄与度から、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因子と時刻モードの第1・3・4 因子の寄与度が大きく、平日朝・日中・夕方に渋滞が

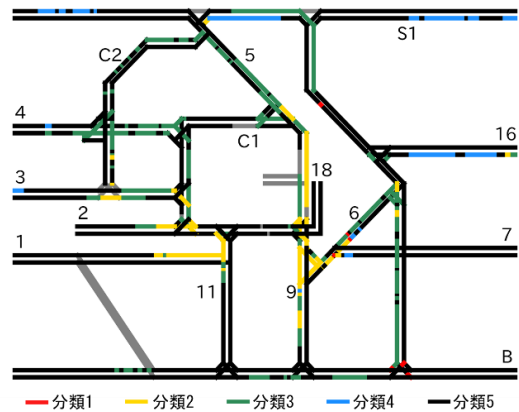


図-9 地点の分類結果

発生する地点を表す。分類3は日モードの第1因子と時刻モードの第4因子の寄与度が突出しており、平日の12時・18時前後の時間帯に渋滞が発生する地点を表す。これらの結果は、平日は分類2の地点で長時間渋滞し、12時・18時前後には分類3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、非負値タッカー分解の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。各分類において、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、時刻・地点に関する渋滞回数行列の表現によって、渋滞の延伸・解消の様子など、より詳細な渋滞の時空間パターンを捉えた。

5. おわりに

本研究は、首都高速道路の15路線を対象に、交通観測データから生成した渋滞の時空間分布データに対してテンソル分解を適用し、渋滞発生の時間的・空間的特徴を分析した。

多項関係を有するデータのパターン抽出への有用性が確認されている非負値タッカー分解を用いて、日、時刻、地点の要素軸を持つコアテンソルと3つの因子行列に縮約した。渋滞発生パターンとして解釈可能な因子が抽出され、非負値タッカー分析の渋滞パターン分析への適用可能性を確認した。

分析の結果、モード毎に、季節・曜日等の周期性や、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された時空間パターンが抽出され、モード間の関係の強さを表す指標の分布傾向から、限られた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。

さらに、抽出結果を用いた日・地点のクラスター分析によって、類似の渋滞特性をもつ日・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。

本分析では、コアテンソルの値である寄与度に着目して、圧縮次元数を定量的に決定したが、この方法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と言える。解釈性とモデル性能の両方を考慮した次元数決定方法が課題である。

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、本研究はJSPS 科研費 (20J21744, 21H01448) の助成を受けた。ここに記し、感謝を表します。

参考文献

- 1) 例えば, Dauwels, J., Aslam, A., Asif, M. T., Zhao, X., Vie, N. M., Cichocki, A., and Jaillet, P.: Predicting traffic speed in urban transportation subnetworks for multiple horizons, In: *Proceedings of the 2014 13th International Conference on Control Automation Robotics & Vision (ICARCV), Singapore*, pp. 547–552, 2014.
- 2) 例えば, Rempe, F., Huber, G., and Bogenberger, K.: Spatio-temporal congestion patterns in urban traffic networks, *Transportation Research Procedia*, Vol.15, pp.513–524, 2016.
- 3) 例えば, Pang, L. X., Chawla, S., Liu, W., and Zheng, Y.: On mining anomalous patterns in road traffic streams, *International Conference on Advanced Data Mining and Applications*, Vol.22, pp.237–251, 2011.
- 4) Zhou, H., and Hirasawa, K.: Spatiotemporal traffic network analysis: technology and applications, *Knowledge and Information Systems*, Vol. 60, pp. 25–61, 2019.
- 5) 酒井高良, 赤松隆: 首都圏高速道路網における渋滞パターンと Macroscopic Fundamental Diagram の安定性, 土木学会論文集 D3 (土木計画学), Vol.75, No.2, pp. 97-108, 2019.
- 6) Sun, L., and Axhausen, K.W.: Understanding urban mobility patterns with a probabilistic tensor factorization

- framework, *Transportation Research Part B: Methodological*, Vol. 91, pp. 511–524, 2016.
- 7) Yang, S., Wu, J., Xu, Y., 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*, Vol. 526, 2019.
- 8) Tucker, L. R.: Some mathematical notes on three-mode factor analysis. *Psychometrika*, Vol. 31, No. 3, pp. 279–311, 1966.
- 9) Harshman, R. A.: *Foundations of the PARAFAC procedure: Models and conditions for an “explanatory” multi-modal factor analysis*. Technical report, UCLA Working Papers in Phonetics, 1970.
- 10) Shashua, A., and Hazan, T.: Non-negative tensor factorization with applications to statistics and computer vision, In: *Proceedings of the 22nd International Conference on Machine Learning*, pp.792–799, 2005.
- 11) Allen, G.: Sparse higher-order principal components analysis, In: *Proceedings of the 15th International Conference on Artificial Intelligence and Statistics*, pp.27–36, 2012.
- 12) Fan, Z., Song, X., and Shibasaki, R.: Cityspectrum: a non-negative tensor factorization approach. In: *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing, ACM*, pp. 213–223, 2014.
- 13) 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*, Vol.14, pp.36–49, 2016.
- 14) 桑野将司, 木下礼央, 福山敬, 谷本圭志, 菅原一孔, 川村 尚生: 非負値テンソル因子分解を用いた検索ログデータからのバス利用特性の分析, 土木学会論文集 D3 (土木計画学), Vol. 73, No. 5, pp. I_569-I_578, 2017.
- 15) Kim, Y. D., and Choi, S.: Non-negative tucker decomposition. In: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 8 pages, 2007.10.1109/CVPR.2007.383405
- 16) McQueen, J.: Some methods for classification and analysis of multivariate observations, In: *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, pp.281–297, 1967.
- 17) Han, Y., and Moutarde, F.: Statistical traffic state analysis in large-scale transportation networks using locality-preserving non-negative matrix factorization, *IET Intelligent Transport Systems*, Vol. 7, No.3, pp. 283–295, 2013.

(2022.3.6 受付)