

自己回帰テンソル分解による時空間データ予測

Autoregressive Tensor Factorization for Spatio-temporal Predictions

竹内 孝^{*1*2} 鹿島 久嗣^{*2*3} 上田 修功^{*1*3}
 Koh Takeuchi Hisashi Kashima Naonori Ueda

^{*1}日本電信電話株式会社 コミュニケーション科学基礎研究所
 NTT Communication Science Laboratories

^{*2}京都大学大学院 情報学研究科
 Graduate School of Informatics, Kyoto University

^{*3}理化学研究所 革新知能統合研究センター
 RIKEN Center for Advanced Intelligence Project

Spatio-temporal data analysis is a common research topic that requires interpolations of unknown locations and predictions of future observations by utilizing information about where and when data were observed. Tensor factorization methods are popular in this field because of their capability of handling multiple types of spatio-temporal data, dealing with missing values, and providing computationally efficient parameter estimation procedures. However, any methods have been proposed to address the most difficult problems, predicting future observations of unknown locations. In this paper, we propose a new tensor factorization method that estimates low-rank latent factors by simultaneously learning the spatial and temporal autocorrelations. We introduce new spatial autoregressive regularizers based on existing spatial autoregressive models and provide an efficient estimation procedure.

1. はじめに

高機能センシングデバイスと高速移動通信網の普及に伴い、センサーネットワークを介した都市の広範囲リアルタイムモニタリングシステムが実現されつつある。システムを通して観測される人の行動やモノの移動、大気環境など多種多様な時空間データを解析し、都市のマネジメントや企画に役立てようとする試みが注目を集めており、データ解析技術は今後ますます重要になると考えられる。

時空間データの解析は、空間統計学、空間経済学、気候学、リモートセンシング [1, 2] などの分野で盛んに行われてきた。これらの研究は、センサー未設置場所の観測データの推定 (空間内挿) あるいはセンサー設置場所での将来観測値の予測 (時間外挿) を目的とする場合が多い。その中でもセンサー未設置場所での将来観測値の予測問題は、空間内挿と時間外挿を同時に行うため、最も難度の高い問題の1つとされている。この問題が解決されれば時空間的な需要予測が可能となるため、タクシー配車やライドシェアの効率化、バイクシェアサービスの車両再配置、宿泊施設の宿泊数予測などへの応用が期待される。

現在のモニタリングシステムは、数百から数千のセンシングデバイスから構成され、複数種類のデータを観測する場合が多い。例えばニューヨーク市のあるバイクシェアサービスでは、344個のバイクステーションから30分毎に返却あるいは貸出された自転車数が送信される。ただ、システム構築に大量のセンシングデバイスを必要とするため、安価だが安定性の低い機器が利用される場合が多く、デバイス故障、データ転送エラー、ストレージ故障などによるデータ欠損が発生しやすい [3]。そのため、データ欠損を抜いしつつ大規模多種類の時空間データを処理するための技術が必要となる。

多変量時系列データの予測問題では、多変量自己回帰 (VAR) やガウス過程 (GP) [4] などが広く利用されている。これらの手法は、変数間の線形あるいは非線形の相関情報を利用し高い予測精度を示すが、パラメータ推定の計算量がデータ次元やデータ点数に対して指数的に増加するため、大規模データの解

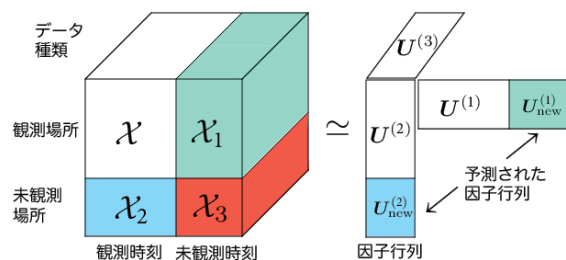


図 1: 時空間データ解析のためのテンソル分解モデル. \mathcal{X}_1 , \mathcal{X}_2 , \mathcal{X}_3 の推定は、センサー設置場所のデータ予測、センサー未設置場所のデータ補完と予測の問題に対応する。

析には適さない [5].

大規模な時空間データの解析を行うための手法として、行列分解やその一般化であるテンソル分解が提案されている [5-7]. テンソル分解は、時空間データを高次元配列のテンソル \mathcal{X} として扱い、テンソルを少数の因子に分解することでデータに含まれる重要な情報を抽出する手法である。テンソル分解の計算量はデータの観測時刻数、観測場所数、データ種類数に対して線形にスケールし、さらに欠損データが存在しても因子推定が可能のため、欠損値の復元に広く利用されてきた。図 1 に、時空間データを観測時刻、観測場所、データ種類数の 3 つのモードからなる 3 階テンソルとして扱い分解する例を示す。例えばタクシーの降車履歴データから新たなサービス提供場所での需要を予測する問題に取り組むとする。観測済みのデータを \mathcal{X} (白色ブロック) とすると、サービス提供場所での将来需要予測、サービス未提供場所の過去需要補完、将来需要予測の問題は、因子行列を用いて 3 つの異なるブロック \mathcal{X}_1 (緑), \mathcal{X}_2 (青), \mathcal{X}_3 (赤) を推定する問題として扱える。

テンソル分解に時空間的な相関情報を取り入れるために、グラフ正則化技術を使用したテンソル分解が提案されている [7]. グラフ正則化は、因子内のパラメータが空間的に隣接する場所であったり前後する時刻に対応する場合、両者が類似した推定

値をもつように促すため、時空間的に滑らかな因子の抽出を可能にする。このようなアイデアは空間経済学分野などで利用されているが、どのようなデータにも適用可能なグラフ構築方法は存在しないとされている [8]。また、既存のグラフ正則化では、グラフラプラシアンを使用するため、パラメータ間の正と負の相関を区別して使用できない。

これらの問題を解決するために、時間方向のグラフ正則化に自己回帰モデルの回帰係数を用いる Temporal regularized matrix factorization (TRMF) [6] が提案されている。TRMF は因子の推定と因子毎のパラメータ間の自己相関の学習を交互に行うことで、事前にグラフを構築する問題を回避した。さらに学習された自己回帰モデルを用いて将来の因子行列 $\mathbf{U}_{\text{new}}^{(1)}$ の予測を行うことで、観測値の予測において既存法よりも大幅な性能改善を示した。しかし、TRMF は空間に対応する因子の学習ではグラフ正則化を用いるのみで、空間的な自己相関を学習する問題は扱っていない。

本研究では、センサーの未設置場所でのデータ補完と予測の問題を解決するために、空間上と時間方向のそれぞれの因子において自己回帰モデルの学習を行いつつ、因子の推定を行う新たなテンソル分解を提案する。まず、コンピュータビジョンの分野で提案された 2 次元自己回帰モデル [9] を拡張し、空間情報を用いた有向グラフ自己回帰正則化を提案し、これを TRMF を組み合わせることで自己回帰テンソル分解を提案する。さらに時空間データは観測値が非負値となる場合があるため、因子パラメータを非負に制約する非負自己回帰テンソル分解の提案も行う。次に、提案したテンソル分解が因子の推定を行うための効率的なアルゴリズムの導出を行う。実世界で観測されたタクシー履歴データおよびバイクシェアリングのデータを用いた実験を行い提案手法の優位性を示す。

2. テンソル分解

T 個の観測時刻、 P 個の観測場所、 M 種類のデータからなる時空間データの 3 階テンソルを $\mathcal{X} \in \mathbb{R}^{T \times P \times M}$ とし、データ内の観測要素の集合を $\Omega = \{(t, p, m) \mid x_{t,p,m} \text{ is observed.}\}$ とする。 K ($\ll \min(T, P, M)$) 個の因子ベクトルからなる因子行列を $\mathbf{U}^{(1)} \in \mathbb{R}^{P \times K}$, $\mathbf{U}^{(2)} \in \mathbb{R}^{T \times K}$, $\mathbf{U}^{(3)} \in \mathbb{R}^{M \times K}$ とする。本研究では、テンソル分解モデルのなかで最もシンプルな CP 分解を用いる。正則化テンソル分解の経験損失関数 $\mathcal{L}(\mathbf{U}^{(1)}, \mathbf{U}^{(2)}, \mathbf{U}^{(3)})$ は、以下のように定義される。

$$\sum_{(t,p,m) \in \Omega} \left(x_{p,t,r} - \sum_{k=1}^K u_{t,k}^{(1)} u_{p,k}^{(2)} u_{m,k}^{(3)} \right)^2 + \sum_{n=1}^3 \lambda_n g_n(\mathbf{U}^{(n)}). \quad (1)$$

なお、 g_n , λ_n は正則化関数、正則化重みとした。(1) の第 1 項は観測値 $x_{t,p,m}$ を因子ベクトルの線形和 $\hat{x}_{t,p,m} = \sum_{k=1}^K u_{t,k}^{(1)} u_{p,k}^{(2)} u_{r,k}^{(3)}$ で近似した際の二乗誤差、第 2 項は正則化関数である。テンソル分解は、式 (1) の最小化問題を解き、局所最小解 $\{\mathbf{U}^{(1)*}, \mathbf{U}^{(2)*}, \mathbf{U}^{(3)*}\}$ を求める。

テンソル分解では様々な正則化関数が利用されているが、非負領域への指示関数 $P_{\geq 0}(\mathbf{U}^{(n)}) = 0$ (if $u_{i,k}^{(n)} \geq 0, \forall (i, k)$), or $+\infty$ (otherwise) を用いるものは、非負テンソル分解 (Non-negative Tensor Factorization: NTF) として知られている。NTF は因子ベクトルを非負に制約することで、非負観測値のみからなるテンソルの欠損値補完において高い精度を示し、また解釈の容易な因子を抽出するとされている [10, 11]。

グラフラプラシアン正則化は、モードの特徴量をノード、特徴量間の類似度を重み付きエッジとするグラフを用いて、因

子ベクトルがグラフ構造に対して滑らかとなるよう制約する技術である。 n 番目のモードに対応する重み付き隣接行列を $\mathbf{W}^{(n)} \in \mathbb{R}^{I_n \times I_n}$ とすると、ラプラシアン行列 $\mathbf{L}^{(n)}$ は次のように定められる。 $\mathbf{L}^{(n)} = \mathbf{D}^{(n)} - \mathbf{W}^{(n)}$, この時、 $\mathbf{D}^{(n)}$ は $\mathbf{W}^{(n)}$ の行毎の要素和を対角成分にもつ対角行列とする。するとグラフラプラシアン正則化の正則化関数は $g_n(\mathbf{U}^{(n)}) = \sum_{k=1}^K \mathbf{u}_k^{(n)\top} \mathbf{L}^{(n)} \mathbf{u}_k^{(n)}$ となる。グラフラプラシアン正則化は、特徴量の時間や空間の構造を表現したグラフを利用することで時空間データの欠損値補完に高い精度を示すが、重みを事前に設定する必要がある。

2.1 自己回帰モデルと正則化

VAR は、多変量時系列の相関をもちいて将来の観測値を予測する手法である。時刻 t に対応する因子行列の行ベクトルを $\mathbf{u}_t^{(1)} \in \mathbb{R}^K$, ラグセットを L , 回帰係数を $A = \{\mathbf{A}_l \in \mathbb{R}^{K \times K} \mid l \in L\}$ とする。VAR は $\mathbf{u}_t^{(1)}$ を過去の列ベクトルの重み付き線形和と多変量ガウス分布からサンプリングしたノイズベクトル ϵ_t を用いて、次のように近似する。 $\mathbf{u}_t^{(1)} = \sum_{l \in L} \mathbf{A}_l \mathbf{u}_{t-l}^{(1)} + \epsilon_t$. TRMF [6] は、VAR の回帰係数を時刻モードのグラフラプラシアンの重みに用いた正則化関数を用いることで、グラフの重みの事前設計を回避し、さらに将来の因子ベクトルの予測を行う技術である。

$$\frac{1}{2} \sum_{t=\bar{L}}^T \left\| \mathbf{u}_t^{(1)} - \sum_{l \in L} \mathbf{A}_l \mathbf{u}_{t-l}^{(1)} \right\|_2^2 + \frac{\eta}{2} \sum_{t=1}^T \left\| \mathbf{u}_t^{(1)} \right\|_2^2, \quad (2)$$

なお、 $\bar{L} = 1 + \max(L)$ とし、 η は重みパラメータとした。

2 次元自己回帰モデル (2D-AR) [9] は、コンピュータビジョンで提案された画像の欠損値補完のための技術である。ある画像に対応するベクトルを $\mathbf{y} \in \mathbb{R}^P$ とし、ベクトルの要素 y_p は (i, j) 番目のピクセルの観測値とする。 L 階の空間ラグの集合を $E = \{(r, s) \mid 1 \leq r < L \cup 1 \leq s < L\}$, 回帰係数を $B = \{b_{r,s} \mid (r, s) \in E\}$ と定める。2D-AR は \mathbf{y} の観測値を近傍の観測値の線形和と多変量ガウス分布からサンプリングしたノイズベクトルを用いて $y_{i,j} = \sum_{(r,s) \in E} b_{r,s} y_{i+r,j+s} + \epsilon$ とし定められる。画像欠損値の補完において 2D-AR は優れた性能を示すことが知られているが、このモデルは観測場所が格子状に並ぶことを前提としており、非格子状のデータ解析には適用できない。

3. 自己回帰テンソル分解

テンソル分解の空間モードと時間モードの因子ベクトルに自己回帰正則化を導入した自己回帰テンソル分解を提案する。提案手法の損失関数 $\mathcal{L}(\mathbf{U}^{(1)}, \mathbf{U}^{(2)}, \mathbf{U}^{(3)}, A, B)$ を次のように定める。

$$\begin{aligned} & \|P_{\Omega}(\mathcal{X} - \mathcal{C} \times_1 \mathbf{U}^{(1)} \times_2 \mathbf{U}^{(2)} \times_3 \mathbf{U}^{(3)})\|_2^2 \\ & + \lambda_1 g_1(\mathbf{U}^{(1)}, A) + \lambda_2 g_2(\mathbf{U}^{(2)}, B), \end{aligned} \quad (3)$$

このとき、 $g_1, g_2, P_{\Omega}(\cdot)$ は、それぞれ、自己回帰正則化関数 [6]、次節で提案する空間自己回帰正則化関数、観測値と推定値の誤差関数である。 A および B は、自己回帰正則化関数および空間自己回帰正則化関数の回帰係数である。 \mathcal{X} の観測値が非負の場合には、非負制約関数 $\sum_{n=1}^3 P_{\geq 0}(\mathbf{U}^{(n)})$ を加え、非負自己回帰テンソル分解とする。自己回帰テンソル分解は式 (3.2) を最小化する $\arg \min_{\mathbf{U}^{(1)}, \mathbf{U}^{(2)}, \mathbf{U}^{(3)}, A, B} \mathcal{L}$ を求めるものとする。

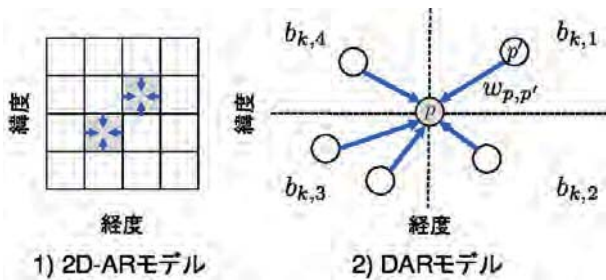


図 2: 2次元自己回帰モデルと空間自己回帰正則化モデル

3.1 空間自己回帰正則化

時空間データの観測場所は、行政区分や都市構造の影響を受けるため、格子状になるとは限らない。そこで観測場所が格子状・非格子状にいずれの場合にも対応可能な自己回帰モデルとして、有向自己回帰正則化 (DAR: directed autoregressive regularizer) を提案する。DAR の仮定は非常にシンプルなものである、つまり、回帰係数は観測場所間の方向によって定まり、ある観測値は周囲の観測値を回帰係数と距離重みで線形和によって近似できるとする。DAR と 2DAR を図 2 に示す。DAR は 2π の角度を N 個の区分に分割し、区分ごとに回帰係数を定める。

DAR の回帰係数を $B = \{b_{k,n} \mid \forall(k,n)\}$, p 番目の観測場所とリンクを持つ観測場所の集合を E_p , p 番目と p' 番目の観測場所の正規化距離重みを $w_{p,p'} \in \mathbb{R}_{\geq 0}$ where $\sum_{p' \in E_p} w_{p,p'} = 1$ とする。DAR 正則化関数 $g^{\text{DAR}}(\mathbf{U}^{(2)}, B)$ を次のように定める。

$$\sum_{k=1}^K \sum_{p=1}^P \left(u_{p,k}^{(2)} - \sum_{p' \in E_p} b_{k,n_{p'}} w_{p,p'} u_{p',k}^{(2)} \right)^2 + \frac{\eta}{2} \|\mathbf{u}_k^{(2)}\|_2^2, \quad (4)$$

このとき、 $n_{p'}$ は p 番目の観測場所から p' 番目の観測場所への角度区分とする。上式は、行列 $\bar{\mathbf{B}}_k \in \mathbb{R}^{P \times P}$

$$\bar{b}_{k,p,p'} = \begin{cases} b_{k,n_{p'}} w_{p,p'} & (\text{if } p' \in E_p), \\ 0 & (\text{otherwise}) \end{cases}, \quad (5)$$

を用いると、次のように式 (2) と同様に变形できる。

$$\sum_{k=1}^K \left\| \mathbf{u}_k^{(2)} - \bar{\mathbf{B}}_k \mathbf{u}_k^{(2)} \right\|_2^2 + \frac{\eta}{2} \|\mathbf{u}_k^{(2)}\|_2^2.$$

3.2 パラメータ推定

式 (3.2) の最小化問題は、次と等価である [12].

$$\min_{\mathbf{U}^{(1)}, \mathbf{U}^{(2)}, \mathbf{U}^{(3)}, A, B, \mathbf{Z}} \|\mathcal{Z} - \mathcal{C} \times_1 \mathbf{U}^{(1)} \times_2 \mathbf{U}^{(2)} \times_3 \mathbf{U}^{(3)}\|_2^2 + \lambda_1 g_1(\mathbf{U}^{(1)}, A) + \lambda_2 g_2(\mathbf{U}^{(2)}, B) \text{ s.t. } P_{\Omega}(\mathcal{X} - \mathcal{Z}). \quad (6)$$

上式の損失関数は非凸関数であるが、パラメータ $\{\mathbf{U}^{(1)}, \mathbf{U}^{(2)}, \mathbf{U}^{(3)}, A, B, \mathbf{Z}\}$ に関しては凸関数 (multi-convex) となっている。紙面の都合上、詳細は割愛するが、それぞれパラメータについての損失関数は 2 次関数となるため、各パラメータについて最小二乗法を行い交互にパラメータ更新することで局所最小解が得られる [13].

4. 実験

ニューヨーク市で観測されたタクシー降車履歴データとバイクシェアリング返却履歴データ*1 を用いた実験を行う。タク

*1 http://www.nyc.gov/html/tlc/html/about/trip_record_data.shtml, <http://www.citibikenyc.com/>

表 1: タクシーデータの予測誤差

手法	未設置場所予測	未設置場所補完	設置場所予測
DAR	0.569±0.079	0.509±0.002	0.317±0.107
DAR+NN	0.568±0.091	0.499±0.001	0.312±0.108
TRMF	0.996±0.041	0.997±0.001	0.350±0.128
FAST	1.095±0.010	1.056±0.002	0.653±0.005
GP	1.256±0.056	1.095±0.060	1.683±0.188

表 2: バイクシェアリングデータの予測誤差

手法	未設置場所予測	未設置場所補完	設置場所予測
DAR	0.843±0.116	0.677±0.021	0.760±0.127
DAR+NN	0.821±0.089	0.678±0.018	0.818±0.295
TRMF	1.065±0.210	0.845±0.032	0.744±0.101
FAST	0.910±0.055	0.818±0.015	0.710±0.058
GP	3.589±5.801	0.743±0.065	3.701±6.139

シー降車履歴データでは、ニューヨーク市内を 500m² からなる 640 個のメッシュに分割し、30 分毎のメッシュ内での降車数を予測対象とした。バイクシェアリングデータでは、344 個のバイクステーションで 30 分毎に返却された自転車の数を予測対象とした。データ毎に観測場所の 10% を乱択してセンサー未設置とし、観測期間の最終週を取り出し各日を予測対象のデータとした。パラメータの推定は、センサー設置場所で観測された予測対象日から 1 週間前までをトレーニングデータを用いて行った。提案手法である自己回帰テンソル分解 (DAR) と非負自己回帰テンソル分解 (DAR+NN) の性能を、TRMF [6]、高速低ランク回帰モデル (FAST) [5]、ガウス過程 (GP) [4] と比較した。

空間自己回帰正則化関数では、観測場所の k 近傍に対し、距離重みを $w_{p,p'} = \exp(-\|\mathbf{r}_p - \mathbf{r}_{p'}\|_2^2 / \sigma^2)$ [14] とし正規化を行った。 \mathbf{r}_p は p 番目の観測場所の GPS 情報、 σ はスケールパラメータで、今回は $\sigma = 100$ に設定した。FAST と TRMF の空間モードのグラフラプリアンには、正規化していない距離重みを使用した。提案手法のハイパーパラメータである角度区分数 N と k 近傍数は $\{4, 8\}$ 、正則化重みは λ_1, λ_2 は $\{0, 0.1, 1.0, 10.0\}$ から、5 交差検定によって選択した。GP では乱択した 5,000 点をトレーニングデータとし、最尤推定によってハイパーパラメータを最適化した。図.1 で示した各問題における精度を相対絶対値誤差 $(1/|\Omega|) \sum_{(p,t,m) \in \Omega} |x_{p,t,m} - \hat{x}_{p,t,m}| / \sum_{(p,t,m) \in \Omega} |x_{p,t,m}|$ によって評価した。

表 1 と 2, 7 日間の試行から得られた相対絶対値誤差の平均

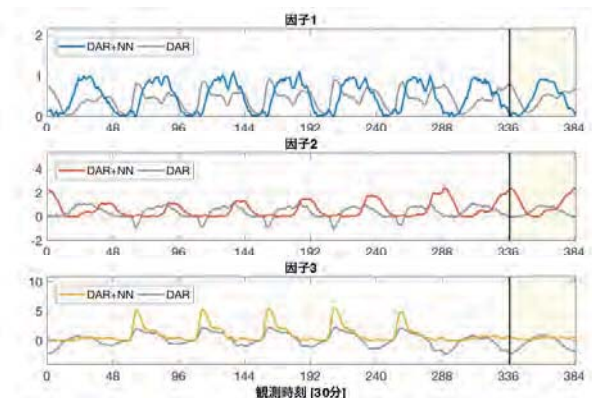


図 3: 時間モードの因子ベクトル (黄色背景はテスト期間)

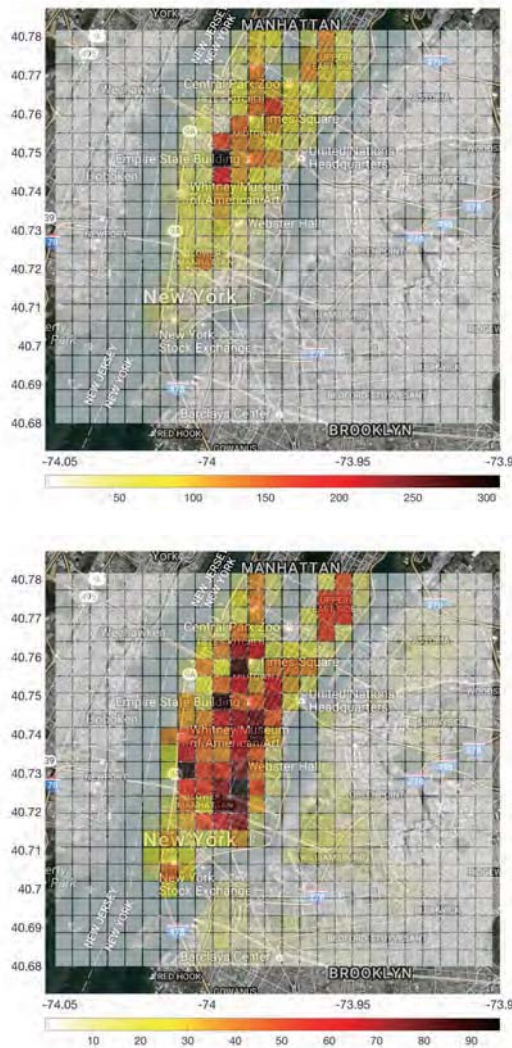


図 4: DAR+NN の空間モードの因子ベクトル (因子 1,2)

と標準偏差を示す。最も高い性能を示した手法の数値を太字とする。表より、提案手法 (DAR, DAR+NN) は観測場所が格子状あるいは非格子状であるかに関わらず、センサー未設置場所のデータ補完および予測において、既存法よりも高い性能を示した。特に予測問題において提案手法は、それぞれの実験データで 42.9%, 9.8% と大幅に予測誤差の改善を達成した。バイクシェアリングデータのセンサー設置場所での予測では FAST が高い性能を示しており、データが完全な場合での低ランク回帰モデルの有効性が示唆された。

提案法の定性的な評価を行うために、タクシーデータから推定された因子の可視化を行う。図 3 に時間モードの 3 つの因子ベクトルを非負制約ありは色線、非負制約なしは灰色で示す。さらに対応する空間モードの非負制約ありの因子ベクトルのヒートマップを図 4 に示す。図 3 より、因子ベクトル毎に異なるダイナミクスが抽出されていることがわかる。特に非負制約ありの場合は、それぞれのピークから日中、夜間、平日の午前中に出現するパターンであるとわかる。さらに図 4 より、因子 1 ではエンパイアステートビル周辺で因子が高い値をもつため、日中に都心部へ移動する降車パターンが抽出されたことが確認できた。因子 2 では異なる降車パターンが抽出されており、マンハッタン内およびブルックリンエリアで高い値が現れている。これは夜間での移動と帰宅時に現れる降車パター

ンが抽出された。以上から、提案手法によってニューヨーク市内でのタクシー降車の特徴的なパターンの抽出が可能になり、これらが予測精度の改善に貢献したものと考えられる。

5. おわりに

本研究では、センサーの未設置場所でのデータ補完と予測の問題を解決するために、自己回帰テンソル分解およびパラメータ推定法を提案した。タクシー降車履歴データおよびバイクシェアリングのデータを用いた実験から、提案手法による補完および予測問題での精度向上を確認した。さらに因子ベクトルの可視化を行い、提案手法の定性的な有効性を定性的に確認した。今後は非定常かつ非線形な自己相関を学習するテンソル分解、低ランク回帰モデルへの拡張などを検討していく。

参考文献

- [1] L. Anselin, "Thirty years of spatial econometrics," *Papers in regional science*, vol. 89, no. 1, pp. 3–25, 2010.
- [2] X. Zhou, S. Shekhar, and R. Y. Ali, "Spatiotemporal change footprint pattern discovery: an inter-disciplinary survey," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 4, no. 1, pp. 1–23, 2014.
- [3] L. Li, J. McCann, N. S. Pollard, and C. Faloutsos, "Dynammo: Mining and summarization of coevolving sequences with missing values," in *Proc. of SIGKDD*, 2009, pp. 507–516.
- [4] C. E. Rasmussen, "Gaussian processes for machine learning." MIT Press, 2006.
- [5] M. T. Bahadori, Q. R. Yu, and Y. Liu, "Fast multivariate spatio-temporal analysis via low rank tensor learning," in *Proc. of NIPS*, 2014, pp. 3491–3499.
- [6] H.-F. Yu, N. Rao, and I. S. Dhillon, "Temporal regularized matrix factorization for high-dimensional time series prediction," in *Proc. of NIPS*, 2016, pp. 847–855.
- [7] N. Rao, H.-F. Yu, P. K. Ravikumar, and I. S. Dhillon, "Collaborative filtering with graph information: Consistency and scalable methods," in *Proc. of NIPS*, 2015, pp. 2107–2115.
- [8] L. Anselin, "Under the hood issues in the specification and interpretation of spatial regression models," *Agricultural economics*, vol. 27, no. 3, pp. 247–267, 2002.
- [9] S. R. Dubois and F. H. Glanz, "An autoregressive model approach to two-dimensional shape classification," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, no. 1, pp. 55–66, 1986.
- [10] A. Cichocki, R. Zdunek, A. H. Phan, and S.-i. Amari, *Non-negative matrix and tensor factorizations: applications to exploratory multi-way data analysis and blind source separation*. John Wiley & Sons, 2009.
- [11] Y. Xu, W. Yin, Z. Wen, and Y. Zhang, "An alternating direction algorithm for matrix completion with nonnegative factors," *Frontiers of Mathematics in China*, vol. 7, no. 2, pp. 365–384, 2012.
- [12] J.-F. Cai, E. J. Candès, and Z. Shen, "A singular value thresholding algorithm for matrix completion," *SIAM Journal on Optimization*, vol. 20, no. 4, pp. 1956–1982, 2010.
- [13] Y. Xu and W. Yin, "A globally convergent algorithm for nonconvex optimization based on block coordinate update," *Journal of Scientific Computing*, pp. 1–35, 2017.
- [14] A. Y. Ng, M. I. Jordan, and Y. Weiss, "On spectral clustering: Analysis and an algorithm," in *Proc. of NIPS*, 2002.