

テンソル分解を用いた首都圏高速道路網 における交通渋滞の時空間特性の解析

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 6-6-06-407) 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 6-6-06-407) 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 6-6-06-408) 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 6-6-06-408) 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。テンソル分解を用いた分析では、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。この方法は、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おり、これまで、交通状態把握や予測、異常検出などを目的とした、テンソル分解によって交通状態の特徴を抽出する試みがなされている。しかしながら、抽出された時間的特徴と空間的特徴は個別に解釈されており、その間の関係性について詳細に解釈している例や、渋滞特性の把握を目的とした分析例は見られない。そこで本研究では、首都圏高速道路網における時空間的な渋滞パターンを解明することを目的とし、テンソル分解を用いて渋滞の日・時刻・地点のパターンを抽出した。その結果、交通渋滞の周期性や規則性、顕著な時空間パターンの存在を明らかにした。

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、その緩和を目指して、交通状態の特徴を捉えた予測や、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ている。近年では、大量の交通観測データが取得可能になり、実証研究が盛んに行われている。特に、深層学習手法を用いた交通データ分析は、時系列データの学習に特化したモデルの精度が高いことで注目されている。しかし、深層学習モデルは解釈性が低いことが知られており、渋滞特性の解明には適さない。

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、渋滞頻発地点が存在す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。

酒井・赤松(2019)¹⁾は、首都圏高速道路網において、渋滞空間分布と渋滞推移過程に着目し、クラスター分析手法である k-means 法に基づいて渋滞特性分析を行い、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。

ここで、複雑な多項関係を有するデータから特徴抽出を行う分析手法に注目する。その一つであるテンソル分解は、行列分解を多次元配列(テンソル)に拡張した手法で、多要素によって構成される複雑なデータの特徴を抽出することができる。交通観測データが時間的特徴・空

間の特徴といっ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もテンソル分解の有用性が期待できる。

交通データにテンソル分解を適用した例として、Lin et al. (2018)²⁾ と Yang et al. (2019)³⁾ を挙げる。Lin et al. (2018)²⁾ は、テンソル分解をニューヨーク市のタクシーの OD データに適用し、抽出結果を通常時の交通状態パターンとした、異常検知方法を提案した。Yang et al. (2019)³⁾ は、テンソル分解を用いた交通状態のパターン抽出を提案した。大規模ネットワークを同様の渋滞レベルをもつサブネットワークに分類することを目的に、速度と密度のそれぞれを日・時間・道路区間方向に分解し、得られた低次元潜在空間を利用して分類した。その結果、各サブネットワークの巨視的な交通状態はそれぞれ異なる交通状態特性を持つことが確認された。

この他にも、欠損データ補完や交通状態予測などのタスクにおいてもテンソル分解の適用が提案され、その有用性が確認されている⁴⁾。しかし、時間や空間のパターンを個別に解釈・利用する分析に留まり、交通状態の時空間パターンについては十分に議論されていない。

以上を踏まえ、本研究では、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て、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、首都高速道路網の観測データから生成した渋滞時空間分布を表すデータに対して、日・時刻・地点に着目してテンソル分解を適用し、時空間的な渋滞パターンを明らかにすることを旨とする。

本論文の構成は以下の通りである。2. で非負値タッカー分解を紹介し、3. で分析対象とテンソルデータの生成方法を示す。4. はテンソル分解を適用した結果を示し、抽出された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. 非負値タッカー分解

テンソル分解は、行列分解を多次元配列に拡張した手法で、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。3階以上のテンソルに対して適用可能で、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、多次元配列であるテンソルを、各モードの特徴を表す因子行列と低ランク配列であるコアテンソルに分解する。

3階テンソル $\mathbf{X} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$ を例に、タッカー分解の概念

図を図-1に示す。タッカー分解モデル \mathbf{X}' は、

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

で表され、 $\mathbf{A} \in \mathbb{R}^{R_1 \times R_2 \times R_3}$ はコアテンソル、 $\mathbf{U}^{(i)} \in \mathbb{R}^{I_i \times R_i}$ ($i = 1, 2, 3$) 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(2)のように因子行列をベクトルで表記すると、タッカー分解は式(3)で書ける。

$$\mathbf{U}^{(i)} = [\mathbf{u}_1^{(i)} \mathbf{u}_2^{(i)} \dots \mathbf{u}_{R_i}^{(i)}] \quad (2)$$

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

式(1)のパラメータは \mathbf{A} , $\mathbf{U}^{(i)}$ で、二乗誤差を最小にするようにパラメータを決定する (式(4))。

$$\begin{aligned} & \min_{\mathbf{A}, \mathbf{U}^{(1)}, \mathbf{U}^{(2)}, \mathbf{U}^{(3)}} \|\mathbf{X} - \mathbf{X}'\|_F^2 \\ & = \|\mathbf{X} - \mathbf{A} \times \mathbf{U}^{(1)} \times \mathbf{U}^{(2)} \times \mathbf{U}^{(3)}\|_F^2 \end{aligned} \quad (4)$$

ここで、 $\|\cdot\|_F$ はフロベニウスノルムを表す。

得られた因子行列の列ベクトルは各モードにおけるパターンを表す。コアテンソルは、モード間の因子ベクトルの組み合わせの係数である。コアテンソル値が大きいほど、その因子ベクトル間の関係が強く、発生度合いが大きいことを表す。

CP 分解は、タッカー分解のコアテンソルが対角であるという制約が加わるため、得られた因子ベクトルは一対一に対応する。一方、タッカー分解は圧縮次元数の設定に関して自由度が高く、より様々なパターンを捉えることが可能であるため、モード間の複雑な関係をもつデータの分析に適している。

さらに、式(4)の最適化問題に非負制約やスパース制約など、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、複雑な関係性を持つ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。

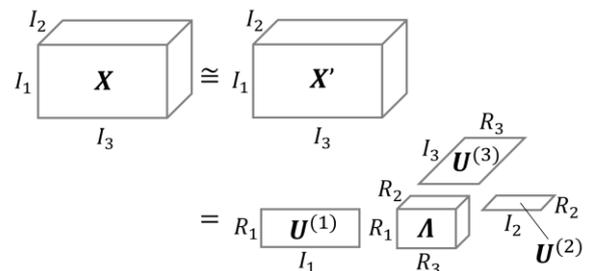


図-1 テンソル分解の概念図
(3階テンソルの場合)

以上より、本研究では、テンソル分解手法として、式(4)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 (Non-negative Tucker Decomposition: NTD) を用いる。

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の路線の、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・交通量データを分析する。

(2) 渋滞回数テンソルの生成

(1) で述べた交通観測データを用い、渋滞の時空間分布を表す渋滞回数テンソルを、a) 渋滞状態の判定、b) 渋滞回数の集計、c) 軸(モード)の設定の手順で生成する。

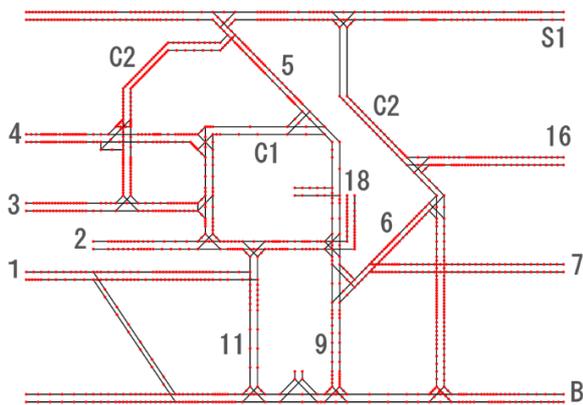


図-2 首都高速道路の模式図
(ノード(赤点)が感知器に対応)

表-1 対象路線一覧

路線記号	路線名	路線記号	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		

a) 渋滞状態の判定

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)¹⁾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(5)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。

$$\delta_{d,t,i} = \begin{cases} 1 & \text{if } (v_{d,t,i} < v^*) \\ 0 & \text{if } (v_{d,t,i} \geq v^*) \end{cases} \quad (5)$$

本分析でも同様に時間平均速度から渋滞を判定し、渋滞状態変数を生成した。

b) 渋滞回数の集計

a) の渋滞状態変数を用い、15分間隔の渋滞回数データを生成する。15分間の渋滞状態変数を集計し、渋滞回数を算出する際、幅120分のガウシアンカーネルを用い、時間方向に平滑化した。各要素の最小値は0、最大値は15である。なお、データが欠損している時間帯・感知器に関しては、渋滞回数を0回とした。ここで、生成されたデータは、「15分間のうち何分間で渋滞と判定されたか」を意味する“渋滞時間”データだが、本稿ではこれを「渋滞回数」と呼ぶ。

c) 軸(モード)の設定

最後に、b)の渋滞回数データに対して、日モード・時刻モード・感知器モードの3軸を設定し、渋滞回数テンソルを生成した(図-3)。テンソルのサイズは、365日×96時刻×1441感知器で、総要素数は50,492,640である。

4. 非負値タッカー分解の適用・結果

本稿では、コアテンソルの次元を(8, 8, 8)に設定した場合の分解結果を示す。非負値タッカー分解によって、日・時刻・感知器の各モードに関する因子行列(365×8, 96×8, 1441×8)とコアテンソル(8×8×8)を得た。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。

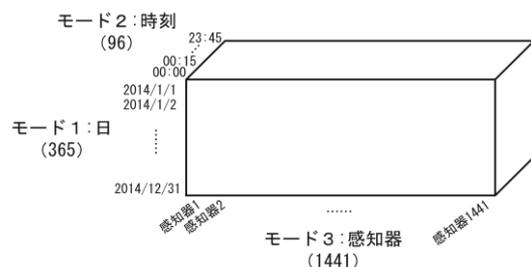


図-3 渋滞回数テンソル
(括弧内に各軸の長さを示す)

(1) 各モードにおける渋滞パターン

渋滞の時空間パターンの考察にあたり、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-4 に示す。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し、縦軸は週、横軸は曜日（月曜始まり）である。

図-4 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 2 因子は日曜日や祝日、GW・お盆・年末年始等の大型連休で色が濃いことから、「休日」を表していると言える。加えて、1 月や 4~6 月では曜日等に関わらず、色が濃く出ており、季節の特徴が現れている。

第 1 因子は、第 2 因子と色の濃さが反転しており、休日以外、すなわち「平日」を表す因子ベクトルである。また、第 2 因子と同様に季節の特徴が 3 月や 7~12 月に現れている。さらに、第 4 因子には「土曜日」のパターンが現れた。

以上のように、日モードでは、交通渋滞の季節性や曜日の周期性が抽出された。

b) 時刻モード

図-5 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もつパターンが得られている。

例えば、第 1 因子は、朝から夕方までの長い時間帯に分布し、一時点ではなく、交通量が多い日中に安定的に見られる渋滞を表す。

一方で、第 3 因子・第 4 因子は、それぞれ 6 時から 11 時、16 時から 20 時の比較的短い時間帯において鋭いピークを持っており、朝ピーク・夕ピークが抽出された。

c) 感知器モード（渋滞空間分布）

感知器モードの各因子ベクトルを図-6 に示す。道路網の模式図上で、閾値以上の値を持つ感知器に対応したリンクを赤で着色し、渋滞空間分布を描画した。

得られた 8 つの渋滞空間分布の全体の傾向として、JCT 付近を先頭に渋滞が発生している様子が見られる。一方で、第 3 因子の川口線 (S1) や、第 4 因子の中央環状線 (C1)、第 7 因子の 6 号三郷線 (16) は、特定の因子ベクトルでのみ渋滞パターンとして現れた。

(2) コアテンソル値を考慮した時空間パターン

コアテンソル値は、2. で述べた通り、3 モードの因子ベクトルの組み合わせに対する係数の役割を持ち、コアテンソル値が大きい各モードの因子の組み合わせは、元のテンソルデータ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。なお、非負値タッカー分解により出力された因子ベクトルのノルムを 1

に正規化し、該当のコアテンソル値に各因子ベクトルのノルムを乗じた。

$$X' = \sum_{r_1}^{R_1} \sum_{r_2}^{R_2} \sum_{r_3}^{R_3} \lambda_{r_1 r_2 r_3} (\mathbf{u}_{r_1}^{(1)} \circ \mathbf{u}_{r_2}^{(2)} \circ \mathbf{u}_{r_3}^{(3)}) \\ = \sum_{r_1}^{R_1} \sum_{r_2}^{R_2} \sum_{r_3}^{R_3} \lambda'_{r_1 r_2 r_3} (\mathbf{u}'_{r_1}{}^{(1)} \circ \mathbf{u}'_{r_2}{}^{(2)} \circ \mathbf{u}'_{r_3}{}^{(3)}) \quad (6)$$

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

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

以降、「コアテンソル値」は調整済みのコアテンソル値 $\lambda'_{r_1 r_2 r_3}$ を指す。

a) コアテンソル値の分布傾向

コアテンソル値の最大値は 243.6、最小値は 0.0、平均値は 3.8 であった。コアテンソル値の分布傾向を示す図-7 から、少数の限られたコアテンソル値が大きな値を持つ一方、その他の多数のコアテンソル値は 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数 $8 \times 8 \times 8 = 512$ のうち、少数の組み合わせが渋滞回数テンソルを表現する上で重要なパターンであると言える。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。

b) コアテンソル値が大きい顕著な時空間パターン

本項では、コアテンソル値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 20 のコアテンソル値を示す。日モードの因子ベクトルの番号に注目すると、上位 4 つは第 1 因子であり、次いで 4 つが第 2 因子である。第 5 因子、第 7 因子は、上位 20 番目までには入っていない。時刻・感知器モードに関しても同様に、因子ベクトルの番号に偏りが生じている。

特にコアテンソル値が大きい上位 3 位までの日・時刻・感知器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。

- ・コアテンソル値最大の組み合わせ (1, 2, 2)

「平日」「夜間」「各 JCT 付近」の渋滞を表す。

- ・コアテンソル値上位 2 位の組み合わせ (1, 3, 3)

「平日」「朝」「川口線、16 号三郷線・6 号向島線・9 号深川線の上り」の渋滞を表す。

- ・コアテンソル値上位 3 位の組み合わせ (1, 1, 1)

「平日」「日中」「都心環状線外回り江戸橋 JCT~竹橋 JCT、2 号目黒線・3 号渋谷線・5 号池袋線・6 号向島線・9 号深川線の上り」の渋滞を表す。

また、日・時刻・感知器のパターンの組み合わせに着目

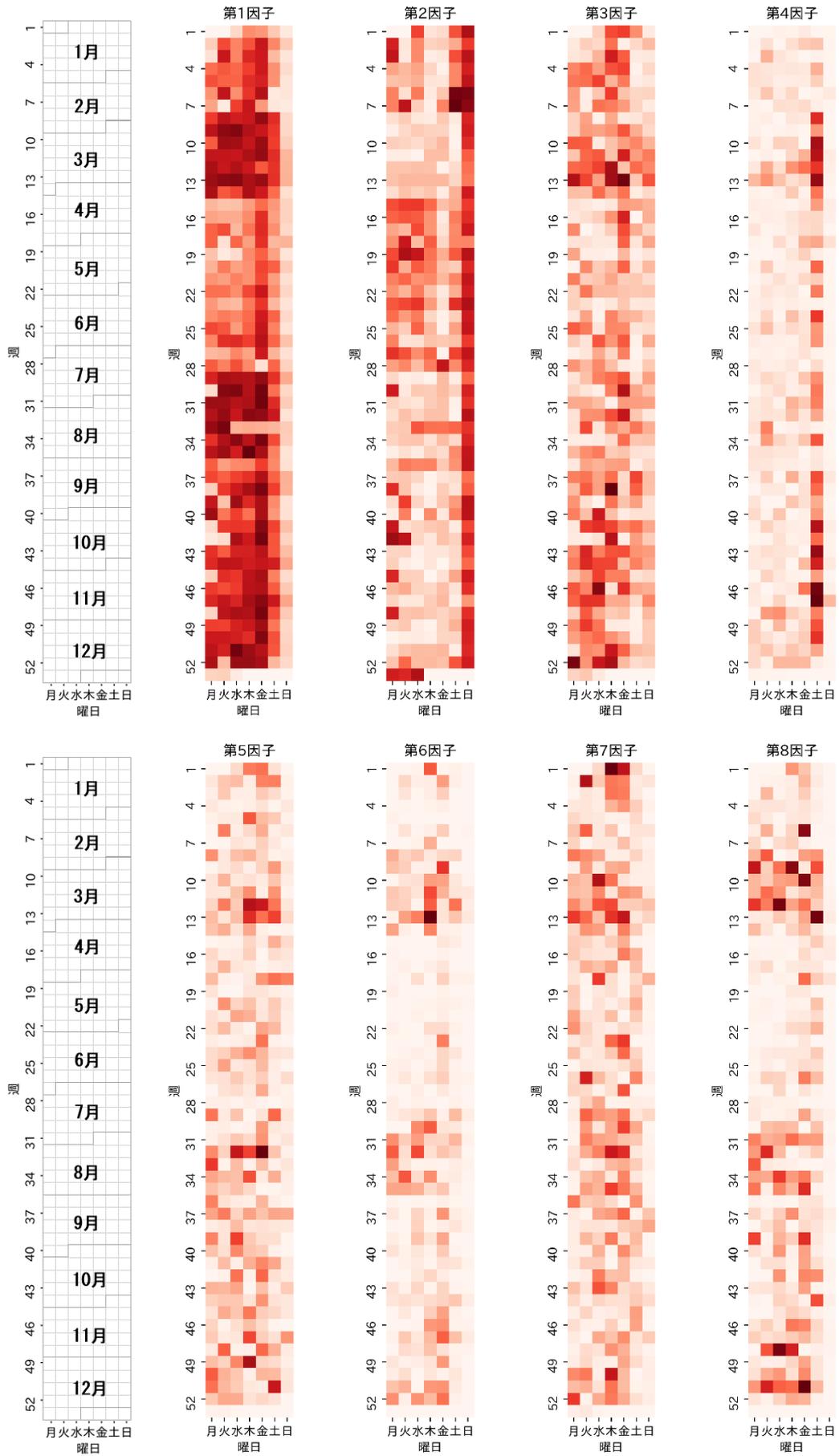


図-4 日モードの因子ベクトル

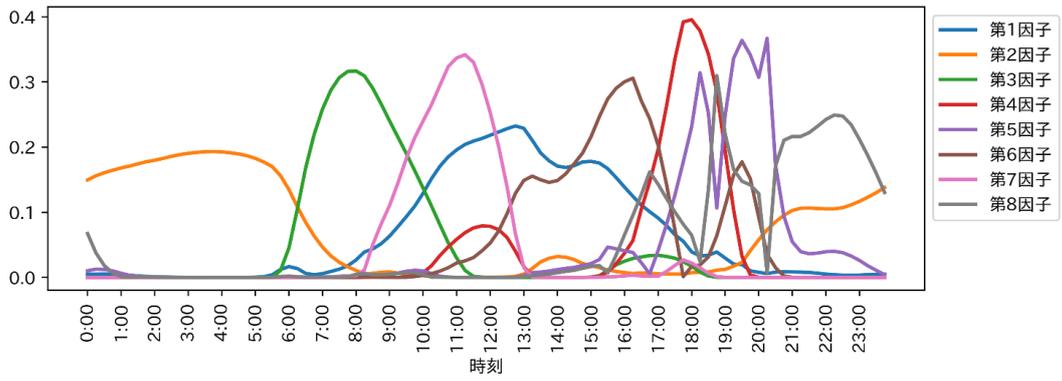


図-5 時刻モードの因子ベクトル

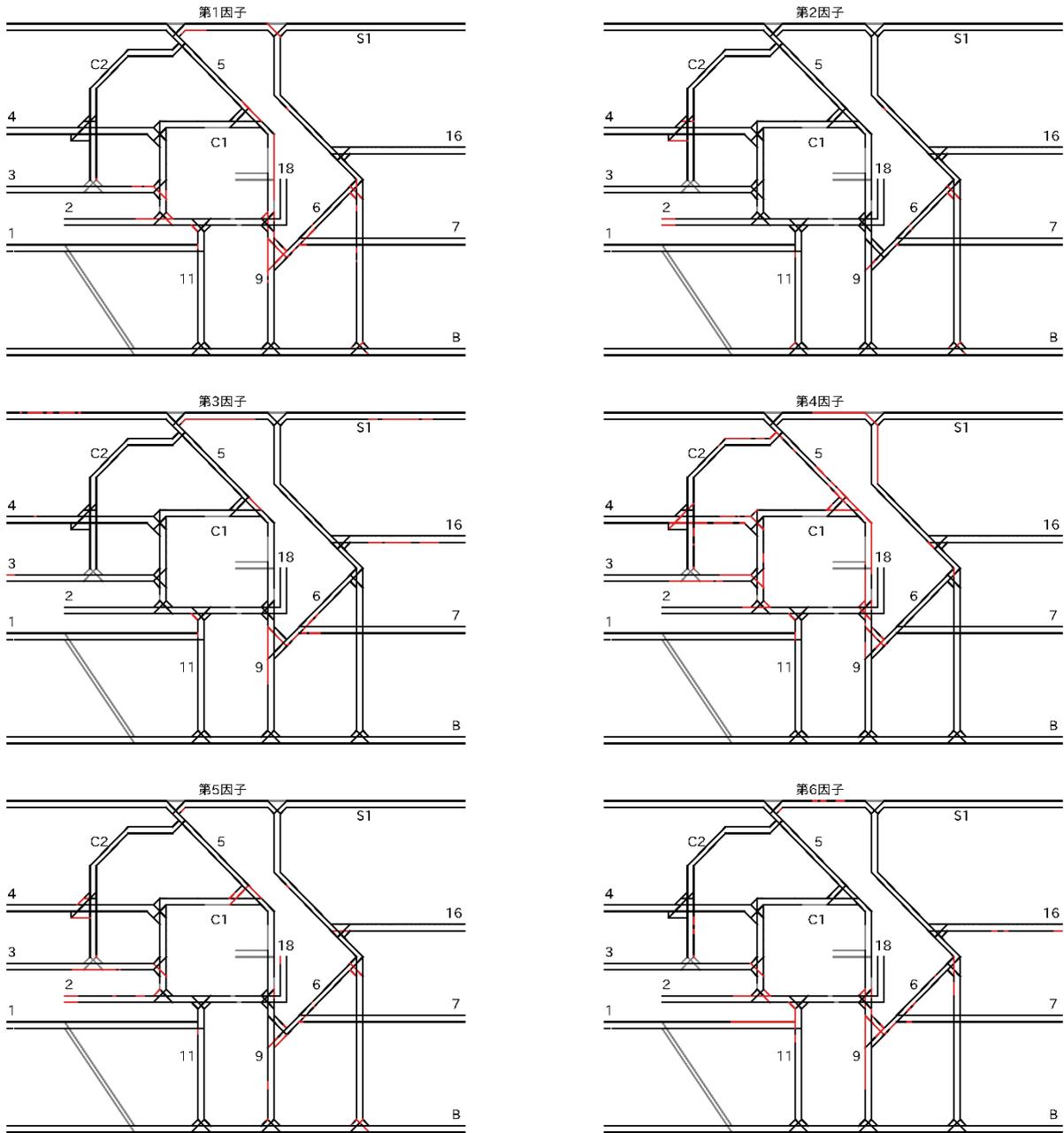


図-6 感知器モードの因子ベクトル
(着色の閾値は0.05)

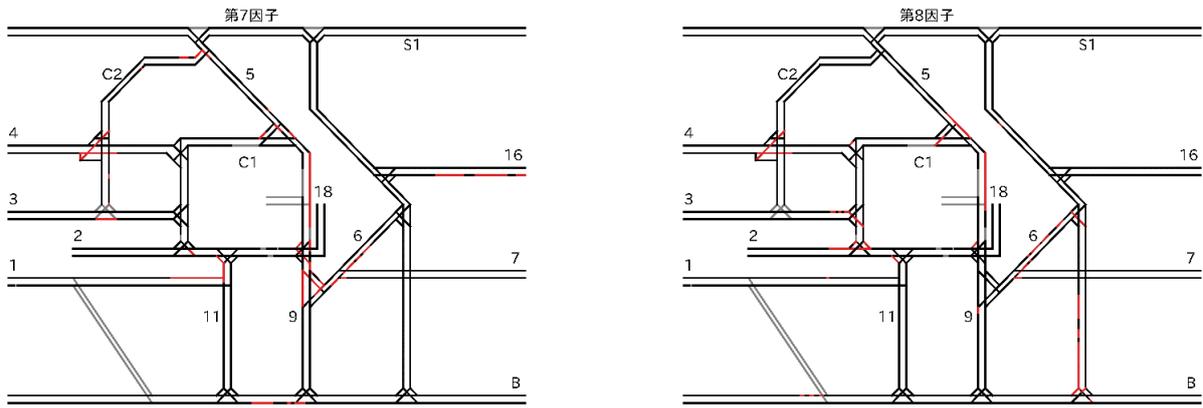


図-6 感知器モードの因子ベクトル (続き)
(着色の閾値は 0.05)

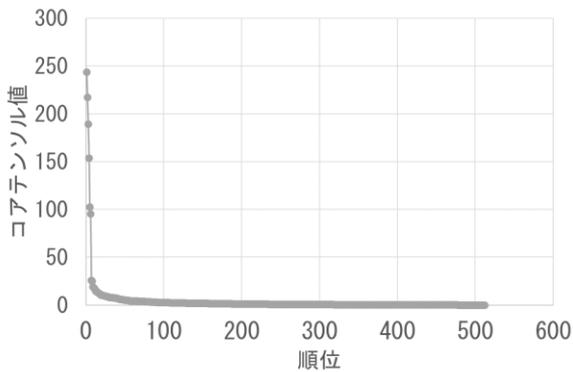


図-7 コアテンソル値の分布傾向

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因子のパターンが2件見られる(1位(1,2,2), 5位(2,2,2))。これは、平日、休日に関わらず、夜間から早朝に、JCT 付近で局所的な渋滞が発生していることを示している。

このように、コアテンソル値を考慮して、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。

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渋滞の時空間分布を表すデータを生成し、テンソル分解に基づいて渋滞時空間パターンを抽出した。

多項関係を有するデータのパターン抽出で有用性が確認されている非負値タッカー分解を、日、時刻、感知器の3要素軸から構成される約5千万要素のテンソルデータに適用し、512要素のコアテンソルとそれぞれ2,920, 768, 1,528要素の3つの因子行列に縮約した。分析結果から渋滞

表-2 コアテンソル値 (上位20)

順位	因子ベクトルの組み合わせ			コアテンソル値
	日	時刻	感知器	
1	1	2	2	243.6
2	1	3	3	217.1
3	1	1	1	189.0
4	1	4	4	153.7
5	2	2	2	102.7
6	2	1	2	95.4
7	2	2	5	25.8
8	2	3	2	24.9
9	3	3	3	19.1
10	2	2	1	18.2
11	1	6	6	17.4
12	8	3	3	15.7
13	1	3	2	14.2
14	2	3	3	14.2
15	1	7	4	13.3
16	4	3	3	12.7
17	6	1	4	12.2
18	6	1	3	11.7
19	3	2	2	10.6
20	2	8	2	10.3

の発生パターンとして解釈可能なパターンが抽出されたことを確認し、非負値タッカー分析が渋滞パターン分析に適用可能であることが明らかとなった。

分析の結果、モード毎に、季節・曜日等の周期性や、時間帯、道路区間によって異なる渋滞パターンが抽出された。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、モード間の関係の強さを表すコアテンソル値の分布傾向から、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。具体的には、平日、休日問わず、夜間から早朝にかけていくつかの JCT 付近で局所的に発生する渋滞パターンや、平日の朝に複数路線の上りで発生する渋滞パターンが確認された。結果は、交通特性の特徴と比較すると、妥当な結果であると考えられる。

本分析では、解釈性に着目し、圧縮次元数を試行錯誤で決定したが、分析上適切な設定である保証はない。圧縮次元数の決定方法に関しては、既往研究でも重要な課題として指摘されており、RMSE や AIC, BIC などのモデルの適合度指標を参考にする方法が提案されている。さらに、「データとモデルを最短に記述できるモデルが最良である」とする記述長最小原理 (MDL 原理) を非負行列分解、非負値 CP 分解へ適用した手法が提案されている^{5,6)}。今後、これらの定量的なモデル評価に基づいて次元数を決定し分析・解釈を行う必要があると考えている。また、交通渋滞の時空間特性の把握に向けて、日や、路線等の条件による固有パターンの検証や抽出された結果の利用についても今後の検討課題とする。

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、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) 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.
- 2) Lin, C., Zhu, Q., Guo, S., Jin, Z., Lin, Y, R., and Cao, N.: Anomaly detection in spatiotemporal data via regularized non-negative tensor analysis, *Data Mining and Knowledge Discovery*, Vol. 32, pp. 1056-1073, 2018.
- 3) 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.
- 4) 例えば, Balasubramaniam, T., Nayak, R., and Yuen, C.: Sparsity constraint nonnegative tensor factorization for mobility pattern mining, *PRICAI 2019: Trends in Artificial Intelligence*, pp. 582-594, 2019.
- 5) Ito, Y., Oeda, S., and Yamanishi, K.: Rank selection for non-negative matrix factorization with normalized maximum likelihood coding, *SIAM International Conference on Data Mining*, pp. 720-728, 2016.
- 6) Fu, Y., Matsushima, S., and Yamanishi, K.: Model selection for non-negative tensor factorization with

minimum description length, *Entropy*, Vol. 21, 632, 2019.

(2021.10.1 受付)